

Reading Notes of “Real Analysis” 3rd Edition
by H. L. Royden

Zigang Pan

October 9, 2017

Preface

This is a reading note of the book Royden (1988) and the MATH 441 & 442 notes by Prof. Peter Leob of University of Illinois at Urbana-Champaign. In Chapter 6, I have also included some material from the book Maunday (1996). The Chapters 6–10 include a significant amount of material from the book Luenberger (1969). Chapter 9 also referenced Bartle (1976). Chapter 11 includes significant amount of self-developed material due to lack of reference on this subject. The proof of Radon-Nikodym Theorem 11.169 was adapted from the MATH 442 notes by Prof. Peck of University of Illinois at Urbana-Champaign. The book Royden (1988) does offer some clues as to how to invent the wheel and the book Bartle (1976) is sometimes used to validate the result. The book Spivak (1965) provides the goal of the Fundamental Theorem of Calculus that Chapter 12 is to prove. But the result in Spivak (1965) may not be correct to begin with. I am also going to work out the definition of the surface measure on hyper surfaces in $\mathbb{R}^m$. This material has hints from MATH 442 notes by Prof. Peck. Then, Chapter 12 is complete.

$A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z$
 $A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z$
 $A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z$

Contents

Preface	3
1 Notations	9
2 Set Theory	19
2.1 Axiomatic Foundations of Set Theory	19
2.2 Relations and Equivalence	20
2.3 Function	21
2.4 Set Operations	22
2.5 Algebra of Sets	23
2.6 Partial Ordering and Total Ordering	24
2.7 Basic Principles	26
3 Topological Spaces	35
3.1 Fundamental Notions	35
3.2 Continuity	38
3.3 Basis and Countability	41
3.4 Products of Topological Spaces	42
3.5 The Separation Axioms	47
3.6 Category Theory	48
3.7 Connectedness	50
3.8 Continuous Real-Valued Functions	55
3.9 Nets and Convergence	59
4 Metric Spaces	73
4.1 Fundamental Notions	73
4.2 Convergence and Completeness	75
4.3 Uniform Continuity and Uniformity	78
4.4 Product Metric Spaces	80
4.5 Subspaces	85
4.6 Baire Category	86
4.7 Completion of Metric Spaces	87
4.8 Metrization of Topological Spaces	92
4.9 Interchange Limits	93

5	Compact and Locally Compact Spaces	97
5.1	Compact Spaces	97
5.2	Countable and Sequential Compactness	103
5.3	Real-Valued Functions and Compactness	105
5.4	Compactness in Metric Spaces	107
5.5	The Ascoli-Arzelá Theorem	109
5.6	Product Spaces	111
5.7	Locally Compact Spaces	114
5.7.1	Fundamental notion	114
5.7.2	Partition of unity	117
5.7.3	The Alexandroff one-point compactification	120
5.7.4	Proper functions	121
5.8	σ -Compact Spaces	123
5.9	Paracompact Spaces	124
5.10	The Stone-Čech Compactification	128
6	Vector Spaces	131
6.1	Group	131
6.2	Ring	133
6.3	Field	134
6.4	Vector Spaces	134
6.5	Product Spaces	136
6.6	Subspaces	137
6.7	Convex Sets	141
6.8	Linear Independence and Dimensions	144
7	Banach Spaces	147
7.1	Normed Linear Spaces	147
7.2	The Natural Metric	153
7.3	Product Spaces	156
7.4	Banach Spaces	158
7.5	Compactness	163
7.6	Quotient Spaces	165
7.7	The Stone-Weierstrass Theorem	167
7.8	Linear Operators	175
7.9	Dual Spaces	179
7.9.1	Basic concepts	179
7.9.2	Duals of some common Banach spaces	180
7.9.3	Extension form of Hahn-Banach Theorem	184
7.9.4	Second dual space	192
7.9.5	Alignment and orthogonal complements	194
7.10	The Open Mapping Theorem	199
7.11	The Adjoints of Linear Operators	203
7.12	Weak Topology	206

8	Global Theory of Optimization	217
8.1	Hyperplanes and Convex Sets	217
8.2	Geometric Form of Hahn-Banach Theorem	220
8.3	Duality in Minimum Norm Problems	222
8.4	Convex and Concave Functionals	225
8.5	Conjugate Convex Functionals	229
8.6	Fenchel Duality Theorem	236
8.7	Positive Cones and Convex Mappings	243
8.8	Lagrange Multipliers	245
9	Differentiation in Banach Spaces	253
9.1	Fundamental Notion	253
9.2	The Derivatives of Some Common Functions	257
9.3	Chain Rule and Mean Value Theorem	261
9.4	Higher Order Derivatives	267
9.4.1	Basic concept	267
9.4.2	Interchange order of differentiation	273
9.4.3	High order derivatives of some common functions	278
9.4.4	Properties of high order derivatives	282
9.5	Mapping Theorems	288
9.6	Global Inverse Function Theorem	300
9.7	Interchange Differentiation and Limit	307
9.8	Tensor Algebra	311
9.9	Analytic Functions	313
10	Local Theory of Optimization	325
10.1	Basic Notion	325
10.2	Unconstrained Optimization	331
10.3	Optimization with Equality Constraints	334
10.4	Inequality Constraints	339
11	General Measure and Integration	349
11.1	Measure Spaces	349
11.2	Outer Measure and the Extension Theorem	355
11.3	Measurable Functions	368
11.4	Integration	383
11.5	General Convergence Theorems	398
11.6	Banach Space Valued Measures	414
11.7	Calculation With Measures	446
11.8	The Radon-Nikodym Theorem	477
11.9	L_p Spaces	493
11.10	Dual of $\mathcal{C}(\mathcal{X}, \mathcal{Y})$	509

12 Differentiation and Integration	533
12.1 Carathéodory Extension Theorem	533
12.2 Change of Variable	539
12.3 Product Measure	544
12.4 Functions of Bounded Variation	571
12.5 Absolute and Lipschitz Continuity	595
12.6 Fundamental Theorem of Calculus	612
12.7 Surface measures	634
12.8 Representation of $(\mathcal{C}_k(\Omega, \mathcal{Y}))^*$	634
13 Hilbert Spaces	637
13.1 Fundamental Notions	637
13.2 Projection Theorems	641
13.3 Dual of Hilbert Spaces	643
13.4 Hermitian Adjoints	645
13.5 Approximation in Hilbert Spaces	647
13.6 Other Minimum Norm Problems	655
14 Probability Theory	659
14.1 Fundamental Notions	659
15 Numerical Methods	661
15.1 Newton's Method	661
A Elements in Calculus	665
A.1 Some Formulas	665
A.2 Convergence of Infinite Sequences	665
Bibliography	667
Index	668

Chapter 1

Notations

$\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$	the sets of natural number, integers, and rational numbers, respectively
$\mathbb{R}$ and $\mathbb{C}$	the sets of real numbers and complex numbers, respectively
$\mathbb{K}$	either $\mathbb{R}$ or $\mathbb{C}$
$\mathbb{Z}_+$, $\mathbb{Z}_-$, $\mathbb{C}_+$, $\mathbb{C}_-$	$\mathbb{N} \cup \{0\}$, $\mathbb{Z} \setminus \mathbb{N}$, the open right half of the complex plane, the open left half of the complex plane, respectively
$\in$	belong to
$\notin$	not belong to
$\subseteq$	contained in
$\supseteq$	contains
$\subset$	strict subset of
$\supset$	strict super set of
$\forall$	for all
$\exists$	exists
$\exists!$	exists a unique
$\because$	because
$\therefore$	therefore
$\ni$	such that
$(x_n)_{n=1}^\infty$	the sequence $x_1, x_2, \dots$
$(x_\alpha)_{\alpha \in \Lambda}$	the ordered collection
id_A	the identity map on a set A
$ \lambda $	the absolute value of a real or complex number λ
$\overline{\lambda}$	the complex conjugate of a complex number λ
$\text{Re}(\lambda)$	the real part of a complex number λ
$\text{Im}(\lambda)$	the imaginary part of a complex number λ
$a \vee b$	the maximum of two real numbers a and b
$a \wedge b$	the minimum of two real numbers a and b
$\emptyset$	the empty set; See Page 19.

$\{x, y\}$	an unordered pair; See Page 19.
X_2	the collection of all subsets of X ; See Page 19.
$\cup$	the set union; See Page 19.
(x, y)	an ordered pair; See Page 20.
$X \times Y$	the Cartesian or direct product of sets X and Y ; See Page 20.
$A = \{x \in B \mid P(x)\}$	Definition of a set A ; See Page 20.
$x \sim y$	x and y are related in a relation; See Page 20.
$X/\equiv$	the quotient of the set X with respect to an equivalence relation $\equiv$; See Page 20.
$f : X \rightarrow Y$	a function of X to Y ; $\{(x, f(x)) \in X \times Y \mid \forall x \in X\}$ is the graph of f ; See Page 21.
$\text{graph}(f)$	the graph of a function f ; See Page 21.
$\text{dom}(f)$	the domain of f ; See Page 21.
$f(A)$	the image of $A \subseteq X$ under f ; See Page 21.
$\text{range}(f)$	the range of f , equals to $f(X)$; See Page 21.
$f_{\text{inv}}(A)$	the preimage of $A \subseteq Y$ under f ; See Page 21.
onto, surjective	$f(X) = Y$; See Page 21.
1-1, injective	$f(x_1) \neq f(x_2)$ if $x_1, x_2 \in X$ and $x_1 \neq x_2$; See Page 21.
bijjective	both surjective and injective; See Page 21.
f_{inv}	the inverse function of f ; See Page 21.
$g \circ f$	the composition of $g : Y \rightarrow Z$ with $f : X \rightarrow Y$; See Page 21.
$f _A$	the restriction of f to A ; See Page 21.
Y^X	the set of all functions of X to Y ; See Page 21.
$\cap$	the set intersection; See Page 22.
$\tilde{A}$	the compliment of a set A , where the whole set is clear from context; See Page 22.
$\setminus$	set minus; See Page 22.
$A \triangle B$	the symmetric difference of A and B , equals to $(A \setminus B) \cup (B \setminus A)$; See Page 22.
$\text{card}(X)$	the number of elements in the finite set X ; See Page 23.
$\prod_{\alpha \in \Lambda} X_\alpha$	the Cartesian or direct product of X_α 's; See Page 33.
$\pi_\alpha(x)$	the projection of an element in a Cartesian product space to one of the coordinates; See Page 33.
$\overline{A}$	the closure of a set A , where the whole set is clear from context; See Page 35.
A°	the interior of a set A , where the whole set is clear from context; See Page 35.
∂A	the boundary of a set A , where the whole set is clear from context; See Page 35.
$\prod_{\alpha \in \Lambda} (X_\alpha, \mathcal{O}_\alpha)$	the product topological space; See Page 42.

$\mathbf{T}_1$	Tychonoff space; See Page 47.
$\mathbf{T}_2$	Hausdorff space; See Page 47.
$\mathbf{T}_3$	regular space; See Page 47.
$\mathbf{T}_4$	normal space; See Page 47.
$\mathbf{T}_{3\frac{1}{2}}$	completely regular space; See Page 58.
$(x_\alpha)_{\alpha \in \mathcal{A}}$	a net; See Page 59.
$\lim_{\alpha \in \mathcal{A}} x_\alpha$	the limit of a net; See Page 59.
$\lim_{x \rightarrow x_0} f(x)$	the limit of $f(x)$ as $x \rightarrow x_0$; See Page 63.
$\mathbb{R}_e$	the set of extended real numbers, which equals to $\mathbb{R} \cup \{-\infty, +\infty\}$; See Page 66.
$\limsup_{\alpha \in \mathcal{A}} x_\alpha$	the limit superior of a real-valued net; See Page 67.
$\liminf_{\alpha \in \mathcal{A}} x_\alpha$	the limit inferior of a real-valued net; See Page 67.
$\limsup_{x \rightarrow x_0} f(x)$	the limit superior of $f(x)$ as $x \rightarrow x_0$; See Page 68.
$\liminf_{x \rightarrow x_0} f(x)$	the limit inferior of $f(x)$ as $x \rightarrow x_0$; See Page 68.
$\mathcal{B}_{\mathcal{X}}(x_0, r)$	the open ball centered at x_0 with radius r ; See Page 73.
$\overline{\mathcal{B}}_{\mathcal{X}}(x_0, r)$	the closed ball centered at x_0 with radius r ; See Page 73.
$\text{dist}(x_0, S)$	the distance from a point x_0 to a set S in a metric space; See Page 75.
$(X, \rho_X) \times (Y, \rho_Y)$	the finite product metric space; See Page 80.
$(\prod_{i=1}^{\infty} X_i, \rho)$	the countably infinite product metric space; See Page 83.
$\text{supp } f$	the support of a function; See Page 117.
$\beta(\mathcal{X})$	the Stone-Čech compactification of a completely regular topological space $\mathcal{X}$; See Page 128.
$(\mathcal{M}(A, \mathcal{Y}), \mathcal{F})$	the vector space of $\mathcal{Y}$ -valued functions of a set A over the field $\mathcal{F}$; See Page 136.
$\vartheta_{\mathcal{X}}$	the null vector of a vector space $\mathcal{X}$; See Page 138.
$\mathcal{N}(A)$	the null space of a linear operator A ; See Page 138.
$\mathcal{R}(A)$	the range space of a linear operator A ; See Page 138.
αS	the scalar multiplication by α of a set S in a vector space; See Page 139.
$S + T$	the sum of two sets S and T in a vector space; See Page 139.
$\text{span}(A)$	the subspace generated by the set A ; See Page 140.
$v(P)$	the linear variety generated by a nonempty set P ; See Page 140.
$\text{co}(S)$	the convex hull generated by S in a vector space; See Page 141.
$\ x\ $	the norm of a vector x ; See Page 147.
$ x $	the Euclidean norm of a vector x ; See Page 147.
$\mathcal{C}_1([a, b])$	the normed linear space of continuously differentiable real-valued functions on the interval $[a, b]$; See Page 148.

l_p	The normed linear space of real-valued sequences with finite p -norm, $1 \leq p \leq +\infty$; See Page 149.
$l_p(\mathcal{X})$	the normed linear space of $\mathcal{X}$ -valued sequences with finite p -norm, $1 \leq p \leq +\infty$; See Page 153.
$V(P)$	the closed linear variety generated by a nonempty set P ; See Page 156.
$^\circ P$	the relative interior of a set P ; See Page 156.
$\mathcal{X} \times \mathcal{Y}$	the finite Cartesian product normed linear space; See Page 156.
$\mathcal{C}([a, b])$	the Banach space of continuous real-valued functions on the interval $[a, b]$; See Page 160.
$\mathcal{C}(\mathcal{K}, \mathcal{X})$	the normed linear space of continuous $\mathcal{X}$ -valued functions on a compact space $\mathcal{K}$; See Page 160.
$\mathcal{X}_{\mathbb{R}}$	the real normed linear space induced by a complex normed linear space $\mathcal{X}$; See Page 164.
$[x]$	the coset of a vector x in a quotient space; See Page 165.
$\mathcal{X}/M$	the quotient space of a vector space $\mathcal{X}$ modulo a subspace M ; See Page 165.
$\mathcal{X}/M$	the quotient space of a normed linear space $\mathcal{X}$ modulo a closed subspace M ; See Page 166.
$\mathcal{C}_v(\mathcal{X}, \mathcal{Y})$	the vector space of continuous $\mathcal{Y}$ -valued functions on $\mathcal{X}$; See Page 168.
$B(\mathcal{X}, \mathcal{Y})$	the set of bounded linear operators of $\mathcal{X}$ to $\mathcal{Y}$; See Page 175.
$\mathcal{X}^*$	the dual of $\mathcal{X}$; See Page 179.
x_*	a vector in the dual; See Page 179.
$\langle\langle x_*, x \rangle\rangle$	the evaluation of a bounded linear functional x_* at the vector x , that is $x_*(x)$; See Page 179.
$c_0(\mathcal{X})$	the subspace of $l_\infty(\mathcal{X})$ consisting of $\mathcal{X}$ -valued sequences with limit $\vartheta_{\mathcal{X}}$; See Page 183.
$\mathcal{X}^{**}$	the second dual of $\mathcal{X}$; See Page 192.
$S^\perp$	the orthogonal complement of the set S ; See Page 194.
$^\perp S$	the pre-orthogonal complement of the set S ; See Page 194.
A'	the adjoint of a linear operator A ; See Page 203.
A''	the adjoint of the adjoint of a linear operator A ; See Page 204.
$\mathcal{O}_{\text{weak}}(\mathcal{X})$	the weak topology on a normed linear space $\mathcal{X}$; See Page 206.
$\mathcal{X}_{\text{weak}}$	the weak topological space associated with a normed linear space $\mathcal{X}$; See Page 207.
$\mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$	the weak* topology on the dual of a normed linear space $\mathcal{X}$; See Page 209.

$\mathcal{X}_{\text{weak}^*}^*$	the weak* topological space associated with a normed linear space $\mathcal{X}$; See Page 210.
K_{supp}	the support of a convex set K ; See Page 222.
$[f, C]$	the epigraph of a convex function $f : C \rightarrow \mathbb{R}$; See Page 225.
C_{conj}	the conjugate convex set; See Page 229.
f_{conj}	the conjugate convex functional; See Page 229.
$[f, C]_{\text{conj}}$	the epigraph of the conjugate convex functional; See Page 230.
$\text{conj}\Gamma$	the pre-conjugate convex set; See Page 232.
$\text{conj}\varphi$	the pre-conjugate convex functional; See Page 232.
$\text{conj}[\varphi, \Gamma]$	the epigraph of the pre-conjugate convex functional; See Page 232.
$\geq$	greater than or equal to (with respect to the positive cone); See Page 243.
$\leq$	less than or equal to (with respect to the positive cone); See Page 243.
$>$	greater than (with respect to the positive cone); See Page 243.
$<$	less than (with respect to the positive cone); See Page 243.
$S^{\oplus}$	the positive conjugate cone of a set S ; See Page 243.
$S^{\ominus}$	the negative conjugate cone of a set S ; See Page 243.
$A_D(x_0)$	the set of admissible deviations in D at x_0 ; See Page 253.
$f^{(1)}(x_0), Df(x_0)$	the Fréchet derivative of f at x_0 ; See Page 254.
$Df(x_0; u)$	the directional derivative of f at x_0 along u ; See Page 254.
$\frac{\partial f}{\partial y}(x_0, y_0)$	partial derivative of f with respect to y at (x_0, y_0) ; See Page 255.
$\text{ro}(A)(B)$	right operate: $\text{ro}(A)(B) = BA$; See Page 260.
$B_k(\mathcal{X}, \mathcal{Y})$	the set of bounded multi-linear $\mathcal{Y}$ -valued functions on $\mathcal{X}^k$; See Page 267.
$B_{S_k}(\mathcal{X}, \mathcal{Y})$	the set of symmetric bounded multi-linear $\mathcal{Y}$ -valued functions on $\mathcal{X}^k$; See Page 267.
$D^k f(x_0), f^{(k)}(x_0)$	the k th order Fréchet derivative of f at x_0 ; See Page 267.
$\mathcal{C}_k, \mathcal{C}_{\infty}$	k -times and infinite-times continuously differentiable functions, respectively; See Page 268.
$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}$	k th-order partial derivative of f ; See Page 273.
$\text{Sm}(D)$	$= \bigcup_{\alpha \in \mathbf{K}} \alpha D$; See Page 273.

$\mathcal{C}_k(\Omega, \mathcal{Y})$	the normed linear space of k -times continuously differentiable $\mathcal{Y}$ -valued functions on a compact set $\Omega \subseteq \mathcal{X}$; See Page 309.
$\mathcal{C}_b(\mathcal{X}, \mathcal{Y})$	the normed linear space of bounded continuous $\mathcal{Y}$ -valued functions on a topological space $\mathcal{X}$; See Page 309.
$\mathcal{C}_{b,k}(\Omega, \mathcal{Y})$	the normed linear space of k -times bounded continuously differentiable $\mathcal{Y}$ -valued functions on a set $\Omega \subseteq \mathcal{X}$; See Page 310.
$A^{T_{n_1, \dots, n_m}}$	the transpose of an m th order tensor A with the permutation $(n_1, \dots, n_m)$; See Page 311.
$A \otimes B$	the outer product of two tensors; See Page 312.
$\mathbf{0}_{m_1 \times \dots \times m_n}$	an n th order $\mathbb{K}$ -valued tensor in $B(\mathbb{K}^{m_n}, \dots, B(\mathbb{K}^{m_1}, \mathbb{K}) \cdots)$ with all elements equal to 0; See Page 313.
$\mathbf{1}_{m_1 \times \dots \times m_n}$	an n th order $\mathbb{K}$ -valued tensor in $B(\mathbb{K}^{m_n}, \dots, B(\mathbb{K}^{m_1}, \mathbb{K}) \cdots)$ with all elements equal to 1; See Page 313.
$\mathcal{S}_+ \mathcal{X}, \mathcal{S}_{\text{psd}} \mathcal{X}$	sets of positive definite and positive semi-definite operators over the real normed linear space $\mathcal{X}$, respectively; See Page 326.
$\mathcal{S}_- \mathcal{X}, \mathcal{S}_{\text{nsd}} \mathcal{X}$	sets of negative definite and negative semi-definite operators over the real normed linear space $\mathcal{X}$, respectively; See Page 326.
$\mathcal{S}_{\mathcal{X}}$	$B_{S_2}(\mathcal{X}, \mathbb{R})$; See Page 326.
$(\mathbb{R}, \mathcal{B}_L, \mu_L)$	Lebesgue measure space; See Page 360.
μ_{Lo}	Lebesgue outer measure; See Page 361.
$\mathcal{B}_B(\mathcal{X})$	Borel sets; See Page 361.
$\mathcal{A}(\mathcal{X})$	the algebra generated by the topology of a topological space $\mathcal{X}$; See Page 361.
μ_B	the Borel measure on $\mathbb{R}$; See Page 362.
$\mathbb{R}$	$((\mathbb{R}, \mathbb{R}, \cdot), \mathcal{B}_B(\mathbb{R}), \mu_B)$; See Page 366.
P a.e. in $\mathcal{X}$	P holds almost everywhere in $\mathcal{X}$; See Page 370.
$P(x)$ a.e. $x \in \mathcal{X}$	P holds almost everywhere in $\mathcal{X}$; See Page 370.
$\mathcal{P} \circ f$	$\mathcal{P} \circ f : X \rightarrow [0, \infty) \subset \mathbb{R}$ defined by $\mathcal{P} \circ f(x) = \ f(x)\ , \forall x \in X$; See Page 380.
$\mathcal{R}(\mathcal{X})$	the collection of all representation of $\mathcal{X}$; See Page 383.
$\mathcal{I}(\mathcal{X})$	the integration system on $\mathcal{X}$; See Page 383.
$\int_X f \, d\mu$	the integral of a function f on a set X with respect to measure μ ; See Page 384.
$\int_X f(x) \, d\mu(x)$	the integral of a function f on a set X with respect to measure μ ; See Page 384.
$\mathcal{P} \circ \mu$	the total variation of a Banach space valued pre-measure μ ; See Page 414.

$\mathcal{P} \circ \mu$	the total variation of a Banach space valued measure μ ; See Page 419.
$\mu_1 + \mu_2$	the $\mathcal{Y}$ -valued measure that equals to the sum of two $\mathcal{Y}$ -valued measures on the same measurable space; See Page 449.
$\alpha\mu$	the $\mathcal{Y}$ -valued measure that equals to the scalar product of $\alpha \in \mathbb{K}$ and $\mathcal{Y}$ -valued measure μ ; See Page 450.
μy	the $\mathcal{Y}$ -valued measure that equals to scalar product of a $\mathbb{K}$ -valued measure μ and $y \in \mathcal{Y}$; See Page 450.
$A\mu$	the $\mathcal{Z}$ -valued measure that equals to product of an bounded linear operator A and a $\mathcal{Y}$ -valued measure μ ; See Page 450.
$\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$	the vector space of σ -finite $\mathcal{Y}$ -valued measures on the measurable space $(X, \mathcal{B})$; See Page 456.
$\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$	the normed linear space of finite $\mathcal{Y}$ -valued measures on the measurable space $(X, \mathcal{B})$; See Page 457.
$\lim_{n \in \mathbb{N}} \mu_n = \nu$	a sequence of σ -finite ($\mathcal{Y}$ -valued) measures $(\mu_n)_{n=1}^\infty$ converges to a σ -finite ($\mathcal{Y}$ -valued) measure ν ; See Page 460.
$\mu_1 \leq \mu_2$	the measures μ_1 and μ_2 on the measurable space $(X, \mathcal{B})$ can be compared if $\mu_1(E) \leq \mu_2(E)$, $\forall E \in \mathcal{B}$; See Page 460.
$\begin{bmatrix} \mu_{1,1} & \cdots & \mu_{1,m} \\ \vdots & & \vdots \\ \mu_{n,1} & \cdots & \mu_{n,m} \end{bmatrix}$	the vector measure; See Page 463.
$\mu_1 \ll \mu_2$	the measure μ_1 is absolutely continuous with respect to the measure μ_2 ; See Page 473.
$\mu_1 \perp \mu_2$	the measures μ_1 and μ_2 are mutually singular; See Page 473.
$\frac{d\nu}{d\mu}$	the Radon-Nikodym derivative of the σ -finite $\mathcal{Y}$ -valued measure ν with respect to the σ -finite $\mathbb{K}$ -valued measure μ ; See Page 483.
$\mathcal{P}_p \circ f$	the function $\ f(\cdot)\ ^p$; See Page 493.
ess sup	the essential supremum; See Page 495.
$\lim_{n \in \mathbb{N}} z_n \doteq z$ in $\bar{\mathcal{L}}_p$	the sequence $(z_n)_{n=1}^\infty \subseteq \bar{\mathcal{L}}_p$ converges to $z \in \bar{\mathcal{L}}_p$ in $\bar{\mathcal{L}}_p$ pseudo-norm; See Page 497.
$\mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$	the normed linear space of finite $\mathcal{Y}$ -valued topological measures on $\mathcal{X}$; See Page 514.
$\mathcal{M}_{\sigma t}(\mathcal{X}, \mathcal{Y})$	the vector space of σ -finite $\mathcal{Y}$ -valued topological measures on $\mathcal{X}$; See Page 514.
$\mathcal{M}_\sigma(X, \mathcal{B})$	the set of σ -finite measures on the measurable space $(X, \mathcal{B})$; See Page 515.
$\mathcal{M}_f(X, \mathcal{B})$	the set of finite measures on the measurable space $(X, \mathcal{B})$; See Page 515.

$\mathcal{M}_{\sigma t}(\mathcal{X})$	the set of σ -finite topological measures on the topological space $\mathcal{X}$; See Page 515.
$\mathcal{M}_{ft}(\mathcal{X})$	the set of finite topological measures on the topological space $\mathcal{X}$; See Page 515.
$\prod_{j=1}^m \mu_j$	the product measure of $\mu_1, \dots, \mu_m$; See Page 547.
$\prod_{i=1}^m \mathbb{X}_i$	product topological measure space of $\mathbb{X}_1, \dots, \mathbb{X}_m$, where $m \in \mathbb{N}$; See Page 564.
$\mathbf{r}_{x_1, x_2}$	the semi-open rectangle in $\mathbb{R}^m$ with corners x_1 and x_2 ; See Page 572.
$\bar{\mathbf{r}}_{x_1, x_2}$	the closed rectangle in $\mathbb{R}^m$ with corners x_1 and x_2 ; See Page 573.
$\mathbf{r}_{x_1, x_2}^\circ$	the open rectangle in $\mathbb{R}^m$ with corners x_1 and x_2 ; See Page 573.
$P(\Omega)$	the principal of a region Ω ; See Page 573.
$\text{VRect}_{i, x_1, x_2}$	the set of vertexes of $\mathbf{r}_{x_1, x_2}$ with i coordinates equal to that of x_1 ; See Page 575.
$\Delta_F(\mathbf{r}_{x_1, x_2})$	the increment of F on $\mathbf{r}_{x_1, x_2}$; See Page 575.
$T_F(\mathbf{r}_{x_1, x_2})$	the total variation of a function F on the semi-open rectangle $\mathbf{r}_{x_1, x_2}$; See Page 575.
$(\mathbb{R}^m, \mathcal{B}_{Lm}, \mu_{Lm}), \mu_{Lmo}$	m -dimensional Lebesgue measure space and the m -dimensional Lebesgue outer measure; See Page 594.
$\int_U g(x) dF(x)$	the integral of function g with respect to $\mathcal{Y}$ -valued measure space $(P(\Omega), \mathcal{B}_B(P(\Omega)), \mu)$ whose distribution function is $F : \Omega \rightarrow \mathcal{Y}$ over the set $U \in P(\Omega)$; See Page 595.
$\pi_{\hat{i}_0} : \mathbb{R}^m \rightarrow \mathbb{R}^{m-1}$	$\pi_{\hat{i}_0}(x) = (\pi_1(x), \dots, \pi_{i_0-1}(x), \pi_{i_0+1}(x), \dots, \pi_m(x))$, $\forall x \in \mathbb{R}^m$; See Page 595.
$\int_a^b f(x) dx$	the integral of function f from $a \in \mathbb{R}$ to $b \in \mathbb{R}$ with respect to μ_B ; See Page 606.
$\text{dia}(S)$	the diameter of a subset of a metric space; See Page 613.
$e_{m,i}$	the i th unit vector in $\mathbb{R}^m$; See Page 616.
$\langle x, y \rangle$	the inner product of vectors x and y ; See Page 637.
$x \perp y$	the vectors x and y are orthogonal in a pre-Hilbert space; See Page 639.
$x \perp S$	the vector x is orthogonal to the set S in a pre-Hilbert space; See Page 639.
x^*	$x^* \in \mathcal{X}^*$ that satisfies $\langle x^*, y \rangle = \langle y, x \rangle$, $\forall y \in \mathcal{X}$; See Page 643.
$M \oplus N$	the direct sum of M and N ; See Page 645.
A^*	the Hermitian adjoint of A ; See Page 645.
$\text{Gram}(y_1, \dots, y_n)$	the Gram matrix of $y_1, \dots, y_n$; See Page 648.
$\text{gram}(y_1, \dots, y_n)$	the Gram determinant of $y_1, \dots, y_n$; See Page 648.
$E(x)$	expectation of x ; See Page 659.

$E(x|\hat{\mathcal{B}})$ conditional expectation of x given $\hat{\mathcal{B}}$; See Page 659.

Chapter 2

Set Theory

2.1 Axiomatic Foundations of Set Theory

We will list the nine axioms of ZFC axiom system. The ninth axiom, which is the Axiom of Choice, will be introduced in Section 2.7. Let A and B be sets and x and y be objects (which is another name for sets).

Axiom 1 (Axiom of Extensionality) $A = B$ if $\forall x \in A$, we have $x \in B$; and $\forall x \in B$, we have $x \in A$.

Axiom 2 (Axiom of Empty Set) There exists an empty set $\emptyset$, which does not contain any element.

Axiom 3 (Axiom of Pairing) For any objects x and y , there exists a set $\{x, y\}$, which contains only x and y .

Axiom 4 (Axiom of Regularity) Any nonempty set $A \neq \emptyset$, there exists $a \in A$, such that $\forall b \in A$, we have $b \notin a$.

Axiom 5 (Axiom of Replacement) $\forall x \in A$, let there be one and only one y to form an ordered pair (x, y) . Then, the collection of all such y 's is a set B .

Axiom 6 (Axiom of Power Set) The collection of all subsets of A is a set denoted by 2^A .

Axiom 7 (Axiom of Union) For any collection of sets $(A_\lambda)_{\lambda \in \Lambda}$, where Λ is a set, then $\bigcup_{\lambda \in \Lambda} A_\lambda$ is a well defined set.

Axiom 8 (Axiom of Infinity) There exists a set A such that $\emptyset \in A$ and $\forall x \in A$, we have $\{\emptyset, x\} \in A$.

By Axiom 2, there exists the empty set $\emptyset$, which we may call 0. Now, by Axiom 3, there exists the set $\{\emptyset\}$, which is nonempty and we may call 1. Again, by Axiom 3, there exists the set $\{\emptyset, \{\emptyset\}\}$, which we will call 2. After we define n , we may define $n + 1 := \{0, n\}$, which exists by Axiom of Pairing. This allows us to define all natural numbers. By Axiom 8, these natural numbers can form the set, $\mathbb{N} := \{1, 2, \dots\}$, which is the set of natural numbers. Furthermore, by Axiom 6, we may define the set of all real numbers, $\mathbb{R}$.

For any $x \in A$ and $y \in B$, we may apply Axiom 3 to define the ordered pair $(x, y) := \{\{\{x\}, 1\}, \{\{y\}, 2\}\}$. Then, the set $A \times B$ is defined by $\bigcup_{x \in A} \bigcup_{y \in B} \{(x, y)\}$, which is a valid set by Axiom 7. By Axiom 5, any portion A_s of a well-defined set A is again a set, which is called a *subset* of A , we will write $A_s \subseteq A$. Thus, the formula

$$\{x \in A \mid p(x) \text{ is true.}\}$$

defines a set as long as A is a set and $p(x)$ is unambiguous logic expression.

2.2 Relations and Equivalence

Definition 2.1 Let A and B be sets. A relation R from A to B is a subset of $A \times B$. $\forall x \in A, \forall y \in B$, we say $x \sim y$ if $(x, y) \in R$. We will say that R is a relation on A if it is a relation from A to A . We define

$$\begin{aligned} \text{dom}(R) &:= \{x \in A \mid \exists y \in B, \text{ such that } x \sim y\} \\ \text{range}(R) &:= \{y \in B \mid \exists x \in A, \text{ such that } x \sim y\} \end{aligned}$$

which are well-defined subsets.

Definition 2.2 Let A be a set and R be a relation on A . $\forall x, y, z \in A$,

1. R is reflexive if $x \sim x$.
2. R is symmetric if $x \sim y$ implies $y \sim x$.
3. R is transitive if $x \sim y$ and $y \sim z$ implies $x \sim z$.
4. R is a equivalence relationship if it is reflexive, symmetric, and transitive, and will be denote by " $\equiv$ ".
5. R is antisymmetric if $x \sim y$ and $y \sim x$ implies $x = y$.

Let $\equiv$ be an equivalence relationship on A , then, it partitions A into disjoint equivalence classes $A_x := \{y \in A \mid x \equiv y\}, \forall x \in A$. The collection of all equivalence classes, $A/\equiv := \{A_x \subseteq A \mid x \in A\}$, is called the *quotient* of A with respect to $\equiv$.

2.3 Function

Definition 2.3 Let X and Y be sets and $D \subseteq X$. A function f of D to Y , denoted by $f : D \rightarrow Y$, is a relation from D to Y such that $\forall x \in D$, there is exactly one $y \in Y$, such that $(x, y) \in f$; we will denote that y as $f(x)$. The graph of f is the set $\text{graph}(f) := \{(x, f(x)) \in X \times Y \mid x \in D\}$. The domain of f is $\text{dom}(f) = D$. $\forall A \subseteq X$, the image under f of A is $f(A) := \{y \in Y \mid \exists x \in A \cap D \text{ such that } f(x) = y\}$, which is a subset of Y . The range of f is $\text{range}(f) = f(X)$. $\forall B \subseteq Y$, the inverse image under f of B is $f_{\text{inv}}(B) := \{x \in D \mid f(x) \in B\}$, which is a subset of D . f is said to be surjective if $f(X) = Y$; and f is said to be injective if $f(x_1) \neq f(x_2)$, $\forall x_1, x_2 \in D$ with $x_1 \neq x_2$; f is said to be bijective if it is both surjective and injective, in which case it is invertible and the inverse function is denoted by $f_{\text{inv}} : Y \rightarrow D$. We will say that f is a function from X to Y .

Let $f : D \rightarrow Y$ and $g : Y \rightarrow Z$ be functions, we may define a function $h : D \rightarrow Z$ by $h(x) = g(f(x))$, then h is called the *composition* of g with f , and is denoted by $g \circ f$. Let $A \subseteq X$. We may define a function $l : A \cap D \rightarrow Y$ by $l(x) = f(x)$, $\forall x \in A \cap D$. This function is called the *restriction* of f to A , and denoted by $f|_A$. Let $f : D \rightarrow Y$, $g : Y \rightarrow Z$, and $h : Z \rightarrow W$, we have $(h \circ g) \circ f = h \circ (g \circ f)$. Let $f : D \rightarrow D$ and $k \in \mathbb{Z}_+$, we will write $f^k := \underbrace{f \circ \dots \circ f}_k$, where $f^0 := \text{id}_D$.

A function $f : X \rightarrow Y$ is a subset of $X \times Y$. Then, $f \in {}^{X \times Y}2$. The collection of all functions of X to Y is then a set given by

$$Y^X := \{f \in {}^{X \times Y}2 \mid \forall x \in X, \exists! y \in Y \ni (x, y) \in f\}$$

We have the following result concerning the inverse of a function.

Proposition 2.4 Let $\phi : X \rightarrow Y$, where X and Y are sets. Then, ϕ is bijective if, and only if, $\exists \psi_i : Y \rightarrow X$, $i = 1, 2$, such that $\phi \circ \psi_1 = \text{id}_Y$ and $\psi_2 \circ \phi = \text{id}_X$. Furthermore, $\phi_{\text{inv}} = \psi_1 = \psi_2$.

Proof “Sufficiency” Let $\psi_i : Y \rightarrow X$, $i = 1, 2$, exist. $\forall y \in Y$, $\phi \circ \psi_1(y) = \text{id}_Y(y) = y$, which implies that $y \in \text{range}(\phi)$, and hence, ϕ is surjective. Suppose that ϕ is not injective, then $\exists x_1, x_2 \in X$ with $x_1 \neq x_2$ such that $\phi(x_1) = \phi(x_2)$. Then, we have

$$x_1 = \text{id}_X(x_1) = \psi_2(\phi(x_1)) = \psi_2(\phi(x_2)) = \text{id}_X(x_2) = x_2$$

which is a contradiction. Hence, ϕ is injective. This proves that ϕ is bijective.

“Necessity” Let ϕ be bijective. Then, $\phi_{\text{inv}} : Y \rightarrow X$ exists. $\forall x \in X$, let $y = \phi(x)$, then $x = \phi_{\text{inv}}(y)$, hence $\phi_{\text{inv}}(\phi(x)) = x$. Therefore, we have $\phi_{\text{inv}} \circ \phi = \text{id}_X$. $\forall y \in Y$, let $x = \phi_{\text{inv}}(y)$, then $y = \phi(x)$, hence $\phi(\phi_{\text{inv}}(y)) = y$. Therefore, we have $\phi \circ \phi_{\text{inv}} = \text{id}_Y$. Hence, $\psi_1 = \psi_2 = \phi_{\text{inv}}$.

Let ψ_1 and ψ_2 satisfy the assumption of the proposition, and ϕ_{inv} be the inverse function of ϕ . Then, we have

$$\begin{aligned}\psi_1 &= \text{id}_X \circ \psi_1 = (\phi_{\text{inv}} \circ \phi) \circ \psi_1 = \phi_{\text{inv}} \circ (\phi \circ \psi_1) = \phi_{\text{inv}} \circ \text{id}_Y = \phi_{\text{inv}} \\ \psi_2 &= \psi_2 \circ \text{id}_Y = \psi_2 \circ (\phi \circ \phi_{\text{inv}}) = (\psi_2 \circ \phi) \circ \phi_{\text{inv}} = \text{id}_X \circ \phi_{\text{inv}} = \phi_{\text{inv}}\end{aligned}$$

This completes the proof of the proposition. $\square$

For bijective functions $f : X \rightarrow Y$ and $g : Y \rightarrow Z$, $g \circ f$ is also bijective and $(g \circ f)_{\text{inv}} = f_{\text{inv}} \circ g_{\text{inv}}$.

2.4 Set Operations

Let X be a set, $^X 2$ is the set consisting of all subsets of X . $\forall A, B \subseteq X$, we will define

$$\begin{aligned}A \cup B &:= \{x \in X \mid x \in A \text{ or } x \in B\} \\ A \cap B &:= \{x \in X \mid x \in A \text{ and } x \in B\} \\ \tilde{A} &:= \{x \in X \mid x \notin A\} \\ A \setminus B &:= \{x \in A \mid x \notin B\} = A \cap \tilde{B} \\ A \triangle B &:= (A \setminus B) \cup (B \setminus A)\end{aligned}$$

We have the following results.

Proposition 2.5 *Let $A, B, D, A_\lambda \in {}^X 2$, $f : D \rightarrow Y$, $C, E, C_\lambda \in {}^Y 2$, where X and Y are sets, $\lambda \in \Lambda$, and Λ is an index set. Then, we have*

1. $A \cup B = B \cup A$ and $A \cap B = B \cap A$;
2. $A \subseteq A \cup B$ and $A = A \cup B$ if, and only if, $B \subseteq A$;
3. $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$, $A \cup X = X$, and $A \cap X = A$;
4. $\tilde{\emptyset} = X$, $\tilde{\tilde{A}} = A$, $A \cup \tilde{A} = X$, $A \cap \tilde{A} = \emptyset$, and $A \subseteq B$ if, and only if, $\tilde{B} \subseteq \tilde{A}$;
5. The De Morgan's Laws:

$$\left(\bigcup_{\lambda \in \Lambda} A_\lambda\right)^\sim = \bigcap_{\lambda \in \Lambda} \tilde{A}_\lambda; \quad \left(\bigcap_{\lambda \in \Lambda} A_\lambda\right)^\sim = \bigcup_{\lambda \in \Lambda} \tilde{A}_\lambda$$

$$6. B \cup \left(\bigcap_{\lambda \in \Lambda} A_\lambda\right) = \bigcap_{\lambda \in \Lambda} (B \cup A_\lambda) \text{ and } B \cap \left(\bigcup_{\lambda \in \Lambda} A_\lambda\right) = \bigcup_{\lambda \in \Lambda} (B \cap A_\lambda);$$

$$7. f\left(\bigcup_{\lambda \in \Lambda} A_\lambda\right) = \bigcup_{\lambda \in \Lambda} f(A_\lambda) \text{ and } f\left(\bigcap_{\lambda \in \Lambda} A_\lambda\right) \subseteq \bigcap_{\lambda \in \Lambda} f(A_\lambda);$$

8. $f_{\text{inv}}\left(\bigcup_{\lambda \in \Lambda} C_\lambda\right) = \bigcup_{\lambda \in \Lambda} f_{\text{inv}}(C_\lambda)$ and $f_{\text{inv}}\left(\bigcap_{\lambda \in \Lambda} C_\lambda\right) = \bigcap_{\lambda \in \Lambda} f_{\text{inv}}(C_\lambda)$;
9. $f_{\text{inv}}(C \setminus E) = f_{\text{inv}}(C) \setminus f_{\text{inv}}(E)$, $f(f_{\text{inv}}(C)) = C \cap \text{range}(f)$, and $f_{\text{inv}}(f(A)) \supseteq A \cap \text{dom}(f) = A \cap D$.

The proof of the above results are standard and is therefore omitted.

2.5 Algebra of Sets

Definition 2.6 A set X is said to be finite if it is either empty or the range of a function of $\{1, 2, \dots, n\}$, with $n \in \mathbb{N}$. In this case, $\text{card}(X)$ denotes the number of elements in X . It is said to be countable if it is either empty or the range of a function of $\mathbb{N}$.

Definition 2.7 Let X be a set and $\mathcal{A} \subseteq {}^X 2$. $\mathcal{A}$ is said to be an algebra of sets on X (or a Boolean algebra on X) if

- (i) $\emptyset, X \in \mathcal{A}$;
- (ii) $\forall A, B \in \mathcal{A}, A \cup B \in \mathcal{A}$ and $\tilde{A} \in \mathcal{A}$.

$\mathcal{A}$ is said to be a σ -algebra on X if it is an algebra on X and countable unions of sets in $\mathcal{A}$ is again in $\mathcal{A}$.

Let $\mathcal{M} \subseteq {}^X 2$, where X is a set, then, there exists a smallest algebra on X , $\mathcal{A}_0 \subseteq {}^X 2$, containing $\mathcal{M}$, which means that any algebra on X , $\mathcal{A}_1 \subseteq {}^X 2$, that contains $\mathcal{M}$, we have $\mathcal{A}_0 \subseteq \mathcal{A}_1$. This algebra $\mathcal{A}_0$ is said to be the algebra on X generated by $\mathcal{M}$. Also, there exists a smallest σ -algebra on X , $\mathcal{A} \subseteq {}^X 2$, containing $\mathcal{M}$, which is said to be the σ -algebra on X generated by $\mathcal{M}$.

Proposition 2.8 Let X be a set, $\mathcal{E}$ be a nonempty collection of subsets of X , $\mathcal{A}$ be the algebra on X generated by $\mathcal{E}$, and

$$\begin{aligned} \bar{\mathcal{A}} := & \left\{ A \subseteq X \mid \exists n, m \in \mathbb{N}, \forall i_1, \dots, i_{2n} \in \{1, \dots, m\}, \exists F_{i_1, \dots, i_{2n}} \subseteq X \right. \\ & \text{with } F_{i_1, \dots, i_{2n}} \in \mathcal{E} \text{ or } (F_{i_1, \dots, i_{2n}})^\sim \in \mathcal{E}, \text{ such that} \\ & \left. A = \bigcup_{i_1=1}^m \bigcap_{i_2=1}^m \cdots \bigcap_{i_{2n}=1}^m F_{i_1, \dots, i_{2n}} \right\} \end{aligned}$$

Then, $\mathcal{A} = \bar{\mathcal{A}}$.

Proof $\forall E \in \mathcal{E}$, let $n = 1$, $m = 1$, and $F_{1,1} = E$. Then, $E = \bigcup_{i_1=1}^1 \bigcap_{i_2=1}^1 F_{i_1, i_2} \in \bar{\mathcal{A}}$. Hence, we have $\mathcal{E} \subseteq \bar{\mathcal{A}}$. It is clear that $\bar{\mathcal{A}} \subseteq \mathcal{A}$. All we need to show is that $\bar{\mathcal{A}}$ is an algebra on X . Then, $\mathcal{A} \subseteq \bar{\mathcal{A}}$ and the result follows.

Fix $E \in \mathcal{E} \neq \emptyset$. $E \subseteq X$. Let $n = 1$, $m = 2$, $F_{1,1} = E$, $F_{1,2} = \tilde{E}$, $F_{2,1} = E$, and $F_{2,2} = \tilde{E}$. Then, $\emptyset = \bigcup_{i_1=1}^2 \bigcap_{i_2=1}^2 F_{i_1,i_2} \in \bar{\mathcal{A}}$. Let $n = 1$, $m = 2$, $F_{1,1} = E$, $F_{1,2} = E$, $F_{2,1} = \tilde{E}$, and $F_{2,2} = \tilde{E}$. Then, $X = \bigcup_{i_1=1}^2 \bigcap_{i_2=1}^2 F_{i_1,i_2} \in \bar{\mathcal{A}}$.

$\forall A, B \in \bar{\mathcal{A}}, \exists n_A, m_A \in \mathbb{N}, \forall i_1, \dots, i_{2n_A} \in \{1, \dots, m_A\}, \exists F_{i_1, \dots, i_{2n_A}}^{(A)} \subseteq X$ with $F_{i_1, \dots, i_{2n_A}}^{(A)} \in \mathcal{E}$ or $\left(F_{i_1, \dots, i_{2n_A}}^{(A)}\right)^\sim \in \mathcal{E}$ such that $A = \bigcup_{i_1=1}^{m_A} \bigcap_{i_2=1}^{m_A} \dots \bigcap_{i_{2n_A}=1}^{m_A} F_{i_1, \dots, i_{2n_A}}^{(A)}$, and $\exists n_B, m_B \in \mathbb{N}, \forall i_1, \dots, i_{2n_B} \in \{1, \dots, m_B\}, \exists F_{i_1, \dots, i_{2n_B}}^{(B)} \subseteq X$ with $F_{i_1, \dots, i_{2n_B}}^{(B)} \in \mathcal{E}$ or $\left(F_{i_1, \dots, i_{2n_B}}^{(B)}\right)^\sim \in \mathcal{E}$ such that $B = \bigcup_{i_1=1}^{m_B} \bigcap_{i_2=1}^{m_B} \dots \bigcap_{i_{2n_B}=1}^{m_B} F_{i_1, \dots, i_{2n_B}}^{(B)}$.

Note that $\tilde{A} = \bigcap_{i_1=1}^{m_A} \bigcup_{i_2=1}^{m_A} \dots \bigcup_{i_{2n_A}=1}^{m_A} \left(F_{i_1, \dots, i_{2n_A}}^{(A)}\right)^\sim$. Let $n = n_A + 1$, $m = m_A$, $\forall i_1, \dots, i_{2n} \in \{1, \dots, m\}$, $G_{i_1, \dots, i_{2n}} = \left(F_{i_2, \dots, i_{2n-1}}^{(A)}\right)^\sim$. Then, $\tilde{A} = \bigcup_{i_1=1}^m \bigcap_{i_2=1}^m \dots \bigcap_{i_{2n}=1}^m G_{i_1, \dots, i_{2n}} \in \bar{\mathcal{A}}$.

Without loss of generality, assume $n_A \geq n_B$. Let $n = n_A$ and $m = m_A + m_B$. Define $\tilde{i} = 1 + (i \bmod m_A)$ and $\tilde{i} = 1 + (i \bmod m_B)$, $\forall i \in \mathbb{N}$. $\forall i_1, \dots, i_{2n} \in \{1, \dots, m\}$, let

$$G_{i_1, \dots, i_{2n}} = \begin{cases} F_{i_1, \tilde{i}_2, \dots, \tilde{i}_{2n_A}}^{(A)} & \text{if } i_1 \leq m_A \\ F_{i_1 - m_A, \tilde{i}_2, \dots, \tilde{i}_{2n_B}}^{(B)} & \text{if } i_1 > m_A \end{cases}$$

Then, it is easy to check that $A \cup B = \bigcup_{i_1=1}^m \bigcap_{i_2=1}^m \dots \bigcap_{i_{2n}=1}^m G_{i_1, \dots, i_{2n}} \in \bar{\mathcal{A}}$. Hence, $\bar{\mathcal{A}}$ is an algebra on X .

This completes the proof of the proposition. $\square$

2.6 Partial Ordering and Total Ordering

Definition 2.9 Let A be a set and $\preceq$ be a relation on A . $\preceq$ will be called a partial ordering if it is reflexive and transitive. It will be called a total ordering if it is an antisymmetric partial ordering and satisfies $\forall x, y \in A$ with $x \neq y$, we have either $x \preceq y$ or $y \preceq x$ (not both).

As an example, the set containment “ $\subseteq$ ” is a partial ordering on any collection of sets; while “ $\leq$ ” is a total ordering on any subset of $\mathbb{R}$.

Definition 2.10 Let A be a set with a partial ordering “ $\preceq$ ”.

1. $a \in A$ is said to be minimal if, $\forall x \in A$, $x \preceq a$ implies $a \preceq x$;
2. $a \in A$ is said to be the least element if, $\forall x \in A$, $a \preceq x$, and $x \preceq a$ implies that $x = a$.
3. $a \in A$ is said to be maximal if, $\forall x \in A$, $a \preceq x$ implies $x \preceq a$;

4. $a \in A$ is said to be the greatest element if, $\forall x \in A, x \preceq a$, and $a \preceq x$ implies that $x = a$.

Definition 2.11 Let A be a set with a partial ordering " $\preceq$ ", and $E \subseteq A$.

1. $a \in A$ is said to be an upper bound of E if $x \preceq a, \forall x \in E$. It is the least upper bound of E if it is the least element in the set of all upper bounds of E ;
2. $a \in A$ is said to be a lower bound of E if $a \preceq x, \forall x \in E$. It is the greatest lower bound of E if it is the greatest element in the set of all lower bounds of E ;

We have the following results.

Proposition 2.12 Let A be a set with a partial ordering " $\preceq$ ". Then, the following holds.

- (i) If $a \in A$ is the least element, then it is minimal.
- (ii) There is at most one least element in A .
- (iii) Define a relation $\succeq$ by $\forall x, y \in A, x \succeq y$ if $y \preceq x$. Then, $\succeq$ is a partial ordering on A . Furthermore, $\succeq$ is antisymmetric if $\preceq$ is antisymmetric.
 1. $a \in A$ is the least element for $(E, \preceq)$ if, and only if, it is the greatest element for $(E, \succeq)$.
 2. $a \in A$ is minimal for $(E, \preceq)$ if, and only if, it is maximal for $(E, \succeq)$.
- (iv) If $a \in A$ is the greatest element, then it is maximal.
- (v) There is at most one greatest element in A .
- (vi) If $\preceq$ is antisymmetric, then $a \in A$ is minimal if, and only if, there does not exist $x \in A$ such that $x \preceq a$ and $x \neq a$.
- (vii) If $\preceq$ is antisymmetric, then $a \in A$ is maximal if, and only if, there does not exist $x \in A$ such that $a \preceq x$ and $x \neq a$.
- (viii) If $\preceq$ is antisymmetric, then it is a total ordering if, and only if, $\forall x_1, x_2 \in A$, we have $x_1 \preceq x_2$ or $x_2 \preceq x_1$.

Proof (i) is straightforward from Definition 2.10.

For (ii), let a_1 and a_2 be least elements of A . By a_1 being the least element, we have $a_1 \preceq a_2$. By a_2 being the least element, we then have $a_1 = a_2$. Hence, the least element is unique if it exists.

For (iii), $\forall x, y, z \in A$, Since $x \preceq x$ implies $x \succeq x$, then $\succeq$ is reflexive. If $x \succeq y$ and $y \succeq z$, we have $y \preceq x$ and $z \preceq y$, which implies $z \preceq x$, and hence

$x \succeq z$. This shows that $\succeq$ is transitive. Hence, $\succeq$ is a partial ordering on A .

When $\preceq$ is antisymmetric, $x \succeq y$ and $y \succeq x$ implies that $x \preceq y$ and $y \preceq x$, and therefore $x = y$. Hence, $\succeq$ is also antisymmetric.

For 1, “only if” let $a \in A$ be the least element in $(E, \preceq)$. $a \preceq x$ implies $x \succeq a$, $\forall x \in A$. $\forall x \in A$ with $a \succeq x$, we have $x \preceq a$, by a being the least element in $(E, \preceq)$, we have $x = a$. Hence, a is the greatest element in $(E, \succeq)$. The “if” part is similar to the “only if” part.

For 2, “only if” let $a \in A$ be a minimal element for $(E, \preceq)$. Then, $\forall x \in A$ with $a \succeq x$ implies $x \preceq a$ and hence $a \preceq x$, which yields $x \succeq a$. Hence, a is a maximal element for $(E, \succeq)$. The “if” part is similar to the “only if” part.

(iv) is straightforward from Definition 2.10.

For (v), Let a_1 and a_2 be greatest elements of A . By a_1 being the greatest element, we have $a_2 \preceq a_1$. By a_2 being the greatest element, we then have $a_1 = a_2$. Hence, the greatest element is unique if it exists.

For (vi), “if”, $\forall x \in A$ with $x \preceq a$, then, we have $a = x$, which means that $a \preceq x$; hence a is minimal. “Only if”, suppose that $\exists x \in A$ such that $x \preceq a$ and $x \neq a$. Note that $a \preceq x$ since a is minimal. Then, $a = x$, since $\preceq$ is antisymmetric, which is a contradiction.

For (vii), “if”, $\forall x \in A$ with $a \preceq x$, then, we have $a = x$, which means that $x \preceq a$; hence a is maximal. “Only if”, suppose that $\exists x \in A$ such that $a \preceq x$ and $x \neq a$. Note that $x \preceq a$ since a is maximal. Then, $a = x$, since $\preceq$ is antisymmetric, which is a contradiction.

For (viii), “if”, $\forall x_1, x_2 \in A$ with $x_1 \neq x_2$, we must have $x_1 \preceq x_2$ or $x_2 \preceq x_1$. They can not hold at the same time since, otherwise, $x_1 = x_2$, which is a contradiction. “Only if”, $\forall x_1, x_2 \in A$, when $x_1 = x_2$, then $x_1 \preceq x_2$; when $x_1 \neq x_2$, then $x_1 \preceq x_2$ or $x_2 \preceq x_1$; hence, in both cases, we have $x_1 \preceq x_2$ or $x_2 \preceq x_1$.

This completes the proof of the proposition. $\square$

2.7 Basic Principles

Now, we introduce the last axiom in ZFC axiom system.

Axiom 9 (Axiom of Choice) *Let $(A_\lambda)_{\lambda \in \Lambda}$ be a collection of nonempty sets, Λ is a set, (this collection is a set by Axiom 5), then, there exists a function $f : \Lambda \rightarrow \bigcup_{\lambda \in \Lambda} A_\lambda$, such that, $\forall \lambda \in \Lambda$, we have $f(\lambda) \in A_\lambda$.*

With Axioms 1–8 holding, the Axiom of Choice is equivalent to the following three results.

Theorem 2.13 (Hausdorff Maximal Principle) *Let $\preceq$ be a partial ordering on a set E . Then, there exists a maximal (with respect to set containment $\subseteq$) subset $F \subseteq E$, such that $\preceq$ is a total ordering on F .*

Theorem 2.14 (Zorn's Lemma) *Let $\preceq$ be an antisymmetric partial ordering on a nonempty set E . If every nonempty totally order subset F of E has an upper bound in E , then, there is a maximal element in E .*

Definition 2.15 *A well ordering of a set is a total ordering such that every nonempty subset has a least element.*

Theorem 2.16 (Well-Ordering Principle) *Every set can be well ordered.*

To prove the equivalence we described above, we need the following result.

Lemma 2.17 *Let E be a nonempty set and $\preceq$ is an antisymmetric partial ordering on E . Assume that every nonempty subset S of E , on which $\preceq$ is a total ordering, has a least upper bound in E . Let $f : E \rightarrow E$ be a mapping such that $x \preceq f(x)$, $\forall x \in E$. Then, f has a fixed point on E , i. e., $\exists w \in E$, $f(w) = w$.*

Proof Fix a point $a \in E$, since $E \neq \emptyset$. We define a collection of “good” sets:

$$\mathcal{B} = \{ B \subseteq E \mid (i) a \in B \quad (ii) f(B) \subseteq B \quad (iii) \forall F \subseteq B, F \neq \emptyset \\ F \text{ is totally ordered with } \preceq \text{ implies that the least upper} \\ \text{bound of } F \text{ belongs to } B. \}$$

Consider the set

$$B_0 := \{ x \in E \mid a \preceq x \}$$

Clearly, B_0 is nonempty since $a \in B_0$ and

$$f(B_0) = \{ f(x) \in E \mid a \preceq x \preceq f(x) \} \subseteq B_0$$

since f satisfies $x \preceq f(x)$, $\forall x \in E$.

For any $F \subseteq B_0$, such that F is totally ordered with $\preceq$ and $F \neq \emptyset$. Let e_0 be the least upper bound of F in E . Then, $\exists x_0 \in F$ such that $a \preceq x_0 \preceq e_0$, and therefore $e_0 \in B_0$.

This shows that $B_0 \in \mathcal{B}$, and $\mathcal{B}$ is nonempty.

The following result holds for the collection $\mathcal{B}$.

Claim 2.17.1 *Let $\{B_\alpha \mid \alpha \in \Lambda\}$ be any nonempty subcollection of $\mathcal{B}$, then $\bigcap_{\alpha \in \Lambda} B_\alpha \in \mathcal{B}$.*

Proof of claim: (i) $a \in B_\alpha$, $\forall \alpha \in \Lambda$. This implies $a \in \bigcap_{\alpha \in \Lambda} B_\alpha$.

(ii) By Proposition 2.5, we have $f(\bigcap_{\alpha \in \Lambda} B_\alpha) \subseteq \bigcap_{\alpha \in \Lambda} f(B_\alpha) \subseteq \bigcap_{\alpha \in \Lambda} B_\alpha$, where the last $\subseteq$ follows from the fact $f(B_\alpha) \subseteq B_\alpha$, $\forall \alpha \in \Lambda$.

(iii) Let $F \subseteq \bigcap_{\alpha \in \Lambda} B_\alpha$, which is totally ordered by $\preceq$ and $F \neq \emptyset$.

For any $\alpha \in \Lambda$, $F \subseteq B_\alpha$ implies that the least upper bound of F is an element of B_α . Therefore, the least upper bound of F is in the intersection $\bigcap_{\alpha \in \Lambda} B_\alpha$.

This establishes $\bigcap_{\alpha \in \Lambda} B_\alpha \in \mathcal{B}$, and completes the proof of the claim. $\square$

The claim shows that the collection $\mathcal{B}$ is closed under arbitrary intersection, as long as the collection is nonempty. Define $A := \bigcap_{B \in \mathcal{B}} B$. By the above claim, we have $A \in \mathcal{B}$, i. e., A is the smallest set in $\mathcal{B}$.

Hence, $A \subseteq B_0$, i. e., the set A satisfies, in addition to (i) – (iii),

(iv) $\forall x \in A, a \preceq x$.

Define the relation $\prec$ on E as $\forall x, y \in E, x \prec y$ if, and only if, $x \preceq y$ and $x \neq y$.

Define the set P by

$$P = \{x \in A \mid \forall y \in A, y \prec x \Rightarrow f(y) \preceq x\}$$

Clearly, $a \in P$, since there does not exist any $y \in A$ such that $y \prec a$, by $\preceq$ being antisymmetric. Therefore, P is nonempty.

We claim that

Claim 2.17.2 (v) $\forall x \in P, \forall z \in A, \text{ then } z \preceq x \text{ or } f(x) \preceq z$.

Proof of claim: Fix $x \in P$, and let

$$B := \{z \in A \mid z \preceq x\} \cup \{z \in A \mid f(x) \preceq z\}$$

We will show that $B \in \mathcal{B}$.

- (i) $a \in A, x \in P \subseteq A$, by (iv), $a \preceq x$, which further implies that $a \in B$.
- (ii) $\forall z \in B \subseteq A$, then $f(z) \in A$ since $A \in \mathcal{B}$. There are three exhaustive scenarios. If $z \prec x$, since $x \in P$ and $z \in B \subseteq A$, then $f(z) \preceq x$. This implies that $f(z) \in B$. If $z = x$, then $f(x) \preceq f(x) = f(z)$. This implies that $f(z) \in B$. If $f(x) \preceq z$, then $f(x) \preceq z \preceq f(z)$. This again implies that $f(z) \in B$. Hence, in all three scenarios, we have $f(z) \in B$. Then, $f(B) \subseteq B$ by the arbitrariness of $z \in B$.
- (iii) Let $F \neq \emptyset$ be any totally ordered subset of B and $e_0 \in E$ be the least upper bound of F . Since $F \subseteq B \subseteq A$ and $A \in \mathcal{B}$, then $e_0 \in A$. There are two exhaustive scenarios. If there exists $y \in F$ such that $f(x) \preceq y$, then, $f(x) \preceq y \preceq e_0$. This implies $e_0 \in B$. If, for any $y \in F, y \preceq x$, then $F \subseteq \{z \in A \mid z \preceq x\}$. This implies that x is an upper bound of F and $e_0 \preceq x$, since e_0 is the least upper bound of F . Therefore, $e_0 \in B$. In both of the cases, we have $e_0 \in B$.

This establishes that $B \in \mathcal{B}$. By A being the smallest set in $\mathcal{B}$, we have $A = B$. Therefore, the claim is proven. $\square$

Now, we show that $P \in \mathcal{B}$.

- (i) $a \in P$ and therefore $P \neq \emptyset$.
- (ii) Fix an $x \in P \subseteq A$. Then, $f(x) \in A$. $\forall y \in A$ such that $y \prec f(x)$. We need to show that $f(y) \preceq f(x)$, which then implies $f(x) \in P$. By (v),

there are two exhaustive scenarios. If $y \preceq x$, then $y \preceq x$. If $f(x) \preceq y$, then $f(x) \preceq y \prec f(x)$ form a contradiction by $\preceq$ being antisymmetric. Therefore, we must have $y \preceq x$, which results in the following two exhaustive scenarios. If $y \prec x$, then $f(y) \preceq x$ since $x \in P$. This implies that $f(y) \preceq x \preceq f(x)$. If $y = x$, then $f(y) = f(x) \preceq f(x)$. In both cases, we have $f(y) \preceq f(x)$. By the arbitrariness of y , we have $f(x) \in P$, which further implies $f(P) \subseteq P$ by the arbitrariness of $x \in P$.

(iii) Let $F \neq \emptyset$ be a totally ordered subset in P . Let $e_0 \in E$ be the least upper bound of F . We have $F \subseteq A$ implies that $e_0 \in A$ by $A \in \mathcal{B}$. $\forall z \in A$ with $z \prec e_0$, implies that z must not be an upper bound of F . Therefore, $\exists x_0 \in F$ such that $x_0 \not\preceq z$. By (v), we have $z \prec x_0$. Hence, by $x_0 \in F \subseteq P$, $z \in A$, and $z \prec x_0$, we have $f(z) \preceq x_0$. Therefore, $f(z) \preceq e_0$ since e_0 is an upper bound of F . This further implies that $e_0 \in P$ by the arbitrariness of z .

This proves that $P \in \mathcal{B}$.

Since $P \subseteq A$ and A is the smallest set in $\mathcal{B}$, then, $P = A$.

The set A satisfies properties (i) – (v).

For any $x_1, x_2 \in A$, by (v), there are two exhaustive scenarios. If $x_1 \preceq x_2$, then, x_1 and x_2 are related through $\preceq$. If $f(x_2) \preceq x_1$, then, $x_2 \preceq f(x_2) \preceq x_1$, which implies that x_1 and x_2 are related through $\preceq$. Therefore, x_1 and x_2 are related through $\preceq$ in both cases. Then, by Proposition 2.12 (viii), A is totally ordered by $\preceq$ and nonempty. Let $w \in E$ be the least upper bound of A . Then, $w \in A$, since $A \in \mathcal{B}$.

Therefore, $f(w) \in A$ by $f(A) \subseteq A$, which implies that $f(w) \preceq w$. This coupled with $w \preceq f(w)$ yields $f(w) = w$, since $\preceq$ is antisymmetric.

This completes the proof of the lemma. $\square$

Theorem 2.18 *Under the Axioms 1–8, the following are equivalent.*

1. *Axiom of Choice*
2. *Hausdorff Maximum Principle*
3. *Zorn's Lemma*
4. *Well-ordering principle*

Proof 1. $\Rightarrow$ 2. Define

$$\mathcal{E} := \{ A \subseteq E \mid \preceq \text{ defines a total ordering on } A \}$$

Clearly, $\emptyset \in \mathcal{E}$, then $\mathcal{E} \neq \emptyset$. Define a partial ordering on $\mathcal{E}$ by $\subseteq$, which is set containment. This partial ordering $\subseteq$ is clearly reflexive, transitive, and antisymmetric.

$\forall A \in \mathcal{E}$, define a collection

$$\mathcal{A}_A := \begin{cases} \{ B \in \mathcal{E} \mid A \subset B \} & \text{if } \exists B \in \mathcal{E} \text{ such that } A \subset B \\ \{ A \} & \text{otherwise} \end{cases}$$

Clearly, $\mathcal{A}_A \neq \emptyset$. By Axiom of Choice, $\exists T : \mathcal{E} \rightarrow \mathcal{E}$ such that $T(A) = B \in \mathcal{A}_A, \forall A \in \mathcal{E}$.

We will show that T admits a fixed point by Lemma 2.17. Let $\mathcal{B} \subseteq \mathcal{E}$ be any nonempty subset on which $\subseteq$ is a total ordering. Let $C := \bigcup_{B \in \mathcal{B}} B$. Clearly, $C \subseteq E$. We will show $\preceq$ is a total ordering on C . Since $\preceq$ is a partial ordering on E , then it is a partial ordering on C . $\forall x_1, x_2 \in C$, $\exists B_1, B_2 \in \mathcal{B}$ such that $x_1 \in B_1$ and $x_2 \in B_2$. Since $\subseteq$ is a total ordering on $\mathcal{B}$, then, we may without loss of generality assume $B_1 \subseteq B_2$. Then, $x_1, x_2 \in B_2$. Since $B_2 \in \mathcal{B} \subseteq \mathcal{E}$, then $\preceq$ is a total ordering on B_2 , which means that we have $x_1 \preceq x_2$ or $x_2 \preceq x_1$. Furthermore, if $x_1 \preceq x_2$ and $x_2 \preceq x_1$, then $x_1 = x_2$ by $\preceq$ being antisymmetric on B_2 . Therefore, by Proposition 2.12, $\preceq$ is a total ordering on C . Hence, $C \in \mathcal{E}$. This shows that $\mathcal{B}$ admits least upper bound C in $\mathcal{E}$ with respect to $\subseteq$. By the definition of T , it is clear that $A \subseteq T(A), \forall A \in \mathcal{E}$. By Lemma 2.17, T has a fixed point on $\mathcal{E}$, i. e., $\exists A_0 \in \mathcal{E}$ such that $T(A_0) = A_0$.

By the definitions of T and $\mathcal{A}_{A_0}$, there does not exist $B \in \mathcal{E}$ such that $A_0 \subset B$. Hence, by Proposition 2.12 (vii), A_0 is maximal in $\mathcal{E}$ with respect to $\subseteq$.

2. $\Rightarrow$ 3. Let E be a nonempty set with an antisymmetric partial ordering $\preceq$. By Hausdorff Maximum Principle, there exists a maximal (with respect to $\subseteq$) totally ordered (with respect to $\preceq$) subset $F \subseteq E$. We must have $F \neq \emptyset$, otherwise, let $x_0 \in E$ (since $E \neq \emptyset$), $F \subset \{x_0\} \subseteq E$ and $\{x_0\}$ is totally ordered by $\preceq$, which violates the fact that F is maximal (with respect to $\subseteq$). Then, F has an upper bound $e_0 \in E$.

Claim 2.18.1 $e_0 \in F$.

Proof of claim: Suppose $e_0 \notin F$. Define $A := F \cup \{e_0\} \subseteq E$. Clearly, $F \subseteq A$ and $F \neq A$. We will show that $\preceq$ is a total ordering on A . Clearly, $\preceq$ is an antisymmetric partial ordering on A since it is an antisymmetric partial ordering on E . $\forall x_1, x_2 \in A$, we will distinguish 4 exhaustive and mutually exclusive cases: Case 1: $x_1, x_2 \in F$; Case 2: $x_1 \in F, x_2 = e_0$; Case 3: $x_1 = e_0, x_2 \in F$; Case 4: $x_1 = x_2 = e_0$. In Case 1, we have $x_1 \preceq x_2$ or $x_2 \preceq x_1$ since $\preceq$ is a total ordering on F . In Case 2, we have $x_1 \preceq x_2 = e_0$ since e_0 is an upper bound of F . In Case 3, we have $x_2 \preceq x_1 = e_0$. In Case 4, we have $x_1 = e_0 \preceq e_0 = x_2$. Hence, $\preceq$ is a total ordering on A . Note that $F \subseteq A$ and $F \neq A$. By Proposition 2.12 (vii), this contradicts with the fact that F is maximal with respect to $\subseteq$. Therefore, we must have $e_0 \in F$. This completes the proof of the claim. $\square$

$\forall e_1 \in E$ such that $e_0 \preceq e_1$. $\forall x \in F$, we have $x \preceq e_0 \preceq e_1$. Hence, e_1 is an upper bound of F . By Claim 2.18.1, we must have $e_1 \in F$. Then, $e_1 \preceq e_0$ since e_0 is an upper bound of F . This shows that e_0 is maximal in E with respect to $\preceq$.

3. $\Rightarrow$ 4. Let E be a set. It is clear that $\emptyset \subseteq E$ is well-ordered by the empty relation. Define

$$\mathcal{E} := \{(A_\alpha, \preceq_\alpha) \mid A_\alpha \subseteq E, A_\alpha \text{ is well-ordered by } \preceq_\alpha\}$$

Then, $\mathcal{E} \neq \emptyset$. Define an ordering $\preceq$ on $\mathcal{E}$ by $\forall (A_1, \sqsubseteq_1), (A_2, \sqsubseteq_2) \in \mathcal{E}$, we say $(A_1, \sqsubseteq_1) \preceq (A_2, \sqsubseteq_2)$ if the following three conditions hold: (i) $A_1 \subseteq A_2$; (ii) $\sqsubseteq_2 = \sqsubseteq_1$ on A_1 ; (iii) $\forall x_1 \in A_1, \forall x_2 \in A_2 \setminus A_1$, we have $x_1 \sqsubseteq_2 x_2$.

Now, we will show that $\preceq$ defines an antisymmetric partial ordering on $\mathcal{E}$. $\forall (A_1, \sqsubseteq_1), (A_2, \sqsubseteq_2), (A_3, \sqsubseteq_3) \in \mathcal{E}$. Clearly, $(A_1, \sqsubseteq_1) \preceq (A_1, \sqsubseteq_1)$. Hence, $\preceq$ is reflexive. If $(A_1, \sqsubseteq_1) \preceq (A_2, \sqsubseteq_2)$ and $(A_2, \sqsubseteq_2) \preceq (A_3, \sqsubseteq_3)$, we have $A_1 \subseteq A_2 \subseteq A_3$, and (i) holds; (ii) $\sqsubseteq_3 = \sqsubseteq_2$ on A_2 and $\sqsubseteq_2 = \sqsubseteq_1$ on A_1 implies that $\sqsubseteq_3 = \sqsubseteq_1$ on A_1 ; (iii) $\forall x_1 \in A_1, \forall x_2 \in A_3 \setminus A_1$, we have 2 exhaustive scenarios: if $x_2 \in A_2$, then $x_2 \in A_2 \setminus A_1$ which implies $x_1 \sqsubseteq_2 x_2$ and hence $x_1 \sqsubseteq_3 x_2$; if $x_2 \in A_3 \setminus A_2$, then we have $x_1 \in A_2$ and $x_1 \sqsubseteq_3 x_2$, thus, we have $x_1 \sqsubseteq_3 x_2$ in both cases. Therefore, $(A_1, \sqsubseteq_1) \preceq (A_3, \sqsubseteq_3)$ and hence $\preceq$ is transitive. If $(A_1, \sqsubseteq_1) \preceq (A_2, \sqsubseteq_2)$ and $(A_2, \sqsubseteq_2) \preceq (A_1, \sqsubseteq_1)$, then $A_1 \subseteq A_2 \subseteq A_1 \Rightarrow A_1 = A_2$ and $\sqsubseteq_2 = \sqsubseteq_1$ on A_1 . Hence, $(A_1, \sqsubseteq_1) = (A_2, \sqsubseteq_2)$, which shows that $\preceq$ is antisymmetric. Therefore, $\preceq$ defines an antisymmetric partial ordering on $\mathcal{E}$.

Let $\mathcal{A} \subseteq \mathcal{E}$ be any nonempty subset totally ordered by $\preceq$. Take $\mathcal{A} = \{(A_\alpha, \sqsubseteq_\alpha) \mid \alpha \in \Lambda\}$ where $\Lambda \neq \emptyset$ is an index set. Define $A := \bigcup_{\alpha \in \Lambda} A_\alpha$. Define an ordering $\sqsubseteq$ on A by: $\forall x_1, x_2 \in A, \exists (A_1, \sqsubseteq_1), (A_2, \sqsubseteq_2) \in \mathcal{A}$ such that $x_1 \in A_1$ and $x_2 \in A_2$, without loss of generality, assume that $(A_1, \sqsubseteq_1) \preceq (A_2, \sqsubseteq_2)$ since $\mathcal{A}$ is totally ordered by $\preceq$, then $x_1, x_2 \in A_2$, we will say that $x_1 \sqsubseteq x_2$ if $x_1 \sqsubseteq_2 x_2$. We will now show that this ordering is uniquely defined independent of $(A_2, \sqsubseteq_2) \in \mathcal{A}$. Let $(A_3, \sqsubseteq_3) \in \mathcal{A}$ be such that $x_1, x_2 \in A_3$. Since $\mathcal{A}$ is totally ordered by $\preceq$, then there are two exhaustive cases: Case 1: $(A_3, \sqsubseteq_3) \preceq (A_2, \sqsubseteq_2)$; Case 2: $(A_2, \sqsubseteq_2) \preceq (A_3, \sqsubseteq_3)$. In Case 1, we have $A_3 \subseteq A_2$ and $\sqsubseteq_3 = \sqsubseteq_2$ on A_3 , which implies that $x_1 \sqsubseteq x_2 \Leftrightarrow x_1 \sqsubseteq_2 x_2 \Leftrightarrow x_1 \sqsubseteq_3 x_2$. In Case 2, we have $A_2 \subseteq A_3$ and $\sqsubseteq_3 = \sqsubseteq_2$ on A_2 , which implies that $x_1 \sqsubseteq x_2 \Leftrightarrow x_1 \sqsubseteq_2 x_2 \Leftrightarrow x_1 \sqsubseteq_3 x_2$. Hence, the ordering $\sqsubseteq$ is well-defined on A .

Next, we will show that $\sqsubseteq$ is a total ordering on A . $\forall x_1, x_2, x_3 \in A$. $\exists (A_i, \sqsubseteq_i) \in \mathcal{A}$ such that $x_i \in A_i, i = 1, 2, 3$. Since $\mathcal{A}$ is totally ordered by $\preceq$, then, without loss of generality, assume that $(A_1, \sqsubseteq_1) \preceq (A_2, \sqsubseteq_2) \preceq (A_3, \sqsubseteq_3)$. Then, $x_1, x_2, x_3 \in A_3$. Clearly, $x_1 \sqsubseteq x_1$ since $x_1 \sqsubseteq_3 x_1$, which implies that $\sqsubseteq$ is reflexive. If $x_1 \sqsubseteq x_2$ and $x_2 \sqsubseteq x_3$, then, $x_1 \sqsubseteq_3 x_2 \sqsubseteq_3 x_3$, which implies $x_1 \sqsubseteq_3 x_3$ since $\sqsubseteq_3$ is transitive on A_3 , and hence, $x_1 \sqsubseteq x_3$. This shows that $\sqsubseteq$ is transitive. If $x_1 \sqsubseteq x_2$ and $x_2 \sqsubseteq x_1$, then $x_1 \sqsubseteq_3 x_2$ and $x_2 \sqsubseteq_3 x_1$, which implies that $x_1 = x_2$ since $\sqsubseteq_3$ is antisymmetric on A_3 . This shows that $\sqsubseteq$ is antisymmetric. Since $\sqsubseteq_3$ is a well-ordering on A_3 , then we must have $x_1 \sqsubseteq_3 x_2 \Leftrightarrow x_1 \sqsubseteq x_2$ or $x_2 \sqsubseteq_3 x_1 \Leftrightarrow x_2 \sqsubseteq x_1$. Hence, $\sqsubseteq$ defines a total ordering on A .

Next, we will show that $\sqsubseteq$ is a well-ordering on A . $\forall B \subseteq A$ with $B \neq \emptyset$. Fix $x_0 \in B$. Then, $\exists (A_1, \sqsubseteq_1) \in \mathcal{A}$ such that $x_0 \in A_1$. Note that $\emptyset \neq B \cap A_1 \subseteq A_1$. Since A_1 is well-ordered by $\sqsubseteq_1$, then $\exists e \in B \cap A_1$, which is the least element of $B \cap A_1$. $\forall y \in B \subseteq A, \exists (A_2, \sqsubseteq_2) \in \mathcal{A}$ such that $y \in A_2$. We have 2 exhaustive and mutually exclusive cases: Case 1: $y \in A_1$; Case 2: $y \in A_2 \setminus A_1$. In Case 1, $e \sqsubseteq_1 y$ since e is the least element

of $B \cap A_1$, which implies that $e \trianglelefteq y$. In Case 2, since $\mathcal{A}$ is totally ordered by $\preceq$, we must have $(A_1, \trianglelefteq_1) \preceq (A_2, \trianglelefteq_2)$, which implies that $e \trianglelefteq_2 y$, by (iii) in the definition of $\preceq$, and hence $e \trianglelefteq y$. In both cases, we have shown that $e \trianglelefteq y$. Since $\trianglelefteq$ is a total ordering on A , then e is the least element of B . Therefore, $\trianglelefteq$ is a well ordering on A , which implies $(A, \trianglelefteq) \in \mathcal{E}$.

$\forall (A_1, \trianglelefteq_1) \in \mathcal{A}$. (i) $A_1 \subseteq A$. (ii) $\forall x_1, x_2 \in A_1$, $x_1 \trianglelefteq_1 x_2 \Leftrightarrow x_1 \trianglelefteq x_2$; hence $\trianglelefteq = \trianglelefteq_1$ on A_1 . (iii) $\forall x_1 \in A_1$, $\forall x_2 \in A \setminus A_1$, $\exists (A_2, \trianglelefteq_2) \in \mathcal{A}$ such that $x_2 \in A_2 \setminus A_1$; since $\mathcal{A}$ is totally ordered by $\preceq$, then, we must have $(A_1, \trianglelefteq_1) \preceq (A_2, \trianglelefteq_2)$; hence $x_1 \trianglelefteq_2 x_2$ and $x_1 \trianglelefteq x_2$. Therefore, we have shown $(A_1, \trianglelefteq_1) \preceq (A, \trianglelefteq)$. Hence, $(A, \trianglelefteq) \in \mathcal{E}$ is an upper bound of $\mathcal{A}$.

By Zorn's Lemma, there is a maximal element $(F, \trianglelefteq_F) \in \mathcal{E}$. We claim that $F = E$. We will prove this by an argument of contradiction. Suppose $F \subset E$, then $\exists x_0 \in E \setminus F$. Let $H := F \cup \{x_0\}$. Define an ordering $\trianglelefteq_H$ on H by: $\forall x_1, x_2 \in H$, if $x_1, x_2 \in F$, we say $x_1 \trianglelefteq_H x_2$ if $x_1 \trianglelefteq_F x_2$; if $x_1 \in F$ and $x_2 = x_0$, then we let $x_1 \trianglelefteq_H x_2$; if $x_1 = x_2 = x_0$, we let $x_1 \trianglelefteq_H x_2$. Now, we will show that $\trianglelefteq_H$ is a well ordering on H . $\forall x_1, x_2, x_3 \in H$. If $x_1 \in F$, then, $x_1 \trianglelefteq_F x_1$ and $x_1 \trianglelefteq_H x_1$; if $x_1 = x_0$, then $x_1 \trianglelefteq_H x_1$. Hence, $\trianglelefteq_H$ is reflexive. If $x_1 \trianglelefteq_H x_2$ and $x_2 \trianglelefteq_H x_3$. We have 4 exhaustive and mutually exclusive cases: Case 1: $x_1, x_3 \in F$; Case 2: $x_1 \in F$ and $x_3 = x_0$; Case 3: $x_3 \in F$ and $x_1 = x_0$; Case 4: $x_1 = x_3 = x_0$. In Case 1, we must have $x_2 \in F$ and then $x_1 \trianglelefteq_F x_2$ and $x_2 \trianglelefteq_F x_3$, which implies that $x_1 \trianglelefteq_F x_3$, and hence $x_1 \trianglelefteq_H x_3$. In Case 2, we have $x_1 \trianglelefteq_H x_3$. In Case 3, we must have $x_2 = x_0$, which leads to a contradiction $x_0 \trianglelefteq_H x_3$, hence, this case is impossible. In Case 4, we have $x_1 \trianglelefteq_H x_3$. In all cases except that is impossible, we have $x_1 \trianglelefteq_H x_3$. Hence, $\trianglelefteq_H$ is transitive. If $x_1 \trianglelefteq_H x_2$ and $x_2 \trianglelefteq_H x_1$. We have 4 exhaustive and mutually exclusive cases: Case 1: $x_1, x_2 \in F$; Case 2: $x_1 \in F$ and $x_2 = x_0$; Case 3: $x_2 \in F$ and $x_1 = x_0$; Case 4: $x_1 = x_2 = x_0$. In Case 1, we have $x_1 \trianglelefteq_F x_2$ and $x_2 \trianglelefteq_F x_1$, which implies that $x_1 = x_2$ since $\trianglelefteq_F$ is antisymmetric on F . In Case 2, we have $x_0 \trianglelefteq_H x_1$, which is a contradiction, and hence this case is impossible. In Case 3, we have $x_0 \trianglelefteq_H x_2$, which is a contradiction, and hence this case is impossible. In Case 4, we have $x_1 = x_2$. In all cases except those impossible, we have $x_1 = x_2$. Hence, $\trianglelefteq_H$ is antisymmetric. When $x_1, x_2 \in F$, then, we must have $x_1 \trianglelefteq_F x_2$ or $x_2 \trianglelefteq_F x_1$ since $\trianglelefteq_F$ is a well ordering on F , and hence $x_1 \trianglelefteq_H x_2$ or $x_2 \trianglelefteq_H x_1$. When $x_1 \in F$ and $x_2 = x_0$, then $x_1 \trianglelefteq_H x_2$. When $x_2 \in F$ and $x_1 = x_0$, then $x_2 \trianglelefteq_H x_1$. When $x_1 = x_2 = x_0$, then $x_1 \trianglelefteq_H x_2$. This shows that $\trianglelefteq_H$ is a total ordering on H . $\forall B \subseteq H$ with $B \neq \emptyset$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $B = \{x_0\}$; Case 2: $B \neq \{x_0\}$. In Case 1, x_0 is the least element of B . In Case 2, $B \setminus \{x_0\} \subseteq F$ and is nonempty, and hence admits a least element $e_0 \in B \setminus \{x_0\} \subseteq F$ with respect to $\trianglelefteq_F$. $\forall x \in B$, if $x \in B \setminus \{x_0\}$, then $e_0 \trianglelefteq_F x$ and hence $e_0 \trianglelefteq_H x$; if $x = x_0$, then $e_0 \trianglelefteq_H x$. Hence, e_0 is the least element of B since $\trianglelefteq_H$ is a total ordering on H . Therefore, $\trianglelefteq_H$ is a well ordering on H and $(H, \trianglelefteq_H) \in \mathcal{E}$.

Clearly, $F \subset H$, $\trianglelefteq_F = \trianglelefteq_H$ on F , and $\forall x_1 \in F$ and $\forall x_2 \in H \setminus F$, we have $x_2 = x_0$ and $x_1 \trianglelefteq_H x_2$. This implies that $(F, \trianglelefteq_F) \preceq (H, \trianglelefteq_H)$. Since $(F, \trianglelefteq_F)$

is maximal in $\mathcal{E}$ with respect to $\preceq$, we must have $(H, \trianglelefteq_H) \preceq (F, \trianglelefteq_F)$, and hence, $H \subseteq F$. This is a contradiction. Therefore, $F = E$ and E is well ordered by $\trianglelefteq_F$.

4. $\Rightarrow$ 1. Let $(A_\lambda)_{\lambda \in \Lambda}$ be a collection of nonempty sets, and Λ is a set. Let $A := \bigcup_{\lambda \in \Lambda} A_\lambda$. By Well-Ordering Principle, A may be well ordered by $\trianglelefteq$. $\forall \lambda \in \Lambda$, $A_\lambda \subseteq A$ is nonempty and admits the least element $e_\lambda \in A_\lambda$. This defines a function $f : \Lambda \rightarrow A$ by $f(\lambda) = e_\lambda \in A_\lambda$, $\forall \lambda \in \Lambda$.

This completes the proof of the theorem. $\square$

Example 2.19 Let Λ be an index set and $(A_\alpha)_{\alpha \in \Lambda}$ be a collection of sets. We will try to define the *Cartesian (direct) product* $\prod_{\alpha \in \Lambda} A_\alpha$. Let $A = \bigcup_{\alpha \in \Lambda} A_\alpha$, which is a set by the Axiom of Union. Then, as we discussed in Section 2.3, A^Λ is a set, which consists of all functions of Λ to A . Define the *projection* functions $\pi_\alpha : A^\Lambda \rightarrow A$, $\forall \alpha \in \Lambda$, by, $\forall f \in A^\Lambda$, $\pi_\alpha(f) = f(\alpha)$. Then, we may define the set

$$\prod_{\alpha \in \Lambda} A_\alpha := \{ f \in A^\Lambda \mid \pi_\alpha(f) \in A_\alpha, \forall \alpha \in \Lambda \}$$

When all of A_α 's are nonempty, then, by Axiom of Choice, the product $\prod_{\alpha \in \Lambda} A_\alpha$ is also nonempty. $\diamond$

Chapter 3

Topological Spaces

3.1 Fundamental Notions

Definition 3.1 A topological space $(X, \mathcal{O})$ consists of a set X and a collection $\mathcal{O}$ of subsets (namely, open subsets) of X such that

- (i) $\emptyset, X \in \mathcal{O}$;
- (ii) $\forall O_1, O_2 \in \mathcal{O}$, we have $O_1 \cap O_2 \in \mathcal{O}$;
- (iii) $\forall (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}$, where Λ is an index set, we have $\bigcup_{\alpha \in \Lambda} O_\alpha \in \mathcal{O}$.

The collection $\mathcal{O}$ is called a topology for the set X .

Definition 3.2 Let $(X, \mathcal{O})$ be a topological space and $F \subseteq X$. The complement of F is $\tilde{F} := X \setminus F$. F is said to be closed if $\tilde{F} \in \mathcal{O}$. The closure of F is given by $\overline{F} := \bigcap_{\substack{F \subseteq B \\ B \in \mathcal{O}}} B$, which is clearly a closed set. The interior

of F is given by $F^\circ := \bigcup_{\substack{B \subseteq F \\ B \in \mathcal{O}}} B$, which is clearly an open set. A point of closure of F is a point in $\overline{F}$. An interior point of F is a point in F° . A boundary point of F is a point $x \in X$ such that $\forall O \in \mathcal{O}$ with $x \in O$, we have $O \cap F \neq \emptyset$ and $O \cap \tilde{F} \neq \emptyset$. The boundary of F , denoted by ∂F , is the set of all boundary points of F . An exterior point of F is a point in $\tilde{F}^\circ$, where $\tilde{F}^\circ$ is called the exterior of F . An accumulation point of F is a point $x \in X$ such that $\forall O \in \mathcal{O}$ with $x \in O$, we have $O \cap (F \setminus \{x\}) \neq \emptyset$.

Clearly, $\emptyset$ and X are both closed and open.

Proposition 3.3 Let $(X, \mathcal{O})$ be a topological space and A, B, E are subsets of X . Then,

- (i) $E \subseteq \overline{E}$, $\overline{\overline{E}} = \overline{E}$, $E^\circ \subseteq E$, $(E^\circ)^\circ = E^\circ$, and $\widetilde{E}^\circ = \widetilde{E}$;
- (ii) $\forall x \in X$, x is a point of closure of E if, and only if, $\forall O \in \mathcal{O}$ with $x \in O$, we have $O \cap E \neq \emptyset$;
- (iii) $\forall x \in X$, x is an interior point of E if, and only if, $\exists O \in \mathcal{O}$ with $x \in O$ such that $O \subseteq E$;
- (iv) $\overline{A \cup B} = \overline{A} \cup \overline{B}$, $(A \cap B)^\circ = A^\circ \cap B^\circ$;
- (v) E is closed if, and only if, $E = \overline{E}$;
- (vi) $\overline{E} = E^\circ \cup \partial E$.
- (vii) X equals to the disjoint union $E^\circ \cup \partial E \cup (\widetilde{E})^\circ$;

Proof (i) Clearly, $E \subseteq \overline{E}$. Then, $\overline{E} \subseteq \overline{\overline{E}}$. $\forall C \supseteq E$ with $\widetilde{C} \in \mathcal{O}$, we have $C \supseteq \overline{E}$. Then, $\overline{E} = \bigcap_{\substack{C \supseteq E \\ \widetilde{C} \in \mathcal{O}}} C \supseteq \bigcap_{\substack{C \supseteq \overline{E} \\ \widetilde{C} \in \mathcal{O}}} C = \overline{\overline{E}}$. Hence, we have $\overline{E} = \overline{\overline{E}}$.

Clearly, $E^\circ \subseteq E$. Note that

$$\widetilde{E} = \left(\bigcap_{\substack{B \supseteq E \\ \widetilde{B} \in \mathcal{O}}} B \right)^\sim = \bigcup_{\substack{B \supseteq E \\ \widetilde{B} \in \mathcal{O}}} \widetilde{B} = \bigcup_{\substack{O \subseteq \widetilde{E} \\ O \in \mathcal{O}}} O = \widetilde{E}^\circ$$

Furthermore,

$$(E^\circ)^\circ = \left(\widetilde{E}^\circ \right)^\circ = \left(\widetilde{\widetilde{E}} \right)^\circ = \widetilde{\widetilde{\widetilde{E}}} = \widetilde{\widetilde{E}} = \left(\widetilde{E} \right)^\circ = E^\circ$$

(ii) “Only if” $\forall x \in \overline{E}$, we have $x \in \bigcap_{\substack{B \supseteq E \\ \widetilde{B} \in \mathcal{O}}} B$. $\forall O \in \mathcal{O}$ with $x \in O$, let $O_1 := O \cup \widetilde{E} \in \mathcal{O}$. Note that $x \in O_1$ and $E \cap \widetilde{E} = \emptyset$. Suppose $O \cap E = \emptyset$. Then, $E \cap O_1 = \emptyset$, which further implies that $E \subseteq \widetilde{O_1}$. Then, $\overline{E} \subseteq \widetilde{O_1}$ and $x \in \widetilde{O_1}$. This contradicts with $x \in O$. Hence, $O \cap E \neq \emptyset$.

“If” $\forall x \in \widetilde{E} = \widetilde{E}^\circ \in \mathcal{O}$. Then, $\exists O := \widetilde{E}^\circ \in \mathcal{O}$ such that $E \cap O = \emptyset$. Hence, the result holds.

(iii) “Only if” $\forall x \in E^\circ \subseteq E$, then $E^\circ \in \mathcal{O}$.

“If” $\forall x \in X$, $\exists O \in \mathcal{O}$ such that $x \in O \subseteq E$. Then, $x \in O \subseteq \bigcup_{\substack{B \subseteq E \\ B \in \mathcal{O}}} B = E^\circ$. Hence, the result holds.

(iv) Let $\mathcal{B} := \{O \in \mathcal{O} \mid O \subseteq A \cap B\}$, $\mathcal{B}_A := \{O \in \mathcal{O} \mid O \subseteq A\}$, and $\mathcal{B}_B := \{O \in \mathcal{O} \mid O \subseteq B\}$. $\forall O_1 \in \mathcal{B}_A$ and $\forall O_2 \in \mathcal{B}_B$, then, $O_1 \cap O_2 \in \mathcal{B}$. On the other hand, $\forall O \in \mathcal{B}$, we have $O = O \cap O$ and $O \in \mathcal{B}_A$ and $O \in \mathcal{B}_B$. Then,

$$\begin{aligned} (A \cap B)^\circ &= \bigcup_{\substack{O \subseteq A \cap B \\ O \in \mathcal{O}}} O = \bigcup_{\substack{O_1 \subseteq A, O_2 \subseteq B \\ O_1, O_2 \in \mathcal{O}}} (O_1 \cap O_2) \\ &= \left(\bigcup_{\substack{O_1 \subseteq A \\ O_1 \in \mathcal{O}}} O_1 \right) \cap \left(\bigcup_{\substack{O_2 \subseteq B \\ O_2 \in \mathcal{O}}} O_2 \right) = A^\circ \cap B^\circ \end{aligned}$$

We also have

$$\begin{aligned}\overline{A \cup B} &= \left(\left(\widetilde{A \cup B} \right)^\circ \right)^\sim = \left(\left(\widetilde{A} \cap \widetilde{B} \right)^\circ \right)^\sim = \left(\left(\widetilde{A} \right)^\circ \cap \left(\widetilde{B} \right)^\circ \right)^\sim \\ &= \widetilde{\widetilde{A} \cap \widetilde{B}} = \overline{A} \cup \overline{B}\end{aligned}$$

(v) “If” E is closed since $E = \overline{E}$ and $\overline{E}$ is closed.

“Only if” Since E is closed, then $\overline{E} \subseteq E$. Then, we have $E = \overline{E}$. Hence, the result holds.

(vi) This result follows directly from (ii), (iii), and Definition 3.2.

(vii) Note that $X = \overline{E} \cup \widetilde{E} = E^\circ \cup \partial E \cup \widetilde{E}^\circ$. By (iii) and Definition 3.2, E° and ∂E are disjoint. It is obvious that $\widetilde{E}$ is disjoint with $E^\circ \cup \partial E = \overline{E}$. Hence, the result holds. $\square$

To simplify notation in the theory, we will abuse the notation to write $x \in \mathcal{X}$ when $x \in X$ and $A \subseteq \mathcal{X}$ when $A \subseteq X$ for a topological space $\mathcal{X} := (X, \mathcal{O})$. We will later simply discuss a topological space $\mathcal{X}$ without further reference to components of $\mathcal{X}$, where the topology is understood to be $\mathcal{O}_{\mathcal{X}}$. When it is clear from the context, we will neglect the subscript $\mathcal{X}$.

Proposition 3.4 *Let $(X, \mathcal{O})$ be a topological space and $A \subseteq X$. A admits the subset topology $\mathcal{O}_A := \{O \cap A \mid O \in \mathcal{O}\}$.*

Proof Clearly, $\mathcal{O}_A$ is a collection of subsets of A . $\emptyset = \emptyset \cap A \in \mathcal{O}_A$ and $A = X \cap A \in \mathcal{O}_A$. $\forall O_{A1}, O_{A2} \in \mathcal{O}_A$, $\exists O_1, O_2 \in \mathcal{O}$ such that $O_{A1} = O_1 \cap A$ and $O_{A2} = O_2 \cap A$. Then, $O_1 \cap O_2 \in \mathcal{O}$ since $\mathcal{O}$ is a topology. Then, $O_{A1} \cap O_{A2} = (O_1 \cap O_2) \cap A \in \mathcal{O}_A$. $\forall (O_{A\alpha})_{\alpha \in \Lambda} \subseteq \mathcal{O}_A$, where Λ is an index set, we have, $\forall \alpha \in \Lambda$, $\exists O_\alpha \in \mathcal{O}$ such that $O_{A\alpha} = O_\alpha \cap A$. Then, $\bigcup_{\alpha \in \Lambda} O_\alpha \in \mathcal{O}$ since $\mathcal{O}$ is a topology. Therefore, $\bigcup_{\alpha \in \Lambda} O_{A\alpha} = \left(\bigcup_{\alpha \in \Lambda} O_\alpha \right) \cap A \in \mathcal{O}_A$.

Hence, $\mathcal{O}_A$ is a topology on A . $\square$

Let $(X, \mathcal{O})$ be a topological space and $A \subseteq X$. The property of a set $E \subseteq A$ being open or closed is relative with respect to $(X, \mathcal{O})$, that is, this property may change if we consider the subset topology $(A, \mathcal{O}_A)$.

Proposition 3.5 *Let $\mathcal{X}$ be a topological space, $A \subseteq \mathcal{X}$ be endowed with the subset topology $\mathcal{O}_A$, and $E \subseteq A$. Then,*

- (1) E is closed in $\mathcal{O}_A$ if, and only if, $E = A \cap F$, where $F \subseteq \mathcal{X}$ is closed in $\mathcal{O}_{\mathcal{X}}$;
- (2) the closure of E relative to $(A, \mathcal{O}_A)$ (the closure of E in $\mathcal{O}_A$) is equal to $\overline{E} \cap A$, where $\overline{E}$ is the closure of E relative to $\mathcal{X}$.

Proof Here, the set complementation and set closure operation are relative to $\mathcal{X}$.

(1) “If” $A \setminus E = A \setminus (A \cap F) = A \cap \widetilde{A \cap F} = A \cap \widetilde{F}$. Since F is closed in $\mathcal{O}_{\mathcal{X}}$, then $\widetilde{F} \in \mathcal{O}_{\mathcal{X}}$. Then, $A \setminus E \in \mathcal{O}_A$. Hence, E is closed in $\mathcal{O}_A$.

“Only if” $A \setminus E \in \mathcal{O}_A$. Then, $\exists O \in \mathcal{O}_X$ such that $A \setminus E = A \cap O$. Then, $E = A \setminus (A \setminus E) = A \cap \widetilde{A \cap O} = A \cap \widetilde{O}$. Hence, the result holds.

(2) By (1), $\overline{E} \cap A$ is closed in $\mathcal{O}_A$. Then, the closure of E relative to $(A, \mathcal{O}_A)$ is contained in $\overline{E} \cap A$. On the other hand, by Proposition 3.3, if $x \in X$ is a point of closure of E relative to X , then it is a point of closure of E in $\mathcal{O}_A$ if $x \in A$. Then, $\overline{E} \cap A$ is contained in the closure of E relative to $(A, \mathcal{O}_A)$. Hence, the result holds.

This completes the proof of the proposition. $\square$

Definition 3.6 For two topologies over the same set X , $\mathcal{O}_1$ and $\mathcal{O}_2$, we will say that $\mathcal{O}_1$ is stronger (finer) than $\mathcal{O}_2$ if $\mathcal{O}_1 \supset \mathcal{O}_2$, in which case, $\mathcal{O}_2$ is said to be weaker (coarser) than $\mathcal{O}_1$.

Proposition 3.7 Let X be a set and $\mathcal{A} \subseteq {}^X 2$. Then, there exists the weakest topology $\mathcal{O}$ on X such that $\mathcal{A} \subseteq \mathcal{O}$. This topology is called the topology generated by $\mathcal{A}$.

Proof Let $\mathcal{M} := \{ \mathcal{X} \subseteq {}^X 2 \mid \mathcal{A} \subseteq \mathcal{X} \text{ and } \mathcal{X} \text{ is a topology on } X \}$ and $\mathcal{O} = \bigcap_{\mathcal{X} \in \mathcal{M}} \mathcal{X}$. Clearly, ${}^X 2 \in \mathcal{M}$ and hence $\mathcal{O}$ is well-defined. Then,

- (i) $\emptyset, X \in \mathcal{X}, \forall \mathcal{X} \in \mathcal{M}$. Hence, $\emptyset, X \in \mathcal{O}$.
- (ii) $\forall A_1, A_2 \in \mathcal{O}$, we have $A_1, A_2 \in \mathcal{X}, \forall \mathcal{X} \in \mathcal{M}$. Then, $A_1 \cap A_2 \in \mathcal{X}, \forall \mathcal{X} \in \mathcal{M}$. Hence, $A_1 \cap A_2 \in \mathcal{O}$.
- (iii) $\forall (A_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}$, where Λ is an index set, we have, $\forall \alpha \in \Lambda, \forall \mathcal{X} \in \mathcal{M}, A_\alpha \in \mathcal{X}$. Then, $\bigcup_{\alpha \in \Lambda} A_\alpha \in \mathcal{X}, \forall \mathcal{X} \in \mathcal{M}$. Hence, we have $\bigcup_{\alpha \in \Lambda} A_\alpha \in \mathcal{O}$.

Therefore, $\mathcal{O}$ is a topology on X . Clearly, $\mathcal{A} \subseteq \mathcal{O}$ since $\mathcal{A} \subseteq \mathcal{X}, \forall \mathcal{X} \in \mathcal{M}$. Therefore, $\mathcal{O}$ is the weakest topology containing $\mathcal{A}$. $\square$

3.2 Continuity

Definition 3.8 Let $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ be topological spaces, $D \subseteq X$ with the subset topology $\mathcal{O}_D$, and $f : D \rightarrow Y$ (or $f : (D, \mathcal{O}_D) \rightarrow (Y, \mathcal{O}_Y)$ to be more specific). Then, f is said to be continuous if, $\forall O_Y \in \mathcal{O}_Y$, we have $f_{\text{inv}}(O_Y) \in \mathcal{O}_D$. f is said to be continuous at $x_0 \in D$ if, $\forall O_Y \in \mathcal{O}_Y$ with $f(x_0) \in O_Y$, $\exists U \in \mathcal{O}_X$ with $x_0 \in U$ such that $f(U) \subseteq O_Y$. f is said to be continuous on $E \subseteq D$ if it is continuous at $x, \forall x \in E$.

Proposition 3.9 Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces, $D \subseteq \mathcal{X}$ with the subset topology $\mathcal{O}_D$, and $f : D \rightarrow \mathcal{Y}$. f is continuous if, and only if, $\forall x_0 \in D$, f is continuous at x_0 .

Proof “If” $\forall O_Y \in \mathcal{O}_Y, \forall x \in f_{\text{inv}}(O_Y) \subseteq D$. Since f is continuous at x , then $\exists U_x \in \mathcal{O}_X$ with $x \in U_x$ such that $f(U_x) \subseteq O_Y$, which implies, by Proposition 2.5, that $U_x \cap D \subseteq f_{\text{inv}}(O_Y)$. Then,

$f_{\text{inv}}(O_Y) = \bigcup_{x \in f_{\text{inv}}(O_Y)} (U_x \cap D) = (\bigcup_{x \in f_{\text{inv}}(O_Y)} U_x) \cap D \in \mathcal{O}_D$. Hence, f is continuous.

“Only if” $\forall x_0 \in D, \forall O_Y \in \mathcal{O}_Y$ with $f(x_0) \in O_Y$, let $U = f_{\text{inv}}(O_Y) \in \mathcal{O}_D$. By Proposition 3.4, $\exists \bar{U} \in \mathcal{O}_X$ such that $U = \bar{U} \cap D$. Then, $x_0 \in U$. By Proposition 2.5, $f(\bar{U}) = f(U) \subseteq O_Y$. Hence, f is continuous at x_0 .

This completes the proof of the proposition. $\square$

Proposition 3.10 *Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces and $f : \mathcal{X} \rightarrow \mathcal{Y}$. f is continuous if, and only if, $\forall B \subseteq \mathcal{Y}$ with $\widetilde{B} \in \mathcal{O}_Y$, we have $f_{\text{inv}}(B) \in \mathcal{O}_X$, that is, the inverse image of any closed set in $\mathcal{Y}$ is closed in $\mathcal{X}$.*

Proof “If” $\forall O \in \mathcal{O}_Y$, we have, by Proposition 2.5, $f_{\text{inv}}(O) = \widetilde{f_{\text{inv}}(O)} \in \mathcal{O}_X$. Hence, f is continuous.

“Only if” $\forall B \subseteq \mathcal{Y}$ with $\widetilde{B} \in \mathcal{O}_Y$. Since f is continuous, then, by Proposition 2.5, $\widetilde{f_{\text{inv}}(B)} = f_{\text{inv}}(\widetilde{B}) \in \mathcal{O}_X$. Hence, the result holds.

This completes the proof of the proposition. $\square$

Theorem 3.11 *Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces, $f : \mathcal{X} \rightarrow \mathcal{Y}$, and $\mathcal{X} = X_1 \cup X_2$, where X_1 and X_2 are both open or both closed. Let X_1 and X_2 be endowed with subset topologies $\mathcal{O}_{X_1}$ and $\mathcal{O}_{X_2}$, respectively. Assume that $f|_{X_1} : X_1 \rightarrow \mathcal{Y}$ and $f|_{X_2} : X_2 \rightarrow \mathcal{Y}$ are continuous. Then, f is continuous.*

Proof Consider the case that X_1 and X_2 are both open. $\forall x_0 \in \mathcal{X}$. Without loss of generality, assume $x_0 \in X_1$. $\forall O \in \mathcal{O}_Y$ with $f(x_0) \in O$. Since $f|_{X_1}$ is continuous, then, by Proposition 3.9, $\exists U \in \mathcal{O}_{X_1}$ with $x_0 \in U$ such that $f|_{X_1}(U) \subseteq O$. Since $X_1 \in \mathcal{O}_X$, then $U \in \mathcal{O}_X$. Note that $f(U) = f|_{X_1}(U) \subseteq O$, since $U \subseteq X_1$. Hence, f is continuous at x_0 . By the arbitrariness of x_0 and Proposition 3.9, f is continuous.

Consider the case that X_1 and X_2 are both closed. $\forall$ closed subset $B \subseteq \mathcal{Y}$, we have $f_{\text{inv}}(B) \subseteq X$. Then, $f_{\text{inv}}(B) \cap X_1 = (f|_{X_1})_{\text{inv}}(B)$ is closed in $\mathcal{O}_{X_1}$, by Proposition 3.10 and the continuity of $f|_{X_1}$. Similarly, $f_{\text{inv}}(B) \cap X_2 = (f|_{X_2})_{\text{inv}}(B)$ is closed in $\mathcal{O}_{X_2}$. Since X_1 and X_2 are closed sets in $\mathcal{O}_X$, then, $f_{\text{inv}}(B) \cap X_1$ and $f_{\text{inv}}(B) \cap X_2$ are closed in $\mathcal{O}_X$, by Proposition 3.5. Then, $f_{\text{inv}}(B) = (f_{\text{inv}}(B) \cap X_1) \cup (f_{\text{inv}}(B) \cap X_2)$ is closed in $\mathcal{O}_X$. By Proposition 3.10, f is continuous.

This completes the proof of the theorem. $\square$

Proposition 3.12 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be topological spaces, $f : \mathcal{X} \rightarrow \mathcal{Y}$, $g : \mathcal{Y} \rightarrow \mathcal{Z}$, and $x_0 \in \mathcal{X}$. Assume that f is continuous at x_0 and g is continuous at $y_0 := f(x_0)$. Then, $g \circ f : \mathcal{X} \rightarrow \mathcal{Z}$ is continuous at x_0 .*

Proof $\forall O_Z \in \mathcal{O}_Z$ with $g(f(x_0)) \in O_Z$. Since g is continuous at $f(x_0)$, then $\exists O_Y \in \mathcal{O}_Y$ with $f(x_0) \in O_Y$ such that $g(O_Y) \subseteq O_Z$. Since f is continuous at x_0 , then $\exists O_X \in \mathcal{O}_X$ with $x_0 \in O_X$ such that $f(O_X) \subseteq O_Y$. Then, $g(f(O_X)) \subseteq O_Z$. Hence, $g \circ f$ is continuous at x_0 . This completes the proof of the proposition. $\square$

Definition 3.13 Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces and $f : \mathcal{X} \rightarrow \mathcal{Y}$. f is said to be a homeomorphism between $\mathcal{X}$ and $\mathcal{Y}$ if it is bijective and continuous and $f_{\text{inv}} : \mathcal{Y} \rightarrow \mathcal{X}$ is also continuous. The spaces $\mathcal{X}$ and $\mathcal{Y}$ are said to be homeomorphic if there exists a homeomorphism between them.

Any properties invariant under homeomorphisms are called *topological properties*.

Homeomorphisms preserve topological properties in topological spaces. Isomorphisms preserve algebraic properties in algebraic systems. Isometries preserve metric properties in metric spaces.

Definition 3.14 Let $\mathcal{X}$ be a topological space, $D \subseteq \mathcal{X}$ with the subset topology $\mathcal{O}_D$, and $f : D \rightarrow \mathbb{R}$. f is said to be upper semicontinuous if $\forall a \in \mathbb{R}, f_{\text{inv}}((-\infty, a)) \in \mathcal{O}_D$. f is said to be upper semicontinuous at $x_0 \in \mathcal{X}$ if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists U \in \mathcal{O}_{\mathcal{X}}$ with $x_0 \in U$ such that $f(x) < f(x_0) + \epsilon, \forall x \in U \cap D$. f is said to be lower semicontinuous if $-f$ is upper semicontinuous.

Proposition 3.15 Let $\mathcal{X}$ be a topological space, $D \subseteq \mathcal{X}$ with the subset topology $\mathcal{O}_D$, and $f : D \rightarrow \mathbb{R}$. f is upper semicontinuous if, and only if, f is upper semicontinuous at $x_0, \forall x_0 \in D$.

Proof This is straightforward, and is therefore omitted. □

Proposition 3.16 Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces, $f : \mathcal{X} \rightarrow \mathbb{R}$ and $g : \mathcal{X} \rightarrow \mathbb{R}$ be upper semicontinuous at $x_0 \in \mathcal{X}$, $h : \mathcal{Y} \rightarrow \mathcal{X}$ be continuous at $y_0 \in \mathcal{Y}$, and $h(y_0) = x_0$. Then, $f + g$ is upper semicontinuous at x_0 and $f \circ h$ is upper semicontinuous at y_0 . Furthermore, if f is also lower semicontinuous at x_0 , then f is continuous at x_0 .

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists U_f \in \mathcal{O}_{\mathcal{X}}$ with $x_0 \in U_f$ such that $f(x) < f(x_0) + \epsilon, \forall x \in U_f$, by the upper semicontinuity of f . By the upper semicontinuity of $g, \exists U_g \in \mathcal{O}_{\mathcal{X}}$ with $x_0 \in U_g$ such that $g(x) < g(x_0) + \epsilon, \forall x \in U_g$. Then, $x_0 \in U := U_f \cap U_g \in \mathcal{O}_{\mathcal{X}}$ and $(f + g)(x) = f(x) + g(x) < f(x_0) + \epsilon + g(x_0) + \epsilon = (f + g)(x_0) + 2\epsilon, \forall x \in U$. Hence, $f + g$ is upper semicontinuous at x_0 .

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists U_f \in \mathcal{O}_{\mathcal{X}}$ with $x_0 \in U_f$ such that $f(x) < f(x_0) + \epsilon, \forall x \in U_f$, by the upper semicontinuity of f . By the continuity of $h, \exists U_h \in \mathcal{O}_{\mathcal{Y}}$ with $y_0 \in U_h$ such that $h(y) \in U_f, \forall y \in U_h$. Then, $(f \circ h)(y) < f(h(y)) + \epsilon, \forall y \in U_h$. Hence, $f \circ h$ is upper semicontinuous at y_0 .

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the upper semicontinuity of $f, \exists U_1 \in \mathcal{O}_{\mathcal{X}}$ with $x_0 \in U_1$ such that $f(x) < f(x_0) + \epsilon, \forall x \in U_1$. By the lower semicontinuity of $f, \exists U_2 \in \mathcal{O}_{\mathcal{X}}$ with $x_0 \in U_2$ such that $f(x) > f(x_0) - \epsilon, \forall x \in U_2$. Then, $x_0 \in U := U_1 \cap U_2 \in \mathcal{O}_{\mathcal{X}}$ and $|f(x) - f(x_0)| < \epsilon, \forall x \in U$. Hence, f is continuous at x_0 .

This completes the proof of the proposition. □

3.3 Basis and Countability

Definition 3.17 Let $(X, \mathcal{O})$ be a topological space and $\mathcal{B} \subseteq \mathcal{O}$. $\mathcal{B}$ is said to be a basis of the topological space if, $\forall O \in \mathcal{O}, \forall x \in O, \exists B \in \mathcal{B}$ such that $x \in B \subseteq O$. $\mathcal{B}_x \subseteq \mathcal{O}$ with $x \in B, \forall B \in \mathcal{B}_x$, is said to be a basis at $x \in X$ if $\forall O \in \mathcal{O}$ with $x \in O, \exists B \in \mathcal{B}_x$ such that $x \in B \subseteq O$.

If $\mathcal{B}$ is a basis for the topology $\mathcal{O}$, then the topology generated by $\mathcal{B}$ is $\mathcal{O}$ and, $\forall O \in \mathcal{O}, O = \bigcup_{x \in O} B_x$, where $B_x \in \mathcal{B}$ is such that $x \in B_x \subseteq O$.

Proposition 3.18 Let X be a set and $\mathcal{B} \subseteq {}^X 2$. Then, $\mathcal{B}$ is a basis for the topology generated by $\mathcal{B}$, $\mathcal{O}$, if, and only if, the following two conditions hold:

- (i) $\forall x \in X, \exists B \in \mathcal{B}$ such that $x \in B$; (that is, $\bigcup_{B \in \mathcal{B}} B = X$);
- (ii) $\forall B_1, B_2 \in \mathcal{B}, \forall x \in B_1 \cap B_2, \exists B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.

Proof “Only if” Let $\mathcal{B}$ be a basis for $\mathcal{O}$. Since $X \in \mathcal{O}$, then $\forall x \in X, \exists B \in \mathcal{B}$ such that $x \in B \subseteq X$. So (i) is true. $\forall B_1, B_2 \in \mathcal{B} \subseteq \mathcal{O}$, we have $B_1 \cap B_2 \in \mathcal{O}$. $\forall x \in B_1 \cap B_2, \exists B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$. Hence, (ii) is also true.

“If” Define $\bar{\mathcal{O}} := \{O \subseteq X \mid \forall x \in O, \exists B \in \mathcal{B} \text{ such that } x \in B \subseteq O\}$. Clearly, $\mathcal{B} \subseteq \bar{\mathcal{O}}$ and $\emptyset \in \bar{\mathcal{O}}$. $X \in \bar{\mathcal{O}}$ since (i). $\forall O_1, O_2 \in \bar{\mathcal{O}}, \forall x \in O_1 \cap O_2$, we have $x \in O_1$ and $x \in O_2$. Then, by the definition of $\bar{\mathcal{O}}$, $\exists B_1, B_2 \in \mathcal{B}$ such that $x \in B_1 \subseteq O_1$ and $x \in B_2 \subseteq O_2$. Then, $x \in B_1 \cap B_2$. By (ii), $\exists B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2 \subseteq O_1 \cap O_2$. Then, $O_1 \cap O_2 \in \bar{\mathcal{O}}$. $\forall (O_\alpha)_{\alpha \in \Lambda} \subseteq \bar{\mathcal{O}}$, where Λ is an index set, let $O = \bigcup_{\alpha \in \Lambda} O_\alpha$. $\forall x \in O, \exists \alpha \in \Lambda$ such that $x \in O_\alpha$. By the definition of $\bar{\mathcal{O}}$, $\exists B \in \mathcal{B}$ such that $x \in B \subseteq O_\alpha \subseteq O$. Hence, $O \in \bar{\mathcal{O}}$. Therefore, $\bar{\mathcal{O}}$ is a topology on X . By the definition of $\bar{\mathcal{O}}$, $\mathcal{B}$ is a basis for $\bar{\mathcal{O}}$. Note that $\bar{\mathcal{O}} \supseteq \mathcal{O}$ since $\mathcal{O}$ is the weakest topology containing $\mathcal{B}$. Then, $\mathcal{B}$ is a basis for $\mathcal{O}$.

This completes the proof of the proposition. $\square$

Example 3.19 For the real line $\mathbb{R}$, let $\mathcal{A} := \{\text{interval } (a, b) \mid a, b \in \mathbb{R}, a < b\}$. Then, the topology generated by $\mathcal{A}$, $\mathcal{O}_{\mathbb{R}}$, is the usual topology on $\mathbb{R}$ as we know before. By Proposition 3.18, $\mathcal{A}$ is a basis for this topology. $\diamond$

Definition 3.20 A topological space $(X, \mathcal{O})$ is said to satisfy the first axiom of countability if there exists a countable basis at each $x \in X$, i. e., $\forall x \in X, \exists \mathcal{B}_x \subseteq \mathcal{O}$, which is countable, such that, $\forall O \in \mathcal{O}$ with $x \in O$, we have $\exists B \in \mathcal{B}_x$ with $x \in B \subseteq O$; and $\forall B \in \mathcal{B}_x$, we have $x \in B$. In this case, we will say that the topological space is first countable. The topological space is said to satisfy the second axiom of countability if there exists a countable basis $\mathcal{B}$ for $\mathcal{O}$, in which case, we will say that it is second countable.

Clearly, a second countable topological space is also first countable.

Example 3.21 The real line is first countable, where a countable basis at any $x \in \mathbb{R}$ consists of intervals of the form $(x - r, x + r)$ with $r \in \mathbb{Q}$ and $r > 0$. The real line is also second countable, where a countable basis for the topology consists of intervals of the form (r_1, r_2) with $r_1, r_2 \in \mathbb{Q}$ and $r_1 < r_2$. $\diamond$

When basis are available on topological spaces $\mathcal{X}$ and $\mathcal{Y}$, in Definitions 3.8 and 3.14, we may restrict the open sets O_Y and U to be basis open sets without changing the meaning of the definition. In Proposition 3.3, we may restrict O to be a basis open set and the results still hold.

Definition 3.22 A collection $(A_\alpha)_{\alpha \in \Lambda}$ of sets, where Λ is an index set, is said to be a covering of a set X if $X \subseteq \bigcup_{\alpha \in \Lambda} A_\alpha$. It is an open covering if A_α 's are open sets in a specific topological space.

Definition 3.23 A topological space $\mathcal{X}$ is said to be Lindelöf if any open covering of $\mathcal{X}$ has a countable subcovering.

Proposition 3.24 A second countable topological space $\mathcal{X}$ is Lindelöf.

Proof Let $(B_i)_{i \in N}$ be a countable basis for $\mathcal{X}$, where N is a countable index set. Let $(O_\alpha)_{\alpha \in \Lambda}$ be an open covering of $\mathcal{X}$, where Λ is an index set. $\forall \alpha \in \Lambda$, $O_\alpha = \bigcup_{i \in N_\alpha} B_i$, where $N_\alpha \subseteq N$. Then, $\mathcal{X} = \bigcup_{\alpha \in \Lambda} \bigcup_{i \in N_\alpha} B_i$. We may determine a subcollection $(B_i)_{i \in \bar{N}}$ where $\bar{N} \subseteq N$ such that these B_i 's appears at least once in the previous union. Then, $\mathcal{X} = \bigcup_{i \in \bar{N}} B_i$. $\forall i \in \bar{N}$, the collection $\mathcal{A}_i := \{O_\alpha \mid \alpha \in \Lambda, B_i \subseteq O_\alpha\}$ is nonempty. Then, by Axiom of Choice, there exists an assignment $(O_{\alpha_i})_{i \in \bar{N}}$ such that $O_{\alpha_i} \in \mathcal{A}_i$, $\forall i \in \bar{N}$. Then, we have $(O_{\alpha_i})_{i \in \bar{N}}$ is a countable subcover. This completes the proof of the proposition. $\square$

3.4 Products of Topological Spaces

Proposition 3.25 Let Λ be an index set, and $(X_\alpha, \mathcal{O}_\alpha)$ be a topological space, $\forall \alpha \in \Lambda$. The product topology $\mathcal{O}$ on $\prod_{\alpha \in \Lambda} X_\alpha$ is the topology generated by

$$\mathcal{B} := \left\{ \prod_{\alpha \in \Lambda} O_\alpha \mid O_\alpha \in \mathcal{O}_\alpha, \forall \alpha \in \Lambda, \text{ and } O_\alpha = X_\alpha \text{ for all but a finite number of } \alpha \text{'s} \right\}$$

The collection $\mathcal{B}$ forms a basis for the topology $\mathcal{O}$. We will also write $(\prod_{\alpha \in \Lambda} X_\alpha, \mathcal{O}) = \prod_{\alpha \in \Lambda} (X_\alpha, \mathcal{O}_\alpha)$. When $(X_\alpha, \mathcal{O}_\alpha) = (X, \mathcal{O}) =: \mathcal{X}$, $\forall \alpha \in \Lambda$, we will denote $\prod_{\alpha \in \Lambda} (X_\alpha, \mathcal{O}_\alpha)$ by $\mathcal{X}^\Lambda$.

Proof We will prove this by Proposition 3.18. Since $\prod_{\alpha \in \Lambda} X_\alpha \in \mathcal{B}$, then (i) is true. $\forall B_1, B_2 \in \mathcal{B}$, $B_1 = \prod_{\alpha \in \Lambda} O_{1\alpha}$, $O_{1\alpha} \in \mathcal{O}_\alpha$, $\forall \alpha \in \Lambda$, and $O_{1\alpha} = X_\alpha$, $\forall \alpha \in \Lambda \setminus \Lambda_1$, where $\Lambda_1 \subseteq \Lambda$ is a finite set; $B_2 = \prod_{\alpha \in \Lambda} O_{2\alpha}$, $O_{2\alpha} \in \mathcal{O}_\alpha$, $\forall \alpha \in \Lambda$, and $O_{2\alpha} = X_\alpha$, $\forall \alpha \in \Lambda \setminus \Lambda_2$, where $\Lambda_2 \subseteq \Lambda$ is a finite set. Let $B_3 = B_1 \cap B_2 = \prod_{\alpha \in \Lambda} (O_{1\alpha} \cap O_{2\alpha})$. Clearly, $O_{1\alpha} \cap O_{2\alpha} \in \mathcal{O}_\alpha$, $\forall \alpha \in \Lambda$, and $O_{1\alpha} \cap O_{2\alpha} = X_\alpha$, $\forall \alpha \in \Lambda \setminus (\Lambda_1 \cup \Lambda_2)$, where $\Lambda_1 \cup \Lambda_2 \subseteq \Lambda$ is a finite set. Hence, $B_3 \in \mathcal{B}$. Then, (ii) also holds. By Proposition 3.18, $\mathcal{B}$ is a basis for $\mathcal{O}$. This completes the proof of the proposition. $\square$

Definition 3.26 Let X_α be a topological space, $\forall \alpha \in \Lambda$, where Λ is an index set, and $\mathcal{Y} = \prod_{\alpha \in \Lambda} X_\alpha$ be the product topological space. Define a collection of projection functions $\pi_\alpha : \mathcal{Y} \rightarrow X_\alpha$, $\forall \alpha \in \Lambda$, by $\pi_\alpha(y) = y_\alpha$ for $y = (y_\alpha)_{\alpha \in \Lambda} \in \mathcal{Y}$.

Proposition 3.27 Let $(X_\alpha, \mathcal{O}_\alpha)$ be a topological space, $\forall \alpha \in \Lambda$, where Λ is an index set, and $(Y, \mathcal{O}) = \prod_{\alpha \in \Lambda} (X_\alpha, \mathcal{O}_\alpha)$ be the product topological space. Then, $\mathcal{O}$ is the weakest topology on which the projection functions π_α , $\forall \alpha \in \Lambda$, are continuous.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\Lambda = \emptyset$; Case 2: $\Lambda \neq \emptyset$. Case 1: $\Lambda = \emptyset$. $(Y, \mathcal{O}) = (\{\emptyset\}, \{\emptyset, \{\emptyset\}\})$. Then, $\mathcal{O}$ is the only topology on Y . Hence, the result is true.

Case 2: $\Lambda \neq \emptyset$. Define

$$\mathcal{A} := \left\{ O \subseteq Y \mid O = \prod_{\alpha \in \Lambda} O_\alpha, \exists \alpha_0 \in \Lambda \text{ such that } O_{\alpha_0} \in \mathcal{O}_{\alpha_0} \text{ and } O_\alpha = X_\alpha, \forall \alpha \in \Lambda \setminus \{\alpha_0\} \right\}$$

Clearly, $\mathcal{A} \neq \emptyset$ is the collection of inverse images of open sets under π_α , $\forall \alpha \in \Lambda$. It is also clear that $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$, where $\mathcal{B}$ is the basis for $\mathcal{O}$ that we introduced in Proposition 3.25. $\forall B \in \mathcal{B}$, we have $B = \bigcap_{i=1}^k A_i$ for some $k \in \mathbb{N}$ and $A_1, \dots, A_k \in \mathcal{A}$. Hence, the topology generated by $\mathcal{A}$ contains $\mathcal{B}$. Then, the topology generated by $\mathcal{A}$ equals to $\mathcal{O}$. Hence, $\mathcal{O}$ is the weakest topology containing $\mathcal{A}$. This completes the proof of the proposition. $\square$

Proposition 3.28 Let $\mathcal{X}_i := (X_i, \mathcal{O}_i)$ be a second countable topological space, $\forall i \in N \subseteq \mathbb{N}$, and N is a countable index set. Then, the product topological space $\mathcal{X} := (X, \mathcal{O}) := \prod_{i \in N} \mathcal{X}_i$ is second countable.

Proof $\forall i \in N$, let $\mathcal{B}_i \subseteq \mathcal{O}_i$ be a countable basis for $\mathcal{X}_i$, without loss of generality, assume $X_i \in \mathcal{B}_i$. Let $\mathcal{B}$ be the basis for the product topological space $\mathcal{X}$ as defined in Proposition 3.25. Define

$$\bar{\mathcal{B}} := \left\{ \prod_{i \in N} B_i \subseteq X \mid B_i \in \mathcal{B}_i, \forall i \in N, \text{ and } B_i = X_i \text{ for all but a finite number of } i\text{'s} \right\} \subseteq \mathcal{B} \subseteq \mathcal{O}$$

Since $\mathcal{B}_i$'s are countable, then $\bar{\mathcal{B}}$ is countable. We will show that $\bar{\mathcal{B}}$ is a basis for $\mathcal{O}$. $\forall O \in \mathcal{O}$, $\forall x \in O$, $\exists O_B \in \mathcal{B}$ such that $x \in O_B \subseteq O$, where $O_B = \prod_{i \in N} O_{Bi}$, $O_{Bi} \in \mathcal{O}_i$, $\forall i \in N$, $O_{Bi} = X_i$, $\forall i \in N \setminus \bar{N}$, and $\bar{N} \subseteq N$ is a finite set. $\forall i \in \bar{N}$, $\pi_i(x) \in O_{Bi}$. Then, $\exists B_i \in \mathcal{B}_i$ such that $\pi_i(x) \in B_i \subseteq O_{Bi}$. Let $B_i := X_i$, $\forall i \in N \setminus \bar{N}$. Let $B := \prod_{i \in N} B_i \in \bar{\mathcal{B}}$. Clearly, $x \in B \subseteq O_B \subseteq O$. Hence, $\bar{\mathcal{B}}$ is a basis for $\mathcal{X}$. Then, $\mathcal{X}$ is second countable. This completes the proof of the proposition. $\square$

Proposition 3.29 *Let $\mathcal{X}_\alpha := (X_\alpha, \mathcal{O}_\alpha)$ be a topological space, $\forall \alpha \in \Lambda$, where Λ is an index set. Let $\mathcal{X} = (X, \mathcal{O}) := \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha$ be the product space. Assume that $A_\alpha \subseteq \mathcal{X}_\alpha$, $\forall \alpha \in \Lambda$. Then, $\overline{\prod_{\alpha \in \Lambda} A_\alpha} = \prod_{\alpha \in \Lambda} \overline{A_\alpha}$.*

Proof Note that

$$\left(\prod_{\alpha \in \Lambda} \overline{A_\alpha} \right)^\sim = \bigcup_{\alpha \in \Lambda} \pi_{\alpha \text{ inv}}(\widetilde{\overline{A_\alpha}})$$

where $\pi_\alpha : \mathcal{X} \rightarrow \mathcal{X}_\alpha$ is the projection function, $\forall \alpha \in \Lambda$. $\forall \alpha \in \Lambda$, by Definition 3.2, $\widetilde{\overline{A_\alpha}} \in \mathcal{O}_\alpha$. By Proposition 3.27, π_α is continuous. Then, $\pi_{\alpha \text{ inv}}(\widetilde{\overline{A_\alpha}}) \in \mathcal{O}$. Then, $\prod_{\alpha \in \Lambda} \overline{A_\alpha}$ is a closed set in $\mathcal{O}$. Clearly, $\prod_{\alpha \in \Lambda} A_\alpha \subseteq \prod_{\alpha \in \Lambda} \overline{A_\alpha}$. Then, we have $\overline{\prod_{\alpha \in \Lambda} A_\alpha} \subseteq \prod_{\alpha \in \Lambda} \overline{A_\alpha}$.

On the other hand, $\forall x \in \prod_{\alpha \in \Lambda} \overline{A_\alpha}$, $\forall$ basis open set $O \in \mathcal{O}$ with $x \in O$, by Proposition 3.25, we have $O = \prod_{\alpha \in \Lambda} O_\alpha$, where $O_\alpha \in \mathcal{O}_\alpha$, $\forall \alpha \in \Lambda$, and $O_\alpha = X_\alpha$ for all but finitely many α 's. $\forall \alpha \in \Lambda$, $\pi_\alpha(x) \in O_\alpha$ and $\pi_\alpha(x) \in \overline{A_\alpha}$. By Proposition 3.3, $O_\alpha \cap A_\alpha \neq \emptyset$. Then, by Axiom of Choice, we have $O \cap (\prod_{\alpha \in \Lambda} A_\alpha) = \prod_{\alpha \in \Lambda} (O_\alpha \cap A_\alpha) \neq \emptyset$. Thus, by Proposition 3.3, $x \in \overline{\prod_{\alpha \in \Lambda} A_\alpha}$. Hence, we have $\prod_{\alpha \in \Lambda} \overline{A_\alpha} \subseteq \overline{\prod_{\alpha \in \Lambda} A_\alpha}$.

Therefore, $\overline{\prod_{\alpha \in \Lambda} A_\alpha} = \prod_{\alpha \in \Lambda} \overline{A_\alpha}$. This completes the proof of the proposition. $\square$

An immediate consequence of the above proposition is that if $A_\alpha \subseteq \mathcal{X}_\alpha$ is closed, $\forall \alpha \in \Lambda$, where $\mathcal{X}_\alpha$'s are topological spaces, then $\prod_{\alpha \in \Lambda} A_\alpha$ is closed in the product topology.

Proposition 3.30 *Let $(\Lambda_\beta)_{\beta \in \Gamma}$ be a collection of pairwise disjoint sets, where Γ is an index set. Let $(X_\alpha, \mathcal{O}_\alpha)$ be a topological space, $\forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta$. $\forall \beta \in \Gamma$, $(\prod_{\alpha \in \Lambda_\beta} X_\alpha, \mathcal{O}^{(\beta)})$ is the product space with the product topology. $(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, \mathcal{O}^{(\Gamma)})$ is the product space of product spaces with the product topology. Let $(\prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha, \mathcal{O})$ be the product space with the product topology. Then, $(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, \mathcal{O}^{(\Gamma)})$ and $(\prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha, \mathcal{O})$ are homeomorphic.*

Proof Define a mapping $E : \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \rightarrow \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha$ by, $\forall x \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$, $\forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta$, $\exists! \beta_\alpha \in \Gamma \ni \alpha \in \Lambda_{\beta_\alpha}$, $\pi_\alpha(E(x)) = \pi_{\alpha}^{(\beta_\alpha)}(\pi_{\beta_\alpha}^{(\Gamma)}(x))$.

$\forall y \in \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha$. $\forall \beta \in \Gamma$, define $x_\beta \in \prod_{\alpha \in \Lambda_\beta} X_\alpha$ be such that $\pi_\alpha^{(\beta)}(x_\beta) = \pi_\alpha(y)$, $\forall \alpha \in \Lambda_\beta$. Define $x \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$ by $\pi_\beta^{(\Gamma)}(x) = x_\beta$, $\forall \beta \in \Gamma$. Then, $\forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta$, $\exists! \beta_\alpha \ni \alpha \in \Lambda_{\beta_\alpha}$, and $\pi_\alpha(E(x)) = \pi_\alpha^{(\beta_\alpha)}(\pi_{\beta_\alpha}^{(\Gamma)}(x)) = \pi_\alpha(y)$ and hence $E(x) = y$. Hence, E is surjective.

$\forall x, z \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$ with $E(x) = E(z)$. $\forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta$, $\exists! \beta_\alpha \ni \alpha \in \Lambda_{\beta_\alpha}$. Then, $\pi_\alpha^{(\beta_\alpha)}(\pi_{\beta_\alpha}^{(\Gamma)}(x)) = \pi_\alpha(E(x)) = \pi_\alpha(E(z)) = \pi_\alpha^{(\beta_\alpha)}(\pi_{\beta_\alpha}^{(\Gamma)}(z))$. Hence, $\forall \beta \in \Gamma$, $\forall \alpha \in \Lambda_\beta$, $\pi_\alpha^{(\beta)}(\pi_\beta^{(\Gamma)}(x)) = \pi_\alpha^{(\beta)}(\pi_\beta^{(\Gamma)}(z))$ which implies that $\pi_\beta^{(\Gamma)}(x) = \pi_\beta^{(\Gamma)}(z)$. Hence, $x = z$. This shows that E is injective.

Hence, E is bijective and admits an inverse E_{inv} .

Next, we show that E is continuous by showing that E is continuous at x_0 , $\forall x_0 \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$. $\forall x_0 \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$. Let $y_0 = E(x_0)$. $\forall B \in \mathcal{O}$ which is a basis open set with $y_0 \in B$. Then, $B = \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} B_\alpha$, $B_\alpha \in \mathcal{O}_\alpha$, $\forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta$, and $B_\alpha = X_\alpha$ for all α 's except finitely many α 's, say $\alpha \in \Lambda_N$. $\forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta$, $\exists! \beta_\alpha \in \Gamma \ni \alpha \in \Lambda_{\beta_\alpha}$. Let $\Gamma_N = \{\beta_\alpha \mid \alpha \in \Lambda_N\}$. Then, Γ_N is a finite set. $\forall \beta \in \Gamma_N$, let $\Lambda_{N\beta} := \Lambda_N \cap \Lambda_\beta$, which is a nonempty finite set. Then, Λ_N equals to the disjoint union of $\bigcup_{\beta \in \Gamma_N} \Lambda_{N\beta}$. Define $B^{(\beta)} := \prod_{\alpha \in \Lambda_\beta} B_\alpha$, $\forall \beta \in \Gamma$. Then, $\forall \beta \in \Gamma$, $B^{(\beta)}$ is a basis open set in $\left(\prod_{\alpha \in \Lambda_\beta} X_\alpha, \mathcal{O}^{(\beta)}\right)$. Define $B^{(\Gamma)} := \prod_{\beta \in \Gamma} B^{(\beta)}$. $\forall \beta \in \Gamma \setminus \Gamma_N$, $B^{(\beta)} = \prod_{\alpha \in \Lambda_\beta} X_\alpha$. Clearly, $B^{(\Gamma)}$ is a basis open set in $\mathcal{O}^{(\Gamma)}$. $\forall \beta \in \Gamma_N$, $\forall \alpha \in \Lambda_{N\beta}$, $\pi_\alpha^{(\beta)}(\pi_\beta^{(\Gamma)}(x_0)) = \pi_\alpha(y_0) \in \pi_\alpha(B) = B_\alpha$. Then, $x_0 \in B^{(\Gamma)}$. $\forall x \in B^{(\Gamma)}$, $\forall \beta \in \Gamma_N$, $\forall \alpha \in \Lambda_{N\beta}$, we have $\pi_\alpha(E(x)) = \pi_\alpha^{(\beta)}(\pi_\beta^{(\Gamma)}(x)) \in \pi_\alpha^{(\beta)}(\pi_\beta^{(\Gamma)}(B^{(\Gamma)})) = \pi_\alpha^{(\beta)}(B^{(\beta)}) = B_\alpha$. Hence, $E(x) \in B$, which implies that $E(B^{(\Gamma)}) \subseteq B$. Hence, E is continuous at x_0 . Then, by the arbitrariness of x_0 and Proposition 3.9, E is continuous.

Finally, we will show E_{inv} is continuous by showing that, $\forall y_0 \in \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha$, E_{inv} is continuous at y_0 . $\forall y_0 \in \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha$. Let $x_0 = E_{\text{inv}}(y_0) \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$. For any basis open set $B^{(\Gamma)} \in \mathcal{O}^{(\Gamma)}$ with $x_0 \in B^{(\Gamma)}$. Then, $B^{(\Gamma)} = \prod_{\beta \in \Gamma} \bar{B}^{(\beta)}$, $\bar{B}^{(\beta)} \in \mathcal{O}^{(\beta)}$, $\forall \beta \in \Gamma$, and $\bar{B}^{(\beta)} = \prod_{\alpha \in \Lambda_\beta} X_\alpha$ for all $\beta \in \Gamma$ except finitely many β 's, say $\beta \in \Gamma_N$. $\forall \beta \in \Gamma_N$, $\pi_\beta^{(\Gamma)}(x_0) \in \bar{B}^{(\beta)}$. Then, there exists a basis open set $B^{(\beta)} \in \mathcal{O}^{(\beta)}$ such that $\pi_\beta^{(\Gamma)}(x_0) \in B^{(\beta)} \subseteq \bar{B}^{(\beta)}$. Then, $B^{(\beta)} = \prod_{\alpha \in \Lambda_\beta} B_\alpha$, $B_\alpha \in \mathcal{O}_\alpha$, $\forall \alpha \in \Lambda_\beta$, and $B_\alpha = X_\alpha$ for all $\alpha \in \Lambda_\beta$ except finitely many α 's, say $\alpha \in \Lambda_{N\beta}$. $\forall \alpha \in \Lambda_{N\beta}$, $\pi_\alpha(y_0) = \pi_\alpha^{(\beta)}(\pi_\beta^{(\Gamma)}(x_0)) \in B_\alpha$. Let $\Lambda_N = \bigcup_{\beta \in \Gamma_N} \Lambda_{N\beta}$, which is a finite set and the union is pairwise disjoint. Let $B_\alpha = X_\alpha$, $\forall \alpha \in \Lambda_\beta$, $\forall \beta \in \Gamma \setminus \Gamma_N$. Define $B := \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} B_\alpha$. Clearly, B is a basis open set in $\mathcal{O}$ and $y_0 \in B$. $\forall y \in B$. Let $x = E_{\text{inv}}(y)$, then, $y = E(x)$. $\forall \beta \in \Gamma$, $\forall \alpha \in \Lambda_\beta$, $\pi_\alpha(y) = \pi_\alpha(E(x)) = \pi_\alpha^{(\beta)}(\pi_\beta^{(\Gamma)}(x)) \in B_\alpha$. Then, $\forall \beta \in \Gamma_N$, $\pi_\beta^{(\Gamma)}(x) \in \prod_{\alpha \in \Lambda_\beta} B_\alpha = B^{(\beta)} \subseteq \bar{B}^{(\beta)}$. Then, $x \in B^{(\Gamma)}$.

Hence, $E_{\text{inv}}(B) \subseteq B^{(\Gamma)}$. Hence, E_{inv} is continuous at y_0 . By the arbitrariness of y_0 and Proposition 3.9, we have E_{inv} is continuous.

This shows that E is a homeomorphism of $(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, \mathcal{O}^{(\Gamma)})$ to $(\prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha, \mathcal{O})$, and completes the proof of the proposition. $\square$

Proposition 3.31 *Let Λ be an index set, $\mathcal{X}_\alpha := (X_\alpha, \mathcal{O}_{X_\alpha})$ and $\mathcal{Y}_\alpha := (Y_\alpha, \mathcal{O}_{Y_\alpha})$ be topological spaces, $\alpha \in \Lambda$. Assume that $\mathcal{X}_\alpha$ and $\mathcal{Y}_\alpha$ are homeomorphic, $\forall \alpha \in \Lambda$. Then, the product topological spaces $\mathcal{X} = (\prod_{\alpha \in \Lambda} X_\alpha, \mathcal{O}_X) := \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha$ and $\mathcal{Y} = (\prod_{\alpha \in \Lambda} Y_\alpha, \mathcal{O}_Y) := \prod_{\alpha \in \Lambda} \mathcal{Y}_\alpha$ are homeomorphic.*

Proof $\forall \alpha \in \Lambda$, since $\mathcal{X}_\alpha$ and $\mathcal{Y}_\alpha$ are homeomorphic, then, $\exists H_\alpha : \mathcal{X}_\alpha \rightarrow \mathcal{Y}_\alpha$ such that H_α is bijective and both H_α and $H_{\alpha \text{inv}}$ are continuous. Define mapping $H : \mathcal{X} \rightarrow \mathcal{Y}$ by, $\forall x \in \prod_{\alpha \in \Lambda} X_\alpha$, $\pi_\alpha^{(Y)}(H(x)) = H_\alpha(\pi_\alpha^{(X)}(x))$, $\forall \alpha \in \Lambda$.

$\forall y \in \mathcal{Y}$, define $x \in \mathcal{X}$ by $\pi_\alpha^{(X)}(x) = H_{\alpha \text{inv}}(\pi_\alpha^{(Y)}(y))$, $\forall \alpha \in \Lambda$. Then, $H(x) = y$. Hence, H is surjective.

$\forall x_1, x_2 \in \mathcal{X}$ with $x_1 \neq x_2$. Then, $\exists \alpha_0 \in \Lambda$ such that $\pi_{\alpha_0}^{(X)}(x_1) \neq \pi_{\alpha_0}^{(X)}(x_2)$. Then, $\pi_{\alpha_0}^{(Y)}(H(x_1)) = H_{\alpha_0}(\pi_{\alpha_0}^{(X)}(x_1)) \neq H_{\alpha_0}(\pi_{\alpha_0}^{(X)}(x_2)) = \pi_{\alpha_0}^{(Y)}(H(x_2))$, since H_{α_0} is injective. Hence, $H(x_1) \neq H(x_2)$. Therefore, H is injective. Therefore, H is invertible with inverse H_{inv} .

Next, we show that H is continuous. For any basis open set $O_Y \in \mathcal{O}_Y$. Then, $O_Y = \prod_{\alpha \in \Lambda} O_{Y_\alpha}$ with $O_{Y_\alpha} \in \mathcal{O}_{Y_\alpha}$, $\forall \alpha \in \Lambda$, and $O_{Y_\alpha} = Y_\alpha$ for all α 's except finitely many α 's, say $\alpha \in \Lambda_N$. Then, $H_{\text{inv}}(O_Y) = \prod_{\alpha \in \Lambda} H_{\alpha \text{inv}}(O_{Y_\alpha}) \in \mathcal{O}_X$. Hence, H is continuous.

Finally, we show that H_{inv} is continuous. For any basis open set $O_X \in \mathcal{O}_X$. Then, $O_X = \prod_{\alpha \in \Lambda} O_{X_\alpha}$ with $O_{X_\alpha} \in \mathcal{O}_{X_\alpha}$, $\forall \alpha \in \Lambda$, and $O_{X_\alpha} = X_\alpha$ for all α 's except finitely many α 's, say $\alpha \in \Lambda_N$. Then, $H(O_X) = \prod_{\alpha \in \Lambda} H_\alpha(O_{X_\alpha}) \in \mathcal{O}_Y$. Hence, H_{inv} is continuous.

Hence, H is a homeomorphism. This completes the proof of the proposition. $\square$

Proposition 3.32 *Let Λ be an index set, $\mathcal{X}$ be a topological space, $\mathcal{Y}_\alpha := (Y_\alpha, \mathcal{O}_\alpha)$ be a topological space, and $f_\alpha : \mathcal{X} \rightarrow \mathcal{Y}_\alpha$, $\forall \alpha \in \Lambda$. Let $\mathcal{Y} = (Y, \mathcal{O}) := \prod_{\alpha \in \Lambda} \mathcal{Y}_\alpha$ be the product topological space and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be given by, $\forall x \in \mathcal{X}$, $\pi_\alpha(f(x)) = f_\alpha(x)$, $\forall \alpha \in \Lambda$. Then, f is continuous at $x_0 \in \mathcal{X}$ if, and only if, f_α is continuous at x_0 , $\forall \alpha \in \Lambda$.*

Proof “Sufficiency” $\forall O \in \mathcal{O}$ with $f(x_0) \in O$. By Proposition 3.25 and Definition 3.17, $\exists$ a basis open set $B = \prod_{\alpha \in \Lambda} B_\alpha \in \mathcal{O}$ such that $f(x_0) \in B \subseteq O$, where $B_\alpha \in \mathcal{O}_\alpha$, $\forall \alpha \in \Lambda$, and $B_\alpha = Y_\alpha$ for all α 's except finitely many α 's, say $\alpha \in \Lambda_N$. $\forall \alpha \in \Lambda_N$, $f_\alpha(x_0) = \pi_\alpha(f(x_0)) \in \pi_\alpha(B) = B_\alpha$. By the continuity of f_α at x_0 , $\exists U_\alpha \in \mathcal{O}_\mathcal{X}$ with $x_0 \in U_\alpha$ such that $f_\alpha(U_\alpha) \subseteq B_\alpha$. Let $U := \bigcap_{\alpha \in \Lambda_N} U_\alpha \in \mathcal{O}_\mathcal{X}$. Clearly $x_0 \in U$. $\forall x \in U$,

$\forall \alpha \in \Lambda_N$, $\pi_\alpha(f(x)) = f_\alpha(x) \in f_\alpha(U) \subseteq f_\alpha(U_\alpha) \subseteq B_\alpha$. Hence, $f(x) \in B$ and $f(U) \subseteq B \subseteq O$. Then, f is continuous at x_0 .

“Necessity” $\forall \alpha \in \Lambda$, the projection function π_α is continuous, by Proposition 3.27. Note that $f_\alpha = \pi_\alpha \circ f$. Then, f_α is continuous at x_0 by Proposition 3.12.

This completes the proof of the proposition. $\square$

3.5 The Separation Axioms

Definition 3.33 Let $(X, \mathcal{O})$ be a topological space. It is said to be

T₁ (Tychonoff): $\forall x, y \in X$ with $x \neq y$, $\exists O \in \mathcal{O}$ such that $x \in O$ and $y \notin O$.

T₂ (Hausdorff): $\forall x, y \in X$ with $x \neq y$, $\exists O_1, O_2 \in \mathcal{O}$ such that $x \in O_1$, $y \in O_2$, and $O_1 \cap O_2 = \emptyset$.

T₃ (regular): it is Tychonoff and, $\forall x \in X$, $\forall F \subseteq X$ with F being closed and $x \notin F$, $\exists O_1, O_2 \in \mathcal{O}$ such that $x \in O_1$, $F \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$.

T₄ (normal): it is Tychonoff and, $\forall F_1, F_2 \subseteq X$ with F_1 and F_2 being closed and $F_1 \cap F_2 = \emptyset$, $\exists O_1, O_2 \in \mathcal{O}$ such that $F_1 \subseteq O_1$, $F_2 \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$.

Note that $(X, \mathcal{O})$ is Tychonoff implies that, $\forall x \in X$, the singleton set $\{x\}$ is closed since $\widetilde{\{x\}} = \bigcup_{y \in X, y \neq x} O_y$, where $O_y \in \mathcal{O}$ and $y \in O_y$ and $x \notin O_y$. Then, it is clear that $\mathbf{T}_4 \Rightarrow \mathbf{T}_3 \Rightarrow \mathbf{T}_2 \Rightarrow \mathbf{T}_1$.

Proposition 3.34 A topological space $(X, \mathcal{O})$ is Tychonoff if, and only if, $\forall x \in X$, the singleton set $\{x\}$ is closed.

Proof “Only if” $\forall x \in X$, $\forall y \in X$ with $y \neq x$, we have $\exists O_y \in \mathcal{O}$ such that $y \in O_y$ and $x \notin O_y$. Then, $\widetilde{\{x\}} = \bigcup_{y \in \widetilde{\{x\}}} O_y \in \mathcal{O}$. Hence, $\{x\}$ is closed.

“If” $\forall x, y \in X$ with $x \neq y$, we have $\widetilde{\{y\}} \in \mathcal{O}$. Then, $x \notin \widetilde{\{y\}}$ and $y \in \widetilde{\{y\}}$. Hence, $(X, \mathcal{O})$ is Tychonoff.

This completes the proof of the proposition. $\square$

Proposition 3.35 Let $(X, \mathcal{O})$ be a Tychonoff topological space. It is normal if, and only if, for all closed subset $F \subseteq X$ and any open subset $O \in \mathcal{O}$ with $F \subseteq O$, $\exists U \in \mathcal{O}$ such that $F \subseteq U \subseteq \overline{U} \subseteq O$.

Proof “Necessity” Since $(X, \mathcal{O})$ is normal, $F, \widetilde{O}$ are closed, and $F \cap \widetilde{O} = \emptyset$, then $\exists O_1, O_2 \in \mathcal{O}$ such that $F \subseteq O_1, \widetilde{O} \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$. Then, we have $F \subseteq O_1 \subseteq \widetilde{O_2} \subseteq O$. Since $\widetilde{O_2}$ is closed, then, $\overline{O_1} \subseteq \widetilde{O_2}$. Therefore, we have $F \subseteq O_1 \subseteq \overline{O_1} \subseteq \widetilde{O_2} \subseteq O$. So, $U = O_1$.

“Sufficiency” For any closed subsets $F_1, F_2 \subseteq X$ with $F_1 \cap F_2 = \emptyset$, we have $F_1 \subseteq \widetilde{F_2}$. Then, $\exists U \in \mathcal{O}$ such that $F_1 \subseteq U \subseteq \overline{U} \subseteq \widetilde{F_2}$. Let $O_1 = U \in \mathcal{O}$ and $O_2 = \widetilde{\overline{U}} \in \mathcal{O}$. Clearly, $O_1 \cap O_2 = \emptyset$, $F_1 \subseteq O_1$, and $F_2 \subseteq O_2$. Hence, $(X, \mathcal{O})$ is normal.

This completes the proof of the proposition. $\square$

It is easy to show that the product topological space of Hausdorff topological spaces is Hausdorff.

3.6 Category Theory

Definition 3.36 In a topological space $(X, \mathcal{O})$, a subset $D \subseteq X$ is said to be dense if $\overline{D} = X$. $(X, \mathcal{O})$ is said to be separable if there exists a countable dense subset $D \subseteq X$. A subset $M \subseteq X$ is said to be nowhere dense if $\widetilde{M^\circ} = \widetilde{M}$ is dense in X . A subset $F \subseteq X$ is said to be of first category (or meager) if F is the countable union of nowhere dense subsets of X . A subset $S \subseteq X$ is said to be of second category (or nonmeager) if S is not meager. A subset $H \subseteq X$ is said to be a residual set (or comeager) if $\widetilde{H}$ is meager. $(X, \mathcal{O})$ is said to be second category everywhere if every nonempty open subset of X is of second category.

Proposition 3.37 Let $\mathcal{X}$ be a topological space and $Y \subseteq \mathcal{X}$ be dense. Then, $\forall O \in \mathcal{O}$, $\overline{O \cap Y} = \overline{O}$.

Proof Clearly, $\overline{O} \supseteq O \cap Y$ is closed. Then, we have $\overline{O \cap Y} \subseteq \overline{O}$. $\forall x \in \overline{O}$, $\forall U \in \mathcal{O}$ with $x \in U$, by Proposition 3.3, $U \cap O \neq \emptyset$. Then, $(U \cap O) \cap Y \neq \emptyset$ since $Y = \mathcal{X}$ and $U \cap O \in \mathcal{O}$. Hence, we have $U \cap (O \cap Y) \neq \emptyset$. This implies that $x \in \overline{O \cap Y}$, by Proposition 3.3. Hence, we have $\overline{O} \subseteq \overline{O \cap Y}$. Therefore, $\overline{O} = \overline{O \cap Y}$. This completes the proof of the proposition. $\square$

Proposition 3.38 Let $\mathcal{X}$ be a topological space. Then, $\mathcal{X}$ is second category everywhere if, and only if, countable intersection of open dense subsets is dense.

Proof “Only if” Let $(O_n)_{n=1}^\infty$ be a sequence of open dense subsets of $\mathcal{X}$. Let $F_n := \widetilde{O_n}$, $\forall n \in \mathbb{N}$. Clearly, F_n is closed and nowhere dense, $\forall n \in \mathbb{N}$, since $\widetilde{\widetilde{F_n}} = \widetilde{F_n} = \overline{O_n} = \mathcal{X}$. Note that $\bigcap_{n=1}^\infty O_n = (\bigcup_{n=1}^\infty F_n)^\sim$. Now, Suppose that $\bigcap_{n=1}^\infty O_n$ is not dense. Then, $\exists O \in \mathcal{O}_\mathcal{X}$ which is nonempty such that $(\bigcap_{n=1}^\infty O_n) \cap O = \emptyset$. Then, $((\bigcup_{n=1}^\infty F_n)^\sim) \cap O = \emptyset$. Hence, $O = \bigcup_{n=1}^\infty (F_n \cap O)$ and is of first category. This contradicts with the fact that O is of second category. Therefore, $\bigcap_{n=1}^\infty O_n$ must be dense.

The case of finite intersection can be converted to the above scenario by padding $\mathcal{X}$ as additional open dense subsets.

“If” $\forall O \in \mathcal{O}$ with $O \neq \emptyset$. For any countable collection $(E_\alpha)_{\alpha \in \Lambda}$ of nowhere dense subsets of $\mathcal{X}$. $\forall \alpha \in \Lambda$, $\widetilde{E_\alpha}$ is open and dense. Then, $\bigcap_{\alpha \in \Lambda} \widetilde{E_\alpha}$

is dense. Then, $O \cap \left(\bigcap_{\alpha \in \Lambda} \widetilde{E_\alpha} \right) \neq \emptyset$, which further implies that $O \not\subseteq \bigcup_{\alpha \in \Lambda} \overline{E_\alpha}$. Hence, $O \not\subseteq \bigcup_{\alpha \in \Lambda} E_\alpha$. Hence, O is of second category by the arbitrariness of $(E_\alpha)_{\alpha \in \Lambda}$. Hence, $\mathcal{X}$ is second category everywhere by the arbitrariness of O .

This completes the proof of the proposition. $\square$

Proposition 3.39 *Let $(X, \mathcal{O})$ be a topological space. $F, O \subseteq X$ with $O, \widetilde{F} \in \mathcal{O}$. Then,*

- (i) $\overline{O} \setminus O$ and $F \setminus F^\circ$ are nowhere dense;
- (ii) if $(X, \mathcal{O})$ is second category everywhere, and F is of first category, then F is nowhere dense.

Proof (i) $\overline{O} \setminus O = \overline{O} \cap \widetilde{O}$ is closed. Note that

$$\overline{\overline{O} \setminus O} = \overline{\overline{O} \cap \widetilde{O}} = \overline{\overline{O} \cup O} = \overline{\overline{O} \cup \overline{O}} = X$$

where we have applied Proposition 3.3 in the above. Hence $\overline{O} \setminus O$ is nowhere dense.

$F \setminus F^\circ = F \cap \widetilde{F}^\circ$ is closed. Note that

$$\overline{\overline{F \setminus F^\circ}} = \overline{\overline{F \cap \widetilde{F}^\circ}} = \overline{\widetilde{F} \cup F^\circ} = \widetilde{F} \cup \overline{F^\circ} = \widetilde{F}^\circ \cup \overline{F^\circ} = X$$

where we have applied Proposition 3.3 in the above. Hence $F \setminus F^\circ$ is nowhere dense.

(ii) Let $F = \bigcup_{n=1}^{\infty} F_n$, where F_n 's are nowhere dense subsets of X . Then, $\widetilde{F_n}$ is open and dense in $(X, \mathcal{O})$. By Proposition 3.38, we have $\bigcap_{n=1}^{\infty} \widetilde{F_n}$ is dense.

Since F is closed, then $\overline{F} = F$. $\forall m \in \mathbb{N}$. Note that $F \supseteq \bigcup_{n=1}^m F_n$. Then, $\overline{F} \supseteq \overline{\bigcup_{n=1}^m F_n} = \bigcup_{n=1}^m \overline{F_n}$, by Proposition 3.3. Therefore,

$$F = \overline{F} \supseteq \bigcup_{n=1}^{\infty} \overline{F_n} \supseteq \bigcup_{n=1}^{\infty} F_n = F$$

This implies $\overline{F} = \bigcup_{n=1}^{\infty} \overline{F_n}$. Hence, we have $\widetilde{F} = \bigcap_{n=1}^{\infty} \widetilde{F_n}$, which is dense. Hence, F is nowhere dense.

This completes the proof of the proposition. $\square$

Proposition 3.40 *Let $(X, \mathcal{O})$ be a topological space. Then,*

- (i) A closed set $F \subseteq X$ is nowhere dense if, and only if, it contains no nonempty open subset;
- (ii) A subset $E \subseteq X$ is nowhere dense if, and only if, $\forall O \in \mathcal{O}$ with $O \neq \emptyset$, $\exists U \in \mathcal{O}$ with $U \neq \emptyset$ such that $U \subseteq O \setminus E$.

Proof (i) “Only if” Suppose $\exists U \in \mathcal{O}$ with $U \neq \emptyset$ and $U \subseteq F$. Then, $\tilde{F} \cap U = \emptyset$. This contradicts with the fact that $\overline{\tilde{F}} = \overline{F} = X$, by Proposition 3.3. Hence, the result holds.

“If” $\forall x \in X, \forall U \in \mathcal{O}$ with $x \in U$. Then, $U \not\subseteq F$ implies $U \cap \tilde{F} \neq \emptyset$. This implies that x is a point of closure of $\tilde{F}$, by Proposition 3.3. Hence, we have $\overline{\tilde{F}} = \overline{F} = X$. Hence, F is nowhere dense.

(ii) “Only if” $\forall O \in \mathcal{O}$ with $O \neq \emptyset$. Note that

$$O \setminus E = O \cap \tilde{E} \supseteq O \cap \tilde{\tilde{E}}$$

Since E is nowhere dense, then $\tilde{\tilde{E}}$ is open and dense, and hence $O \cap \tilde{\tilde{E}}$ is open and nonempty by the nonemptiness of O and Proposition 3.3. Then, $\exists U = O \cap \tilde{\tilde{E}}$.

“If” $\forall x \in X, \forall O \in \mathcal{O}$ with $x \in O$. Then, $\exists U \in \mathcal{O}$ with $U \neq \emptyset$ such that $U \subseteq O \setminus E$. Note that $U = U^\circ \subseteq \left(O \cap \tilde{\tilde{E}}\right)^\circ = O^\circ \cap \tilde{E}^\circ = O \cap \tilde{\tilde{E}}$, where we have made use of Proposition 3.3. Hence, $O \cap \tilde{\tilde{E}}$ is nonempty. By the arbitrariness of O and Proposition 3.3, x is a point of closure for $\tilde{\tilde{E}}$. By the arbitrariness of x , E is nowhere dense.

This completes the proof of the proposition. $\square$

Theorem 3.41 (Uniform Boundedness Principle) *Let $(X, \mathcal{O})$ be a topological space that is second category everywhere. Let $\mathcal{F}$ be a family of continuous real-valued functions of X . $\forall x \in X, \exists M_x \in [0, \infty) \subset \mathbb{R}$ such that $|f(x)| \leq M_x, \forall f \in \mathcal{F}$. Then, $\exists U \in \mathcal{O}$ with $U \neq \emptyset$ and $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $|f(x)| \leq M, \forall x \in U$ and $\forall f \in \mathcal{F}$.*

Proof Let $E_{m,f} := f_{\text{inv}}([-m, m]) \subseteq X, \forall f \in \mathcal{F}$ and $\forall m \in \mathbb{Z}_+$. Since f is continuous and $[-m, m]$ is a closed interval, then, by Proposition 3.10, $E_{m,f}$ ’s are closed. $\forall m \in \mathbb{Z}_+$, let $E_m := \bigcap_{f \in \mathcal{F}} E_{m,f}$, which is closed. $\forall x \in X, \exists m \in \mathbb{Z}_+$ with $m \geq M_x$ such that $|f(x)| \leq m, \forall f \in \mathcal{F}$. Then, $x \in E_{m,f}, \forall f \in \mathcal{F}$, and hence $x \in E_m$. Therefore, $X = \bigcup_{m=0}^{\infty} E_m$. Since X is second category everywhere, then E_m ’s are not all nowhere dense. Then, $\exists n \in \mathbb{Z}_+$ such that E_n is not nowhere dense, that is $\overline{\tilde{E}_n} = \overline{E_n} \neq X$. Then, by Proposition 3.40, E_n contains a nonempty open subset $U \in \mathcal{O}$. Then, $U \subseteq E_n$ and $U \neq \emptyset$. $\forall x \in U, \forall f \in \mathcal{F}$, we have $|f(x)| \leq n$. This completes the proof of the theorem. $\square$

3.7 Connectedness

Definition 3.42 *A topological space $\mathcal{X}$ is said to be connected if there do not exist nonempty open sets O_1, O_2 such that $\mathcal{X} = O_1 \cup O_2$ and $O_1 \cap O_2 = \emptyset$. Such a pair of O_1 and O_2 is called a separation of $\mathcal{X}$ if it exists.*

Proposition 3.43 *Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be a surjective continuous function. If $\mathcal{X}$ is connected then $\mathcal{Y}$ is connected.*

Proof Suppose $\mathcal{Y}$ is not connected. Then, $\exists O_1, O_2 \in \mathcal{O}_{\mathcal{Y}}$ with $O_1 \neq \emptyset$ and $O_2 \neq \emptyset$ such that $O_1 \cup O_2 = \mathcal{Y}$ and $O_1 \cap O_2 = \emptyset$. Since f is surjective, then $f_{\text{inv}}(O_1) \neq \emptyset$ and $f_{\text{inv}}(O_2) \neq \emptyset$. Since f is continuous, then $f_{\text{inv}}(O_1), f_{\text{inv}}(O_2) \in \mathcal{O}_{\mathcal{X}}$. By Proposition 2.5, $f_{\text{inv}}(O_1) \cup f_{\text{inv}}(O_2) = f_{\text{inv}}(O_1 \cup O_2) = \mathcal{X}$ and $f_{\text{inv}}(O_1) \cap f_{\text{inv}}(O_2) = f_{\text{inv}}(O_1 \cap O_2) = \emptyset$. This contradicts with the assumption that $\mathcal{X}$ is connected. Therefore, $\mathcal{Y}$ is connected. This completes the proof of the proposition. $\square$

Theorem 3.44 (Mean Value Theorem) *Let $\mathcal{X}$ be a topological space and $f : \mathcal{X} \rightarrow \mathbb{R}$ be a continuous function. Assume that $\mathcal{X}$ is connected and $\exists x, y \in \mathcal{X}$ such that $f(x) < c < f(y)$ for some $c \in \mathbb{R}$. Then, $\exists z \in \mathcal{X}$ such that $f(z) = c$.*

Proof Suppose the result is false. Then, $f_{\text{inv}}(\{c\}) = \emptyset$. Let $O_1 := f_{\text{inv}}((c, \infty))$ and $O_2 := f_{\text{inv}}((-\infty, c))$. Then, $x \in O_2 \in \mathcal{O}_{\mathcal{X}}$ and $y \in O_1 \in \mathcal{O}_{\mathcal{X}}$, by assumptions of the theorem. By Proposition 2.5, $O_1 \cap O_2 = \emptyset$ and $O_1 \cup O_2 = \mathcal{X}$. This contradicts with the assumption that $\mathcal{X}$ is connected. Hence, the result is true. This completes the proof of the theorem. $\square$

Proposition 3.45 *Let $\mathcal{X}$ be a topological space, $U \subseteq V \subseteq \overline{U} \subseteq \mathcal{X}$, and U be connected in the subset topology $\mathcal{O}_U$. Then, V is connected in the subset topology $\mathcal{O}_V$.*

Proof Suppose V is not connected in its subset topology. Then, $\exists O_{V1}, O_{V2} \in \mathcal{O}_V$ with $O_{V1} \neq \emptyset$ and $O_{V2} \neq \emptyset$, such that $O_{V1} \cup O_{V2} = V$ and $O_{V1} \cap O_{V2} = \emptyset$. By Proposition 3.4, $\exists O_1, O_2 \in \mathcal{O}$ such that $O_{V1} = O_1 \cap V$ and $O_{V2} = O_2 \cap V$. Let $x_1 \in O_{V1}$ and $x_2 \in O_{V2}$. Then, $x_i \in \overline{U} \cap O_i$, $i = 1, 2$. By Proposition 3.3, $\exists \bar{x}_i \in U \cap O_i =: O_{Ui} \neq \emptyset$, $i = 1, 2$. By Proposition 3.4, $O_{U1}, O_{U2} \in \mathcal{O}_U$. Note that $O_{U1} \cap O_{U2} = U \cap O_1 \cap O_2 = U \cap V \cap O_1 \cap O_2 = U \cap (O_1 \cap V) \cap (O_2 \cap V) = U \cap O_{V1} \cap O_{V2} = \emptyset$ and $O_{U1} \cup O_{U2} = U \cap (O_1 \cup O_2) = U \cap V \cap (O_1 \cup O_2) = U \cap ((V \cap O_1) \cup (V \cap O_2)) = U \cap (O_{V1} \cup O_{V2}) = U \cap V = U$. Hence, the pair (O_{U1}, O_{U2}) is a separation of U . This implies that U is not connected in its subset topology. This contradicts with the assumption. Hence, V is connected. $\square$

Definition 3.46 *Let $\mathcal{X}$ be a topological space, $x_0 \in \mathcal{X}$. Let*

$$\mathcal{M} := \{ M \subseteq \mathcal{X} \mid x_0 \in M, M \text{ is connected in the subset topology.} \}$$

The component containing x_0 is defined by $A := \bigcup_{M \in \mathcal{M}} M$.

Clearly, $\mathcal{X}$ is the union of its components.

Proposition 3.47 *Let $\mathcal{X}$ be a topological space and $x_0 \in \mathcal{X}$. Then, the component A of $\mathcal{X}$ containing x_0 is connected and closed.*

Proof Let $\mathcal{M}$ be as defined in Definition 3.46. Suppose A is not connected. Let $\mathcal{O}_A$ be the subset topology on A . Then, $\exists O_1, O_2 \in \mathcal{O}_A$ with $O_1 \neq \emptyset$ and $O_2 \neq \emptyset$ such that $O_1 \cup O_2 = A$ and $O_1 \cap O_2 = \emptyset$. Without loss of generality, assume that $x_0 \in O_1$. By the definition of A , $\exists A_0 \in \mathcal{M}$ such that $A_0 \cap O_2 \neq \emptyset$. Note that $A_0 = (O_1 \cap A_0) \cup (O_2 \cap A_0)$ and $O_1 \cap A_0 \ni x_0$ and $O_2 \cap A_0$ are nonempty and open in the subset topology on A_0 . Furthermore, $(O_1 \cap A_0) \cap (O_2 \cap A_0) = \emptyset$. This shows that A_0 is not connected, which contradicts with the fact that $A_0 \in \mathcal{M}$. Hence, A is connected.

By Proposition 3.45, $\overline{A}$ is connected. Then, we have $\overline{A} \in \mathcal{M}$ and $\overline{A} = A$. By Proposition 3.3, A is closed in $\mathcal{O}$.

This completes the proof of the proposition. $\square$

Proposition 3.48 *Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $A_\alpha \subseteq X$ be connected (in subset topology), $\forall \alpha \in \Lambda$, where Λ is an index set. Assume that $A_{\alpha_1} \cap A_{\alpha_2} \neq \emptyset$, $\forall \alpha_1, \alpha_2 \in \Lambda$. Then, $A := \bigcup_{\alpha \in \Lambda} A_\alpha$ is connected (in subset topology).*

Proof Suppose A is not connected. Then, $\exists O_1, O_2 \in \mathcal{O}$ such that $O_1 \cap A \neq \emptyset \neq O_2 \cap A$, $(O_1 \cap A) \cap (O_2 \cap A) = \emptyset$, and $(O_1 \cap A) \cup (O_2 \cap A) = (O_1 \cup O_2) \cap A = A$. Let $x_i \in O_i \cap A$, $i = 1, 2$. Then, $\exists \alpha_i \in \Lambda$ such that $x_i \in A_{\alpha_i}$, $i = 1, 2$. By the assumption, let $x_0 \in A_{\alpha_1} \cap A_{\alpha_2} \neq \emptyset$. Without loss of generality, assume $x_0 \in O_1$. Then, $x_0 \in O_1 \cap A_{\alpha_2} \neq \emptyset$, $x_2 \in O_2 \cap A_{\alpha_2} \neq \emptyset$, $(O_1 \cap A_{\alpha_2}) \cap (O_2 \cap A_{\alpha_2}) \subseteq (O_1 \cap A) \cap (O_2 \cap A) = \emptyset$, and $(O_1 \cap A_{\alpha_2}) \cup (O_2 \cap A_{\alpha_2}) = (O_1 \cup O_2) \cap A_{\alpha_2} = A_{\alpha_2}$. Hence, $O_1 \cap A_{\alpha_2}$ and $O_2 \cap A_{\alpha_2}$ form a separation of A_{α_2} . This implies that A_{α_2} is not connected. This is a contradiction. Therefore, A is connected. This completes the proof of the proposition. $\square$

Definition 3.49 *Let $\mathcal{X}$ be a topological space. It is said to be locally connected if there exists a basis $\mathcal{B}$ such that B is connected (in the subset topology), $\forall B \in \mathcal{B}$.*

Proposition 3.50 *Any component of a locally connected topological space is open.*

Proof Let $\mathcal{X}$ be a locally connected topological space and $\mathcal{B}$ be a basis made up of connected sets. $\forall x_0 \in \mathcal{X}$, let A be the component containing x_0 . By Proposition 3.47, A is closed and connected. $\forall x \in A$, $\exists B \in \mathcal{B}$ such that $x \in B$. Since B is connected, then, by Definition 3.46, $B \subseteq A$. Hence, A is open. This completes the proof of the proposition. $\square$

Proposition 3.51 *Let $(X_\alpha, \mathcal{O}_\alpha)$ be a connected topological space, $\forall \alpha \in \Gamma$, where Γ is an index set. Let $(X, \mathcal{O})$ be the product topological space $\prod_{\alpha \in \Gamma} (X_\alpha, \mathcal{O}_\alpha)$. Then, $(X, \mathcal{O})$ is connected.*

Proof Suppose that X is not connected. Then, $\exists O_1, O_2 \in \mathcal{O}$ with $O_1 \neq \emptyset$ and $O_2 \neq \emptyset$ such that $O_1 \cup O_2 = X$ and $O_1 \cap O_2 = \emptyset$. Let $\mathcal{B}$ be

the basis defined in Proposition 3.25. Then, there exists nonempty disjoint index sets Λ_1 and Λ_2 such that $O_1 = \bigcup_{\lambda \in \Lambda_1} B_\lambda$ and $O_2 = \bigcup_{\lambda \in \Lambda_2} C_\lambda$, where $B_\lambda \in \mathcal{B}$, $\forall \lambda \in \Lambda_1$, and $C_\lambda \in \mathcal{B}$, $\forall \lambda \in \Lambda_2$. $\forall \lambda \in \Lambda_1$, $B_\lambda = \prod_{\alpha \in \Gamma} B_{\lambda\alpha}$, where $B_{\lambda\alpha} \in \mathcal{O}_\alpha$, $\forall \alpha \in \Gamma$, and $B_{\lambda\alpha} = X_\alpha$ for all except finitely many α 's, say $\alpha \in \Gamma_\lambda$. Note that $\Gamma_\lambda \neq \emptyset$, since $B_\lambda \subseteq O_1 \subset X$. $\forall \lambda \in \Lambda_2$, $C_\lambda = \prod_{\alpha \in \Gamma} C_{\lambda\alpha}$, where $C_{\lambda\alpha} \in \mathcal{O}_\alpha$, $\forall \alpha \in \Gamma$, and $C_{\lambda\alpha} = X_\alpha$ for all except finitely many α 's, say $\alpha \in \Gamma_\lambda$. Note that $\Gamma_\lambda \neq \emptyset$, since $C_\lambda \subseteq O_2 \subset X$. Note that

$$\emptyset = O_1 \cap O_2 = \bigcup_{\lambda \in \Lambda_1} \bigcup_{\gamma \in \Lambda_2} (B_\lambda \cap C_\gamma)$$

Therefore, $B_\lambda \cap C_\gamma = \emptyset$, $\forall \lambda \in \Lambda_1$, $\forall \gamma \in \Lambda_2$, which implies that $\exists \alpha_{\lambda\gamma} \in \Gamma \ni B_{\lambda\alpha_{\lambda\gamma}} \cap C_{\gamma\alpha_{\lambda\gamma}} = \emptyset$.

Fix $x_1 \in O_1$ and $x_2 \in O_2$. Then, $x_1 \neq x_2$. $\exists \lambda_1 \in \Lambda_1$ such that $x_1 \in B_{\lambda_1}$. Let $\Gamma_1 := \Gamma_{\lambda_1}$. $\exists \lambda_2 \in \Lambda_2$ such that $x_2 \in C_{\lambda_2}$. Let $\Gamma_2 := \Gamma_{\lambda_2}$. Note that $\forall x \in X$ with $\pi_\alpha(x) = \pi_\alpha(x_1)$, $\forall \alpha \in \Gamma_1$, we have $x \in B_{\lambda_1} \subseteq O_1$. Similarly, $\forall x \in X$ with $\pi_\alpha(x) = \pi_\alpha(x_2)$, $\forall \alpha \in \Gamma_2$, we have $x \in C_{\lambda_2} \subseteq O_2$. Therefore, starting with $x_1 \in O_1$ and switch, one by one, its coordinate $\pi_\alpha(x_1)$ to $\pi_\alpha(x_2)$, for all $\alpha \in \Gamma_2$, we will end up with a point $x_3 \in O_2$. Therefore, there must exist a step in this process such that switching one coordinate $\pi_{\alpha_0}(x_1)$ to $\pi_{\alpha_0}(x_2)$, for some $\alpha_0 \in \Gamma_2$, leads to the change of set membership from $\bar{x}_1 \in O_1$ before the switch to $\bar{x}_2 \in O_2$ after the switch. In summary, there exist $\bar{x}_1 \in O_1$, $\bar{x}_2 \in O_2$, and $\alpha_0 \in \Gamma$ such that $\pi_\alpha(\bar{x}_1) = \pi_\alpha(\bar{x}_2)$, $\forall \alpha \in \Gamma \setminus \{\alpha_0\}$. Since $O_1 \cap O_2 = \emptyset$, we must have $\pi_{\alpha_0}(\bar{x}_1) \neq \pi_{\alpha_0}(\bar{x}_2)$.

Define

$$\Lambda_{10} := \{ \lambda \in \Lambda_1 \mid B_\lambda = \prod_{\alpha \in \Gamma} B_{\lambda\alpha}, \pi_\alpha(\bar{x}_1) \in B_{\lambda\alpha}, \forall \alpha \in \Gamma \setminus \{\alpha_0\} \}$$

and

$$\Lambda_{20} := \{ \lambda \in \Lambda_2 \mid C_\lambda = \prod_{\alpha \in \Gamma} C_{\lambda\alpha}, \pi_\alpha(\bar{x}_2) \in C_{\lambda\alpha}, \forall \alpha \in \Gamma \setminus \{\alpha_0\} \}$$

Let $M := \prod_{\alpha \in \Gamma} M_\alpha \subseteq X$ where $M_\alpha = \{\pi_\alpha(\bar{x}_1)\}$, $\forall \alpha \in \Gamma \setminus \{\alpha_0\}$, and $M_{\alpha_0} = X_{\alpha_0}$. Note that $M \subseteq X = O_1 \cup O_2 = (\bigcup_{\lambda \in \Lambda_1} B_\lambda) \cup (\bigcup_{\gamma \in \Lambda_2} C_\gamma)$ implies that

$$M \subseteq \left(\bigcup_{\lambda \in \Lambda_{10}} B_\lambda \right) \cup \left(\bigcup_{\gamma \in \Lambda_{20}} C_\gamma \right) = \bigcup_{\lambda \in \Lambda_{10}} \bigcup_{\gamma \in \Lambda_{20}} (B_\lambda \cup C_\gamma)$$

Therefore, we have $X_{\alpha_0} = \pi_{\alpha_0}(M) \subseteq \bigcup_{\lambda \in \Lambda_{10}} \bigcup_{\gamma \in \Lambda_{20}} (\pi_{\alpha_0}(B_\lambda) \cup \pi_{\alpha_0}(C_\gamma)) = \bigcup_{\lambda \in \Lambda_{10}} \bigcup_{\gamma \in \Lambda_{20}} (B_{\lambda\alpha_0} \cup C_{\gamma\alpha_0}) \subseteq X_{\alpha_0}$, by Proposition 2.5. Then,

$$\begin{aligned} X_{\alpha_0} &= \bigcup_{\lambda \in \Lambda_{10}} \bigcup_{\gamma \in \Lambda_{20}} (B_{\lambda\alpha_0} \cup C_{\gamma\alpha_0}) = \left(\bigcup_{\lambda \in \Lambda_{10}} B_{\lambda\alpha_0} \right) \cup \left(\bigcup_{\gamma \in \Lambda_{20}} C_{\gamma\alpha_0} \right) \\ &=: D_1 \cup D_2 \end{aligned}$$

By derivations in the first paragraph of this proof, we have $\forall \lambda \in \Lambda_{10}$, $\forall \gamma \in \Lambda_{20}$, $B_\lambda \cap C_\gamma = \prod_{\alpha \in \Gamma} (B_{\lambda\alpha} \cap C_{\gamma\alpha}) = \emptyset$. Note that, $\forall \alpha \in \Gamma \setminus \{\alpha_0\}$, $\pi_\alpha(\bar{x}_1) = \pi_\alpha(\bar{x}_2) \in B_{\lambda\alpha} \cap C_{\gamma\alpha} \neq \emptyset$. Then, we must have $B_{\lambda\alpha_0} \cap C_{\gamma\alpha_0} = \emptyset$. Therefore, we have

$$D_1 \cap D_2 = \bigcup_{\lambda \in \Lambda_{10}} \bigcup_{\gamma \in \Lambda_{20}} (B_{\lambda\alpha_0} \cap C_{\gamma\alpha_0}) = \emptyset$$

Clearly, $D_1, D_2 \in \mathcal{O}_{\alpha_0}$, $\pi_{\alpha_0}(\bar{x}_1) \in D_1 \neq \emptyset$, and $\pi_{\alpha_0}(\bar{x}_2) \in D_2 \neq \emptyset$ since $\bar{x}_1 \in O_1$ and $\bar{x}_2 \in O_2$. This shows that D_1 and D_2 form a separation of X_{α_0} . This contradicts with the assumption that $(X_\alpha, \mathcal{O}_\alpha)$ is connected, $\forall \alpha \in \Gamma$. Therefore, $(X, \mathcal{O})$ is connected.

This completes the proof of the proposition. $\square$

Definition 3.52 Let $\mathcal{X}$ be a topological space and $I := [0, 1] \subset \mathbb{R}$ be endowed with the subset topology of $\mathbb{R}$. A curve in $\mathcal{X}$ is a continuous mapping $\gamma : I \rightarrow \mathcal{X}$. $\gamma(0)$ is called the beginning point and $\gamma(1)$ is called the end point. A closed curve is such that $\gamma(0) = \gamma(1)$. Two closed curves γ_1 and γ_2 are said to be homotopic to each other if there exists a continuous function $\phi : I \times I \rightarrow \mathcal{X}$ such that $\phi(t, 0) = \gamma_1(t)$, $\phi(t, 1) = \gamma_2(t)$, and $\phi(0, t) = \phi(1, t)$, $\forall t \in I$.

Definition 3.53 Let $\mathcal{X}$ be a topological space. It is said to be arcwise connected if $\forall x_1, x_2 \in \mathcal{X}$, there exists a curve γ in $\mathcal{X}$ such that $\gamma(0) = x_1$ and $\gamma(1) = x_2$. $\mathcal{X}$ is said to be simply connected if it is arcwise connected and any closed curve is homotopic to a single point (that is a degenerate curve γ with $\gamma(t) = x \in \mathcal{X}$, $\forall t \in [0, 1] \subset \mathbb{R}$.)

Proposition 3.54 Let $(X, \mathcal{O})$ be a topological space. Then, it is connected if it is arcwise connected.

Proof Suppose $(X, \mathcal{O})$ is not connected. Then, $\exists O_1, O_2 \in \mathcal{O}$ with $O_1 \neq \emptyset$ and $O_2 \neq \emptyset$ such that $O_1 \cup O_2 = X$ and $O_1 \cap O_2 = \emptyset$. Fix $x_1 \in O_1$ and $x_2 \in O_2$. Then, $x_1 \neq x_2$. Since $(X, \mathcal{O})$ is arcwise connected, then there exists a curve γ such that $\gamma(0) = x_1$ and $\gamma(1) = x_2$. Define $t := \sup \{s \in [0, 1] \subset \mathbb{R} \mid \gamma([0, s]) \subseteq O_1\}$. Then, $t \in [0, 1] \subset \mathbb{R}$.

Claim 3.54.1 $\gamma(t) \in O_1$.

Proof of claim: Suppose $\gamma(t) \notin O_1$, then $\gamma(t) \in O_2$. Then, $t > 0$, since $\gamma(0) = x_1 \in O_1$. Since γ is continuous, then $\exists t_1 \in [0, t) \subset \mathbb{R}$ such that $\gamma((t_1, t]) \subseteq O_2$. $\forall s \in (t_1, 1] \subset \mathbb{R}$, $\gamma([0, s]) \cap O_2 \neq \emptyset$. Then, $s \notin \{s \in [0, 1] \subset \mathbb{R} \mid \gamma([0, s]) \subseteq O_1\}$. Hence, $s \geq t$, which implies that $t_1 \geq t$. This contradicts $t_1 < t$. Hence, $\gamma(t) \in O_1$. This completes the proof of the claim. $\square$

Since $\gamma(1) = x_2 \in O_2$, then $t < 1$. By the continuity of γ , $\exists t_2 \in (t, 1] \subset \mathbb{R}$ such that $\gamma([t, t_2]) \subseteq O_1$. $\forall s \in [0, t)$, $\exists s_1 \in (s, t]$ such that

$s_1 \in \{s \in [0, 1] \subset \mathbb{R} \mid \gamma([0, s]) \subseteq O_1\}$. Then, $\gamma([0, s_1]) \subseteq O_1$ and hence $\gamma(s) \in O_1$. Therefore, $\gamma([0, t]) \subseteq O_1$. This coupled with $\gamma([t, t_2]) \subseteq O_1$, we have $\gamma([0, t_2]) \subseteq O_1$. Then, $t \geq t_2$. This contradicts with $t < t_2$.

Therefore, $(X, \mathcal{O})$ is connected. This completes the proof of the proposition. $\square$

3.8 Continuous Real-Valued Functions

Theorem 3.55 (Urysohn's Lemma) *Let $(X, \mathcal{O})$ be a normal topological space, $A, B \subseteq X$ be closed subsets, and $A \cap B = \emptyset$. Then, there exists a continuous real-valued function $f : X \rightarrow [0, 1] \subset \mathbb{R}$ such that $f(x) = 0$, $\forall x \in A$, and $f(x) = 1$, $\forall x \in B$.*

Proof Since the set $Q := \mathbb{Q} \cap [0, 1]$ is countable, then, by recursively applying Proposition 3.35, we may find $(O_r)_{r \in Q} \subseteq \mathcal{O}$ such that the following two properties are satisfied:

1. $\forall r \in Q, A \subseteq O_r \subseteq \overline{O_r} \subseteq \tilde{B}$;
2. $\forall r, s \in Q$ with $r < s$, $\overline{O_r} \subseteq O_s$.

Define the real-valued function $f : X \rightarrow \mathbb{R}$ by

$$f(x) = \inf(\{r \in Q \mid x \in O_r\} \cup \{1\})$$

Clearly, $f : X \rightarrow [0, 1]$, $f(x) = 0$, $\forall x \in O_0$, and $f(x) = 1$, $\forall x \in \widetilde{O_1}$. By 1, we have $A \subseteq O_0$ and $O_1 \subseteq \tilde{B}$. Hence, all we need to show is that f is continuous. $\forall x_0 \in X$, we will show that f is continuous at x_0 . Let $a_0 = f(x_0) \in [0, 1]$. $\forall U \subseteq \mathbb{R}$ with U being open and $a_0 \in U$, $\exists a_1, a_2, a_3, a_4 \in \mathbb{Q}$ such that $a_1 < a_2 < a_0 < a_3 < a_4$ and $(a_1, a_4) \subseteq U$. Let $\bar{a}_2 = \max\{a_2, 0\}$ and $\bar{a}_3 = \min\{a_3, 1\}$. Then, we must have $a_1 < \bar{a}_2 \leq a_0 \leq \bar{a}_3 < a_4$ and $\bar{a}_2, \bar{a}_3 \in Q$. We will distinguish three exhaustive and mutually exclusive cases: Case 1: $a_0 \in (0, 1)$; Case 2: $a_0 = 0$; Case 3: $a_0 = 1$.

Case 1: $a_0 \in (0, 1)$. Then, we must have $a_1 < \bar{a}_2 < a_0 < \bar{a}_3 < a_4$. Let $V = \widetilde{\overline{O_{\bar{a}_2}}} \cap O_{\bar{a}_3} \in \mathcal{O}$. $\forall x \in V$, we have $x \in O_{\bar{a}_3}$ and $f(x) \leq \bar{a}_3$. Also, $x \in \widetilde{\overline{O_{\bar{a}_2}}}$ implies that $f(x) \geq \bar{a}_2$. Hence, $f(V) \subseteq [\bar{a}_2, \bar{a}_3] \subset (a_1, a_4) \subseteq U$. $f(x_0) = a_0 < \bar{a}_3$ implies that $x_0 \in O_{\bar{a}_3}$. $f(x_0) = a_0 > \bar{a}_2$ implies that $\exists \tilde{a}_2 \in (\bar{a}_2, a_0) \cap Q$ such that $x_0 \in \widetilde{\overline{O_{\tilde{a}_2}}} \subseteq \widetilde{\overline{O_{\bar{a}_2}}}$. Therefore $x_0 \in V$. This shows that $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $f(V) \subseteq U$.

Case 2: $a_0 = 0$. Then, we must have $a_1 < 0 = a_0 < \bar{a}_3 < a_4$. Take $V = O_{\bar{a}_3} \in \mathcal{O}$. We must have $x_0 \in V$. $\forall x \in V$, $0 \leq f(x) \leq \bar{a}_3$. Hence, $f(V) \subseteq [0, \bar{a}_3] \subset (a_1, a_4) \subseteq U$. Hence, $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $f(V) \subseteq U$.

Case 3: $a_0 = 1$. Then, we must have $a_1 < \bar{a}_2 < a_0 = 1 < a_4$. Take $V = \widetilde{\overline{O_{\bar{a}_2}}} \in \mathcal{O}$. Since $f(x_0) = a_0 = 1$, then $x_0 \in \widetilde{\overline{O_{\frac{1+\bar{a}_2}{2}}}} \subseteq \widetilde{\overline{O_{\bar{a}_2}}} = V$. $\forall x \in V$,

$f(x) \geq \bar{a}_2$. Hence, $f(V) \subseteq [\bar{a}_2, 1] \subset (a_1, a_4) \subseteq U$. Hence, $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $f(V) \subseteq U$.

Therefore, in all cases, $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $f(V) \subseteq U$. Hence, f is continuous at x_0 . By the arbitrariness of x_0 and Proposition 3.9, f is continuous. This completes the proof of the theorem. $\square$

Proposition 3.56 *Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces and $\mathcal{Y}$ be Hausdorff. $f_1 : \mathcal{X} \rightarrow \mathcal{Y}$ and $f_2 : \mathcal{X} \rightarrow \mathcal{Y}$ are continuous. Let $D \subseteq \mathcal{X}$ be dense. Assume that $f_1|_D = f_2|_D$. Then, $f_1 = f_2$.*

Proof Suppose $f_1 \neq f_2$. Then, $\exists x \in \mathcal{X}$ such that $f_1(x) \neq f_2(x)$. Since $\mathcal{Y}$ is Hausdorff, then $\exists O_1, O_2 \in \mathcal{O}_{\mathcal{Y}}$ such that $f_1(x) \in O_1$, $f_2(x) \in O_2$, and $O_1 \cap O_2 = \emptyset$. Since f_1 and f_2 are continuous, we have $U_1 := f_{1\text{inv}}(O_1) \in \mathcal{O}_{\mathcal{X}}$ and $U_2 := f_{2\text{inv}}(O_2) \in \mathcal{O}_{\mathcal{X}}$. Note that $x \in U_1 \cap U_2 \in \mathcal{O}_{\mathcal{X}}$ and $x \in \overline{D}$, then, by Proposition 3.3, $\exists \bar{x} \in D \cap U_1 \cap U_2$. Then, $f_1(\bar{x}) \in O_1$ and $f_2(\bar{x}) \in O_2$, which implies that $f_1|_D(\bar{x}) \neq f_2|_D(\bar{x})$. This is a contradiction. Hence, we must have $f_1 = f_2$.

This completes the proof of the proposition. $\square$

Theorem 3.57 (Tietze's Extension Theorem) *Let $(X, \mathcal{O})$ be a normal topological space, $A \subseteq X$ be closed, and $h : A \rightarrow \mathbb{R}$. Let A be endowed with the subset topology $\mathcal{O}_A$. Assume that h is continuous. Then, there exists a continuous function $k : X \rightarrow \mathbb{R}$ such that $k|_A = h$.*

Proof Let $f := \frac{h}{1+|h|}$. Then, $|f(x)| < 1$, $\forall x \in A$, and by Proposition 3.12, f is continuous.

Claim 3.57.1 *Let $l : A \rightarrow \mathbb{R}$ be a continuous function such that $|l(x)| \leq c_1 \in \mathbb{R}$, $\forall x \in A$, where $c_1 > 0$. Then, there exists a continuous function $g : X \rightarrow \mathbb{R}$ such that $|g(x)| \leq c_1/3$, $\forall x \in X$, and $|l(x) - g(x)| \leq 2c_1/3$, $\forall x \in A$.*

Proof of claim: Let $B := \{x \in A \mid l(x) \leq -c_1/3\}$ and $C := \{x \in A \mid l(x) \geq c_1/3\}$. Then, B and C are closed sets in $\mathcal{O}_A$, by the continuity of l and Proposition 3.10. Since A is closed, then B and C are closed in $\mathcal{O}$, by Proposition 3.5. Clearly, $B \cap C = \emptyset$. By Urysohn's Lemma, there exists a continuous function $g : X \rightarrow \mathbb{R}$ such that $|g(x)| \leq c_1/3$, $\forall x \in X$, $g(x) = -c_1/3$, $\forall x \in B$, and $g(x) = c_1/3$, $\forall x \in C$. Hence, $|l(x) - g(x)| \leq 2c_1/3$, $\forall x \in A$. This completes the proof of the claim. $\square$

By repeated application of Claim 3.57.1, we may define $f_i : X \rightarrow \mathbb{R}$, $\forall i \in \mathbb{N}$, such that f_i is continuous, $|f_i(x)| \leq \frac{2^{i-1}}{3^i}$, $\forall x \in X$, and $|f(x) - \sum_{k=1}^i f_k(x)| \leq \frac{2^i}{3^i}$, $\forall x \in A$.

Define $g : X \rightarrow \mathbb{R}$ by $g(x) = \lim_{i \in \mathbb{N}} \sum_{k=1}^i f_k(x)$, $\forall x \in X$. Clearly, g is well-defined, $g|_A = f$, and $|g(x)| \leq \sum_{i=1}^{\infty} \frac{2^{i-1}}{3^i} = 1$, $\forall x \in X$. $\forall x_0 \in X$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$. $\exists N \in \mathbb{N}$ such that $\left| \sum_{i=N+1}^{\infty} f_i(x) \right| < \epsilon/3$, $\forall x \in X$. By the continuity of $f_1, \dots, f_N$ and Proposition 3.9, $\exists U \in \mathcal{O}$ with $x_0 \in U$ such that $\left| \sum_{i=1}^N f_i(x) - \sum_{i=1}^N f_i(x_0) \right| < \epsilon/3$, $\forall x \in U$. Then, we have, $\forall x \in U$,

$$\begin{aligned} |g(x) - g(x_0)| &\leq \left| g(x) - \sum_{i=1}^N f_i(x) \right| + \left| \sum_{i=1}^N f_i(x) - \sum_{i=1}^N f_i(x_0) \right| \\ &\quad + \left| \sum_{i=1}^N f_i(x_0) - g(x_0) \right| < \epsilon \end{aligned}$$

Therefore, g is continuous at x_0 . Then, g is continuous, by the arbitrariness of x_0 and Proposition 3.9.

Let $D := \{x \in X \mid |g(x)| = 1\}$. Clearly, D is a closed set, by Proposition 3.10. Note that $A \cap D = \emptyset$, since $g|_A = f$ and $|f(x)| < 1$, $\forall x \in A$. Then, by Urysohn's Lemma, there exists a continuous function $\bar{g} : X \rightarrow [0, 1]$ such that $\bar{g}|_A = 1$ and $\bar{g}|_D = 0$. Define $k : X \rightarrow \mathbb{R}$ by $k(x) = \frac{\bar{g}(x)g(x)}{1 - \bar{g}(x)|g(x)|}$, $\forall x \in X$. By Propositions 3.12 and 3.32 and the fact that $1 - \bar{g}(x)|g(x)| \neq 0$, $\forall x \in X$, we have k is continuous. $\forall x \in A$, $k(x) = \frac{g(x)}{1 - |g(x)|} = h(x)$. Hence, $k|_A = h$.

This completes the proof of the theorem. $\square$

Definition 3.58 Let X be a set and $\mathcal{F}$ be a collection of real-valued functions of X . Then, there is the weakest topology on X such that all functions in $\mathcal{F}$ are continuous. This topology is called the weak topology generated by $\mathcal{F}$.

Let X be a set, $I := [0, 1] \subset \mathbb{R}$, and $\mathcal{F}$ be a collection of functions of X to I such that $\forall x, y \in X$ with $x \neq y$, $\exists f \in \mathcal{F}$, we have $f(x) \neq f(y)$. Each $f \in \mathcal{F}$ is a point in I^X and $\mathcal{F}$ can be identified with a subset of I^X . The topology that $\mathcal{F}$ inherits as a subspace of I^X is called the *topology of pointwise convergence*. Now, X can be identified with a subset of $I^{\mathcal{F}}$ by, $\forall x \in X$, $\pi_f(x) = f(x)$, $\forall f \in \mathcal{F}$. Then, the topology of X as a subset of $I^{\mathcal{F}}$ is the weak topology generated by $\mathcal{F}$.

Proposition 3.59 Let $\mathcal{X}$ be a topological space, $I := [0, 1] \subset \mathbb{R}$, and $\mathcal{F}$ be a collection of continuous functions of $\mathcal{X}$ to I such that $\forall x, y \in \mathcal{X}$ with $x \neq y$, $\exists f \in \mathcal{F}$, we have $f(x) \neq f(y)$. Let $E : \mathcal{X} \rightarrow I^{\mathcal{F}}$ be the equivalence map given by, $\forall x \in \mathcal{X}$, $\pi_f(E(x)) = f(x)$, $\forall f \in \mathcal{F}$. Then, E is continuous. Furthermore, if $\forall$ closed set $F \subseteq \mathcal{X}$ and $\forall x \in \mathcal{X}$ with $x \notin F$, $\exists f \in \mathcal{F}$ with $f(x) = 1$ and $f|_F = 0$, then $E : \mathcal{X} \rightarrow E(\mathcal{X})$ is a homeomorphism.

Proof $\forall x_0 \in \mathcal{X}$. Fix a basis open set O in $I^{\mathcal{F}}$ with $E(x_0) \in O$. By Proposition 3.25, $O = \prod_{f \in \mathcal{F}} O_f$, where $O_f \in \mathcal{O}_I$ with $\mathcal{O}_I$ being the subset topology on I , $\forall f \in \mathcal{F}$, and $O_f = I$ for all f 's except finitely many f 's, say

$f \in \mathcal{F}_N$. Let $U = \bigcap_{f \in \mathcal{F}_N} f_{\text{inv}}(O_f) \in \mathcal{O}_X$. By $E(x_0) \in O$, we have $x_0 \in U$. $\forall x \in U$, we have $\pi_f(E(x)) \in O_f = I$, $\forall f \in \mathcal{F} \setminus \mathcal{F}_N$, and $\pi_f(E(x)) = f(x) \in O_f$, $\forall f \in \mathcal{F}_N$. Hence, $E(x) \in O$. Then, $E(U) \subseteq O$. Therefore, E is continuous at x_0 . By the arbitrariness of x_0 and Proposition 3.9, E is continuous.

Under the additional assumption on $\mathcal{F}$, we need to show that E is a homeomorphism between X and $E(X)$. $\forall x, y \in X$ with $x \neq y$, $\exists f \in \mathcal{F}$ such that $\pi_f(E(x)) = f(x) \neq f(y) = \pi_f(E(y))$. Then, $E(x) \neq E(y)$. Hence, $E : X \rightarrow E(X)$ is injective. Clearly, $E : X \rightarrow E(X)$ is surjective. Then, $E : X \rightarrow E(X)$ is bijective and admits inverse $E_{\text{inv}} : E(X) \rightarrow X$. $\forall x_0 \in X$, we will show that E_{inv} is continuous at $E(x_0)$. $\forall O \in \mathcal{O}_X$ with $x_0 \in O$. $\tilde{O}$ is closed and $x_0 \notin \tilde{O}$. Then, $\exists f_0 \in \mathcal{F}$ such that $f_0(x_0) = 1$ and $f_0|_{\tilde{O}} = 0$. Define $U = \prod_{f \in \mathcal{F}} U_f \subseteq I^{\mathcal{F}}$ by $U_f = I$, $\forall f \in \mathcal{F} \setminus \{f_0\}$ and $U_{f_0} = (1/2, 1] \in \mathcal{O}_I$. Clearly, U is open in $I^{\mathcal{F}}$. Clearly, $E(x_0) \in U$. $\forall x \in X$ with $E(x) \in U$, we have $\pi_{f_0}(E(x)) = f_0(x) > 1/2$. Then, $x \notin \tilde{O}$ and $x \in O$. This shows that $E_{\text{inv}}(E(X) \cap U) \subseteq O$. Hence, E_{inv} is continuous at $E(x_0)$. By the arbitrariness of x_0 and Proposition 3.9, $E_{\text{inv}} : E(X) \rightarrow X$ is continuous. This implies that $E : X \rightarrow E(X)$ is a homeomorphism.

This completes the proof of the proposition. $\square$

Definition 3.60 A topological space X is said to be completely regular (or $\mathbf{T}_{3\frac{1}{2}}$) if it is Tychonoff and $\forall x_0 \in X$ and $\forall$ closed set $F \subseteq X$ with $x_0 \notin F$, there exists a continuous real-valued function $f : X \rightarrow [0, 1]$ such that $f(x_0) = 1$ and $f|_F = 0$.

Proposition 3.61 A normal topological space is completely regular. A completely regular topological space is regular.

Proof Let X be a normal topological space. Then, X is Tychonoff. $\forall x_0 \in X$ and $\forall$ closed set $F \subseteq X$ with $x_0 \notin F$, we have $\{x_0\}$ is closed, by Proposition 3.34. By Urysohn's Lemma, there exists a continuous real-valued function $f : X \rightarrow [0, 1]$ such that $f(x_0) = 1$ and $f|_F = 0$. Hence, X is completely regular.

Let X be a completely regular topological space. Then, X is Tychonoff. $\forall x_0 \in X$ and $\forall$ closed set $F \subseteq X$ with $x_0 \notin F$, there exists a continuous real-valued function $f : X \rightarrow [0, 1]$ such that $f(x_0) = 1$ and $f|_F = 0$. Let $O_1 := \{x \in X \mid f(x) > 1/2\}$ and $O_2 := \{x \in X \mid f(x) < 1/2\}$. Then, $O_1, O_2 \in \mathcal{O}$ by the continuity of f . Clearly, $x_0 \in O_1$, $F \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$. Hence, X is regular.

This completes the proof of the proposition. $\square$

Corollary 3.62 Let X be a completely regular topological space, $I = [0, 1] \subset \mathbb{R}$, and $\mathcal{F} := \{f : X \rightarrow I \mid f \text{ is continuous}\}$. Then, the equivalence map: $E : X \rightarrow I^{\mathcal{F}}$ defined by $\pi_f(E(x)) = f(x)$, $\forall x \in X$, $\forall f \in \mathcal{F}$, is a homeomorphism between X and $E(X) \subseteq I^{\mathcal{F}}$.

Proof Since $\mathcal{X}$ is completely regular, then $\mathcal{X}$ is Tychonoff and all singleton subset of $\mathcal{X}$ is closed. Then, it is easy to check that all assumptions in Proposition 3.59 are satisfied. Then, the result follows. This completes the proof of the corollary. $\square$

3.9 Nets and Convergence

Definition 3.63 A directed system is a nonempty set A and a relation on A , $\prec$, such that

- (i) $\prec$ is transitive;
- (ii) $\forall \alpha, \beta \in A, \exists \gamma \in A$ such that $\alpha \prec \gamma$ and $\beta \prec \gamma$.

A net is a mapping of a directed system $\mathcal{A} := (A, \prec)$ to a topological space $\mathcal{X}$. $\forall \alpha \in A$, the image is x_α . The net is denoted by $(x_\alpha)_{\alpha \in \mathcal{A}}$, where we have abuse the notation to say $\alpha \in \mathcal{A}$ when $\alpha \in A$. It is understood that the relation for $\mathcal{A}$ is $\prec_{\mathcal{A}}$, where we will ignore the subscript $\mathcal{A}$ if no confusion arises.

A point $x \in \mathcal{X}$ is a limit of the net $(x_\alpha)_{\alpha \in \mathcal{A}}$ if $\forall O \in \mathcal{O}$ with $x \in O$, $\exists \alpha_0 \in \mathcal{A} \ni \forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$, we have $x_\alpha \in O$. We also say that $(x_\alpha)_{\alpha \in \mathcal{A}}$ converges to x .

A point $x \in \mathcal{X}$ is a cluster point of $(x_\alpha)_{\alpha \in \mathcal{A}}$ if $\forall O \in \mathcal{O}$ with $x \in O$, $\forall \alpha \in \mathcal{A}, \exists \beta \in \mathcal{A}$ with $\alpha \prec \beta \ni x_\beta \in O$.

Clearly, a limit point of a net is a cluster point of the net.

In Definition 3.63, we may restrict O to be a basis open set without changing the meaning of the definition.

Example 3.64 $(\mathbb{N}, \leq)$ is a directed system. A net over $(\mathbb{N}, \leq)$ corresponds to a sequence. $\diamond$

Proposition 3.65 Let $\mathcal{X}$ be a topological space. Then, the following statements holds.

- (i) $\mathcal{X}$ is Hausdorff if, and only if, for all net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathcal{X}$, there exists at most one limit point for the net. We then write $x = \lim_{\alpha \in \mathcal{A}} x_\alpha$ when the limit exists.
- (ii) If $\mathcal{X}$ is Hausdorff, any convergent net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathcal{X}$ with limit $x \in \mathcal{X}$ has exactly one cluster point, which is x .

Proof (i) “Only if” Suppose there exists a net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathcal{X}$ such that $\exists x_A, x_B \in \mathcal{X}$ with $x_A \neq x_B$ and x_A and x_B are limit points of the net. Since $\mathcal{X}$ is Hausdorff, then $\exists O_1, O_2 \in \mathcal{O}$ such that $x_A \in O_1, x_B \in O_2$, and $O_1 \cap O_2 = \emptyset$. Since x_A is the limit of the net, then $\exists \alpha_1 \in \mathcal{A}, \forall \alpha \in \mathcal{A}$ with $\alpha_1 \prec \alpha$, we have $x_\alpha \in O_1$. Similarly, since x_B is the limit of the net, then

$\exists \alpha_2 \in \mathcal{A}$, $\forall \alpha \in \mathcal{A}$ with $\alpha_2 \prec \alpha$, we have $x_\alpha \in O_2$. Since $\mathcal{A}$ is a directed system, $\exists \alpha_3 \in \mathcal{A}$ such that $\alpha_1 \prec \alpha_3$ and $\alpha_2 \prec \alpha_3$. Then, we have $x_{\alpha_3} \in O_1$ and $x_{\alpha_3} \in O_2$, which implies that $O_1 \cap O_2 \neq \emptyset$, which is a contradiction. Therefore, every net in $\mathcal{X}$ has at most one limit point.

“If” Suppose $\mathcal{X}$ is not Hausdorff. Then, $\exists x_A, x_B \in \mathcal{X}$ with $x_A \neq x_B$ such that $\forall O_A, O_B \in \mathcal{O}$ with $x_A \in O_A$ and $x_B \in O_B$, we have $O_A \cap O_B \neq \emptyset$. Let $\Lambda := \{(O_A, O_B) \mid x_A \in O_A \in \mathcal{O}, x_B \in O_B \in \mathcal{O}\}$. Clearly, $(\mathcal{X}, \mathcal{X}) \in \Lambda$, then $\Lambda \neq \emptyset$. Define a relation $\prec$ on Λ by, $\forall (O_{A1}, O_{B1}), (O_{A2}, O_{B2}) \in \Lambda$, we say $(O_{A1}, O_{B1}) \prec (O_{A2}, O_{B2})$ if $O_{A1} \supseteq O_{A2}$ and $O_{B1} \supseteq O_{B2}$. Clearly, $\prec$ is transitive on Λ . $\forall (O_{A1}, O_{B1}), (O_{A2}, O_{B2}) \in \Lambda$, we have $x_A \in O_{A3} := O_{A1} \cap O_{A2} \in \mathcal{O}$ and $x_B \in O_{B3} := O_{B1} \cap O_{B2} \in \mathcal{O}$. Then, we have $(O_{A3}, O_{B3}) \in \Lambda$, $(O_{A1}, O_{B1}) \prec (O_{A3}, O_{B3})$, and $(O_{A2}, O_{B2}) \prec (O_{A3}, O_{B3})$. Hence, $\mathcal{A} := (\Lambda, \prec)$ is a directed system. $\forall (O_A, O_B) \in \Lambda$, $O_A \cap O_B \neq \emptyset$. By Axiom of Choice, we may have a mapping $x_{(O_A, O_B)} \in O_A \cap O_B$, $\forall (O_A, O_B) \in \Lambda$. Then, the net $(x_{(O_A, O_B)})_{(O_A, O_B) \in \Lambda} \subseteq \mathcal{X}$. $\forall O_{A1} \in \mathcal{O}$ with $x_A \in O_{A1}$. Fix $O_{B1} := \mathcal{X} \in \mathcal{O}$ with $x_B \in O_{B1}$. Then, $(O_{A1}, O_{B1}) \in \Lambda$. $\forall (O_{A2}, O_{B2}) \in \Lambda$ with $(O_{A1}, O_{B1}) \prec (O_{A2}, O_{B2})$, we have $x_{(O_{A2}, O_{B2})} \in O_{A2} \cap O_{B2} \subseteq O_{A1} \cap O_{B1} = O_{A1}$. Hence, x_A is a limit point of $(x_{(O_A, O_B)})_{(O_A, O_B) \in \Lambda}$. $\forall O_{B1} \in \mathcal{O}$ with $x_B \in O_{B1}$. Fix $O_{A1} := \mathcal{X} \in \mathcal{O}$ with $x_A \in O_{A1}$. Then, $(O_{A1}, O_{B1}) \in \Lambda$. $\forall (O_{A2}, O_{B2}) \in \Lambda$ with $(O_{A1}, O_{B1}) \prec (O_{A2}, O_{B2})$, we have $x_{(O_{A2}, O_{B2})} \in O_{A2} \cap O_{B2} \subseteq O_{A1} \cap O_{B1} = O_{B1}$. Hence, x_B is a limit point of $(x_{(O_A, O_B)})_{(O_A, O_B) \in \Lambda}$. This contradicts with the assumption that every net has at most one limit point. Therefore, $\mathcal{X}$ is Hausdorff.

(ii) Let $\mathcal{X}$ be Hausdorff and net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathcal{X}$ satisfy $\lim_{\alpha \in \mathcal{A}} x_\alpha = x \in \mathcal{X}$. Clearly, x is a cluster point of the net $(x_\alpha)_{\alpha \in \mathcal{A}}$. Let $y \in \mathcal{X}$ with $y \neq x$ be another cluster point of the net $(x_\alpha)_{\alpha \in \mathcal{A}}$. Then, there exist $O_1, O_2 \in \mathcal{O}$ with $x \in O_1$ and $y \in O_2$ and $O_1 \cap O_2 = \emptyset$. Since x is the limit of the net, then $\exists \alpha_1 \in \mathcal{A}$ such that $\forall \alpha \in \mathcal{A}$ with $\alpha_1 \prec \alpha$, we have $x_\alpha \in O_1$. Then, $x_\alpha \notin O_2$. This contradicts the definition that y is cluster point of $(x_\alpha)_{\alpha \in \mathcal{A}}$. Hence, the net has exactly one cluster point, which is x .

This completes the proof of the proposition. $\square$

Proposition 3.66 *Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces, $D \subseteq \mathcal{X}$ with subset topology $\mathcal{O}_D$, and $f : D \rightarrow \mathcal{Y}$. Then, the following are equivalent.*

- (i) f is continuous at $x_0 \in D$;
- (ii) $\forall$ net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D$ with x_0 as a limit point, we have that the net $(f(x_\alpha))_{\alpha \in \mathcal{A}}$ has a limit point $f(x_0)$.
- (iii) $\forall$ net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D$ with x_0 as a cluster point, we have that the net $(f(x_\alpha))_{\alpha \in \mathcal{A}} \subseteq \mathcal{Y}$ has a cluster point $f(x_0)$.

Proof (i) $\Rightarrow$ (ii). Fix a net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D$ with $x_0 \in D$ as a limit point. $\forall O_Y \in \mathcal{O}_Y$ with $f(x_0) \in O_Y$. By the continuity of f at x_0 , $\exists O_X \in \mathcal{O}_X$ with $x_0 \in O_X$ such that $f(O_X) \subseteq O_Y$. Since x_0 is a limit point of $(x_\alpha)_{\alpha \in \mathcal{A}}$,

then $\exists \alpha_0 \in \mathcal{A}$ such that, $\forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$, we have $x_\alpha \in O_X$. Then, $f(x_\alpha) \in O_Y$. Hence, we have $f(x_0)$ is a limit point of $(f(x_\alpha))_{\alpha \in \mathcal{A}}$.

(ii) $\Rightarrow$ (i). Suppose f is not continuous at $x_0 \in D$. Then, $\exists O_{Y_0} \in \mathcal{O}_Y$ with $f(x_0) \in O_{Y_0}$ such that, $\forall O_X \in \mathcal{O}_X$ with $x_0 \in O_X$, we have $f(O_X) \not\subseteq O_{Y_0}$. Let $\mathcal{M} := \{O \in \mathcal{O}_X \mid x_0 \in O\}$. Clearly, $\mathcal{X} \in \mathcal{M}$ and $\mathcal{M} \neq \emptyset$. Define a relation $\prec$ on $\mathcal{M}$ by, $\forall O_1, O_2 \in \mathcal{M}$, we say $O_1 \prec O_2$ if $O_1 \supseteq O_2$. Clearly, $\prec$ is transitive on $\mathcal{M}$. $\forall O_1, O_2 \in \mathcal{M}$, let $O_3 = O_1 \cap O_2 \in \mathcal{O}_X$ and $x_0 \in O_3$. Then, $O_3 \in \mathcal{M}$, $O_1 \prec O_3$, and $O_2 \prec O_3$. Hence, $\mathcal{A} := (\mathcal{M}, \prec)$ is a directed system. $\forall O \in \mathcal{M}$, $f(O) \setminus O_{Y_0} \neq \emptyset$. By Axiom of Choice, we may define a net $(x_O)_{O \in \mathcal{A}}$ by $x_O \in O \cap D$ with $f(x_O) \notin O_{Y_0}$. Clearly, x_0 is a limit point of $(x_O)_{O \in \mathcal{A}}$. Yet, $f(x_0) \in O_{Y_0} \in \mathcal{O}_Y$ and $f(x_O) \notin O_{Y_0}$, $\forall O \in \mathcal{M}$. Then, $f(x_0)$ is not a limit point of the net $(f(x_O))_{O \in \mathcal{A}}$. This contradicts with the assumption. Therefore, f is continuous at x_0 .

(i) $\Rightarrow$ (iii). Fix a net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D$ with x_0 as a cluster point. $\forall O_Y \in \mathcal{O}_Y$ with $f(x_0) \in O_Y$, by the continuity of f at x_0 , $\exists U \in \mathcal{O}_X$ with $x_0 \in U$ such that $f(U) \subseteq O_Y$. By Definition 3.63, $\forall \alpha \in \mathcal{A}$, $\exists \alpha_0 \in \mathcal{A}$ with $\alpha \prec \alpha_0$, $x_{\alpha_0} \in U$. Then, $f(x_{\alpha_0}) \in O_Y$. Hence, $f(x_0)$ is a cluster point of the net $(f(x_\alpha))_{\alpha \in \mathcal{A}}$.

(iii) $\Rightarrow$ (i). Suppose f is not continuous at x_0 . Let $\mathcal{M} := \{O \in \mathcal{O}_X \mid x_0 \in O\}$. Clearly, $\mathcal{A} := (\mathcal{M}, \supseteq)$ is a directed system. $\exists O_{Y_0} \in \mathcal{O}_Y$ with $f(x_0) \in O_{Y_0}$ such that $\forall U \in \mathcal{M}$, we have $f(U) \not\subseteq O_{Y_0}$. By Axiom of Choice, we may assign to each $U \in \mathcal{M}$ an $x_U \in U \cap D$ such that $f(x_U) \in \widetilde{O_{Y_0}}$. Consider the net $(x_U)_{U \in \mathcal{A}} \subseteq D$. Clearly, x_0 is a limit point of the net, and therefore is a cluster point of the net. Consider the net $(f(x_U))_{U \in \mathcal{A}}$. For the open set $O_{Y_0} \ni f(x_0)$, $\forall U \in \mathcal{A}$, $f(x_U) \in \widetilde{O_{Y_0}}$. Then, $f(x_0)$ is not a cluster point of $(f(x_U))_{U \in \mathcal{A}}$. This contradicts with the assumption. Therefore, f must be continuous at x_0 .

This completes the proof of the proposition. $\square$

Proposition 3.67 *Let $(X_\alpha, \mathcal{O}_\alpha)$ be a topological space, $\forall \alpha \in \Lambda$, where Λ is an index set. Let $(X, \mathcal{O})$ be the product space $\prod_{\alpha \in \Lambda} (X_\alpha, \mathcal{O}_\alpha)$. Let $(x_\beta)_{\beta \in \mathcal{A}} \subseteq X$ be a net. Then, $x_0 \in X$ is a limit point of $(x_\beta)_{\beta \in \mathcal{A}}$ if, and only if, $\forall \alpha \in \Lambda$, $\pi_\alpha(x_0) \in X_\alpha$ is a limit point of $(\pi_\alpha(x_\beta))_{\beta \in \mathcal{A}}$.*

Proof “Only if” $\forall \alpha \in \Lambda$, by Proposition 3.27, π_α is continuous. Then, π_α is continuous at $x_0 \in X$, by Proposition 3.9. By Proposition 3.66, $\pi_\alpha(x_0)$ is a limit point of the net $(\pi_\alpha(x_\beta))_{\beta \in \mathcal{A}}$.

“If” Suppose that $x_0 \in \mathcal{X}$ is not a limit point of the net $(x_\beta)_{\beta \in \mathcal{A}}$. Then, $\exists$ a basis open set $B \in \mathcal{O}$ with $x_0 \in B$ such that, $\forall \beta_0 \in \mathcal{A}$, $\exists \beta \in \mathcal{A}$ with $\beta_0 \prec \beta$, we have $x_\beta \notin B$. Then, $B = \prod_{\alpha \in \Lambda} O_\alpha$, $O_\alpha \in \mathcal{O}_\alpha$, $\forall \alpha \in \Lambda$, and $O_\alpha = X_\alpha$ for all α 's except finitely many α 's, say $\alpha \in \Lambda_N$. Then, $\forall \beta_0 \in \mathcal{A}$, $\exists \beta \in \mathcal{A}$ with $\beta_0 \prec \beta$, we have $x_\beta \notin B$. This implies that $\pi_{\alpha_\beta}(x_\beta) \notin O_{\alpha_\beta}$, for some $\alpha_\beta \in \Lambda_N$. Then, by an argument of contradiction, we may show that $\exists \alpha_0 \in \Lambda_N$ such that, $\forall \beta_0 \in \mathcal{A}$, $\exists \beta \in \mathcal{A}$ with $\beta_0 \prec \beta$, we have $\pi_{\alpha_0}(x_\beta) \notin O_{\alpha_0}$. Hence, $\pi_{\alpha_0}(x_0) \in O_{\alpha_0}$ is not the limit of the net

$(\pi_{\alpha_0}(x_\beta))_{\beta \in \mathcal{A}}$. This contradicts with the assumption. Hence, we have x_0 is a limit point of $(x_\beta)_{\beta \in \mathcal{A}}$.

This completes the proof of the proposition. $\square$

Proposition 3.68 *Let $(X, \mathcal{O})$ be a topological space, $E \subseteq X$, and $x \in X$. $x \in \overline{E}$ if, and only if, $\exists$ a net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq E$ such that x is a limit point of the net.*

Proof “Only if” Let $\mathcal{M} := \{O \in \mathcal{O} \mid x \in O\}$. Clearly, $X \in \mathcal{M}$, then $\mathcal{M} \neq \emptyset$. Clearly, $\mathcal{A} := (\mathcal{M}, \supseteq)$ is a directed system. Since $x \in \overline{E}$, then, by Proposition 3.3, $\forall O \in \mathcal{A}$, $O \cap E \neq \emptyset$. By Axiom of Choice, $\exists$ a net $(x_O)_{O \in \mathcal{A}} \subseteq E$ such that $x_O \in O \cap E$, $\forall O \in \mathcal{A}$. $\forall O \in \mathcal{O}$ with $x \in O$, then $O \in \mathcal{A}$. $\forall O_1 \in \mathcal{A}$ with $O \supseteq O_1$, we have $x_{O_1} \in O_1 \cap E \subseteq O$. Hence, x is a limit point of $(x_O)_{O \in \mathcal{A}}$.

“If” Let $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq E$ be the net such that x is a limit point of the net. $\forall O \in \mathcal{O}$ with $x \in O$, $\exists \alpha_0 \in \mathcal{A}$, $\forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$, we have $x_\alpha \in E \cap O$. Since $(x_\alpha)_{\alpha \in \mathcal{A}}$ is a net, then $\exists \alpha_1 \in \mathcal{A}$ with $\alpha_0 \prec \alpha_1$. Then, $x_{\alpha_1} \in O \cap E \neq \emptyset$. By Proposition 3.3, $x \in \overline{E}$.

This completes the proof of the proposition. $\square$

Definition 3.69 *Let $(X, \mathcal{O})$ be a topological space, $\mathcal{A} := (A, \prec)$ be a directed system, and $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq X$ be a net. Let $A_s \subseteq A$ be a subset with the same relation $\prec$ as A such that $\forall \alpha \in \mathcal{A}$, $\exists \alpha_s \in A_s$ such that $\alpha \prec \alpha_s$. Then, $\mathcal{A}_s := (A_s, \prec)$ is a directed system and $(x_\alpha)_{\alpha \in \mathcal{A}_s}$ is a net, which is called a subnet of $(x_\alpha)_{\alpha \in \mathcal{A}}$.*

Proposition 3.70 *Let $(X, \mathcal{O})$ be a topological space and $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq X$ be a net. Then, $x_0 \in X$ is a limit point of $(x_\alpha)_{\alpha \in \mathcal{A}}$ if, and only if, any subnet $(x_\alpha)_{\alpha \in \mathcal{A}_s}$ has a limit point x_0 .*

Proof “Only if” Since $x_0 \in X$ is a limit point of $(x_\alpha)_{\alpha \in \mathcal{A}}$, then $\forall O \in \mathcal{O}$ with $x_0 \in O$, $\exists \alpha_0 \in \mathcal{A}$ such that, $\forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$, we have $x_\alpha \in O$. Let $(x_\alpha)_{\alpha \in \mathcal{A}_s}$ be a subnet. Then, $\exists \alpha_{s_0} \in \mathcal{A}_s$ such that $\alpha_0 \prec \alpha_{s_0}$. $\forall \alpha_s \in \mathcal{A}_s$ with $\alpha_{s_0} \prec \alpha_s$, we have $\alpha_0 \prec \alpha_s$ and $x_{\alpha_s} \in O$. Hence, x_0 is a limit point of the subnet.

“If” Since $(x_\alpha)_{\alpha \in \mathcal{A}}$ is a subnet of itself, then it has limit x_0 .

This completes the proof of the proposition. $\square$

A cluster point of a subnet is clearly a cluster point of the net.

Proposition 3.71 *Let $(X, \mathcal{O})$ be a topological space and $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq X$ be a net. Then, $x_0 \in X$ is a limit point of $(x_\alpha)_{\alpha \in \mathcal{A}}$ if, and only if, for every subnet $(x_\alpha)_{\alpha \in \mathcal{A}_s}$ of $(x_\alpha)_{\alpha \in \mathcal{A}}$, there exists a subnet $(x_\alpha)_{\alpha \in \mathcal{A}_{ss}}$ that has a limit point x_0 .*

Proof “Sufficiency” We assume that every subnet $(x_\alpha)_{\alpha \in \mathcal{A}_s}$ of $(x_\alpha)_{\alpha \in \mathcal{A}}$, there exists a subnet $(x_\alpha)_{\alpha \in \mathcal{A}_{ss}}$ that has a limit point x_0 . We

will prove the result using an argument of contradiction. Suppose x_0 is not a limit point of $(x_\alpha)_{\alpha \in \mathcal{A}}$. Then, $\exists O_0 \in \mathcal{O}$ with $x_0 \in O_0$, $\forall \alpha_0 \in \mathcal{A}$, $\exists \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$ such that $x_\alpha \in \widetilde{O_0}$. Define $\mathcal{A}_s := (\{\alpha \in \mathcal{A} \mid x_\alpha \in \widetilde{O_0}\}, \prec)$. Clearly, $(x_\alpha)_{\alpha \in \mathcal{A}_s}$ is a subnet of $(x_\alpha)_{\alpha \in \mathcal{A}}$. Any subnet $(x_\alpha)_{\alpha \in \mathcal{A}_{ss}}$ of $(x_\alpha)_{\alpha \in \mathcal{A}_s}$, $\forall \alpha_{ss} \in \mathcal{A}_{ss}$, we have $x_{\alpha_{ss}} \in \widetilde{O_0}$. Then, x_0 is not a limit of $(x_\alpha)_{\alpha \in \mathcal{A}_{ss}}$. This contradicts the assumption. Therefore, x_0 is a limit point of $(x_\alpha)_{\alpha \in \mathcal{A}}$.

“Necessity” Let x_0 be a limit point of $(x_\alpha)_{\alpha \in \mathcal{A}}$ and $(x_\alpha)_{\alpha \in \mathcal{A}_s}$ be a subnet. By Proposition 3.70, x_0 is a limit point of $(x_\alpha)_{\alpha \in \mathcal{A}_s}$, which is a subnet of itself. Then, the result holds.

This completes the proof of the proposition. $\square$

Definition 3.72 Let $\mathcal{X} := (X, \mathcal{O}_X)$ and $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be topological spaces, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, and $x_0 \in \mathcal{X}$ be an accumulation point of D . $y_0 \in \mathcal{Y}$ is said to be a limit point of $f(x)$ as $x \rightarrow x_0$ if $\forall O_Y \in \mathcal{O}_Y$ with $y_0 \in O_Y$, $\exists U \in \mathcal{O}_X$ with $x_0 \in U$ such that $f(U \setminus \{x_0\}) = f((D \cap U) \setminus \{x_0\}) \subseteq O_Y$. We will also say that $f(x)$ converges to y_0 as $x \rightarrow x_0$.

When basis are available on topological spaces $\mathcal{X}$ and $\mathcal{Y}$, in Definition 3.72, we may restrict the open sets O_Y and U to be basis open sets without changing the meaning of the definition.

Proposition 3.73 Let $\mathcal{X} := (X, \mathcal{O}_X)$ and $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be topological spaces, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, and $x_0 \in \mathcal{X}$ be an accumulation point of D . If $\mathcal{Y}$ is Hausdorff, then there is at most one limit point of $f(x)$ as $x \rightarrow x_0$. In this case, we will write $\lim_{x \rightarrow x_0} f(x) = y_0 \in \mathcal{Y}$ when the limit exists.

Proof Suppose $f(x)$ admits limit points $y_A, y_B \in \mathcal{Y}$ as $x \rightarrow x_0$ with $y_A \neq y_B$. Since $\mathcal{Y}$ is Hausdorff, then $\exists U_A, U_B \in \mathcal{O}_Y$ such that $y_A \in U_A$, $y_B \in U_B$, and $U_A \cap U_B = \emptyset$. Since y_A is a limit point of $f(x)$ as $x \rightarrow x_0$, then $\exists V_A \in \mathcal{O}_X$ with $x_0 \in V_A$ such that $f(V_A \setminus \{x_0\}) \subseteq U_A$. Since y_B is a limit point of $f(x)$ as $x \rightarrow x_0$, then $\exists V_B \in \mathcal{O}_X$ with $x_0 \in V_B$ such that $f(V_B \setminus \{x_0\}) \subseteq U_B$. Then, $x_0 \in V := V_A \cap V_B \in \mathcal{O}$. Since x_0 is an accumulation point of D , then $\exists x \in (D \cap V) \setminus \{x_0\}$. Then, we have $f(x) \in U_A$ since $x \in (D \cap V_A) \setminus \{x_0\}$ and $f(x) \in U_B$ since $x \in (D \cap V_B) \setminus \{x_0\}$. Then, $f(x) \in U_A \cap U_B \neq \emptyset$. This contradicts with $U_A \cap U_B = \emptyset$. Hence, the result holds. This completes the proof of the proposition. $\square$

Proposition 3.74 Let $\mathcal{X} := (X, \mathcal{O}_X)$ and $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be topological spaces, $D \subseteq X$ with subset topology $\mathcal{O}_D$, $f : D \rightarrow \mathcal{Y}$, and $x_0 \in D$. Then, the following statements are equivalent.

- (i) f is continuous at x_0 .
- (ii) If x_0 is an accumulation point of D , then $f(x_0)$ is a limit point of $f(x)$ as $x \rightarrow x_0$.

Proof (i) $\Rightarrow$ (ii). This is straightforward.

(ii) $\Rightarrow$ (i). We will distinguish two exhaustive and mutually exclusive cases: Case 1: x_0 is not an accumulation point of D ; Case 2: x_0 is an accumulation point of D . Case 1: x_0 is not an accumulation point of D . $\exists V \in \mathcal{O}_X$ with $x_0 \in V$ such that $V \cap D = \{x_0\}$. $\forall U \in \mathcal{O}_Y$ with $f(x_0) \in U$, we have $f(V) = \{f(x_0)\} \subseteq U$. Hence, f is continuous at x_0 .

Case 2: x_0 is an accumulation point of D . $\forall U \in \mathcal{O}_Y$ with $f(x_0) \in U$, $\exists V \in \mathcal{O}_X$ with $x_0 \in V$ such that $f(V \setminus \{x_0\}) \subseteq U$. Then, we have $f(V) \subseteq U$. Hence, f is continuous at x_0 .

In both cases, f is continuous at x_0 .

This completes the proof of the proposition. $\square$

Proposition 3.75 Let $\mathcal{X} := (X, \mathcal{O}_X)$, $\mathcal{Y} := (Y, \mathcal{O}_Y)$, and $\mathcal{Z} := (Z, \mathcal{O}_Z)$ be topological spaces, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, $x_0 \in \mathcal{X}$ be an accumulation point of D , $y_0 \in \mathcal{Y}$ be a limit point of $f(x)$ as $x \rightarrow x_0$, and $g : \mathcal{Y} \rightarrow \mathcal{Z}$ be continuous at y_0 . Then, $g(y_0) \in \mathcal{Z}$ is a limit point of $g(f(x))$ as $x \rightarrow x_0$. When $\mathcal{Y}$ and $\mathcal{Z}$ are Hausdorff, then we may write $\lim_{x \rightarrow x_0} g(f(x)) = g(\lim_{x \rightarrow x_0} f(x))$.

Proof $\forall O_Z \in \mathcal{O}_Z$ with $g(y_0) \in O_Z$, by the continuity of g at y_0 , $\exists O_Y \in \mathcal{O}_Y$ with $y_0 \in O_Y$ such that $g(O_Y) \subseteq O_Z$. Since y_0 is the limit of $f(x)$ as $x \rightarrow x_0$, then $\exists O_X \in \mathcal{O}_X$ with $x_0 \in O_X$ such that $f(O_X \setminus \{x_0\}) \subseteq O_Y$. Then, $g(f(O_X \setminus \{x_0\})) \subseteq O_Z$. Hence, $g(f(x))$ converges to $g(y_0)$ as $x \rightarrow x_0$. This completes the proof of the proposition. $\square$

Proposition 3.76 Let $\mathcal{X}$ be a topological space, $\bar{D} \subseteq \mathcal{X}$, $x_0 \in \mathcal{X}$ be an accumulation point of $\bar{D}$, $\mathcal{Y}$ and $\mathcal{Z}$ be Hausdorff topological spaces, $D \subseteq \mathcal{Y}$, $y_0 \in \mathcal{Y}$ be an accumulation point of D , $f : \bar{D} \rightarrow D$, and $g : D \rightarrow \mathcal{Z}$. Assume that

(i) $\exists O_0 \in \mathcal{O}_X$ with $x_0 \in O_0$ such that $f(O_0 \setminus \{x_0\}) \subseteq D \setminus \{y_0\}$;

(ii) $\lim_{x \rightarrow x_0} f(x) = y_0$ and $\lim_{y \rightarrow y_0} g(y) = z_0 \in \mathcal{Z}$.

Then, $\lim_{x \rightarrow x_0} g(f(x)) = z_0$.

Proof $\forall O_Z \in \mathcal{O}_Z$ with $z_0 \in O_Z$, by $\lim_{y \rightarrow y_0} g(y) = z_0$, $\exists O_Y \in \mathcal{O}_Y$ with $y_0 \in O_Y$ such that $g(O_Y \setminus \{y_0\}) \subseteq O_Z$. By $\lim_{x \rightarrow x_0} f(x) = y_0$, $\exists O_X \in \mathcal{O}_X$ with $x_0 \in O_X$ such that $f(O_X \setminus \{x_0\}) \subseteq O_Y$. Let $O_1 := O_0 \cap O_X \in \mathcal{O}_X$. Clearly, $x_0 \in O_1$. Then, $\forall x \in (O_1 \cap \bar{D}) \setminus \{x_0\}$, we have $f(x) \in O_Y \cap (D \setminus \{y_0\}) = (O_Y \cap D) \setminus \{y_0\}$ and $g(f(x)) \in O_Z$. Hence, $g(f((O_1 \cap \bar{D}) \setminus \{x_0\})) \subseteq O_Z$. Hence, $\lim_{x \rightarrow x_0} g(f(x)) = z_0$. This completes the proof of the proposition. $\square$

Proposition 3.77 Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be Hausdorff topological spaces, $\bar{D} \subseteq \mathcal{X}$, $x_0 \in \mathcal{X}$ be an accumulation point of $\bar{D}$, $D \subseteq \mathcal{Y}$, $y_0 \in \mathcal{Y}$ be an accumulation point of D , $f : \bar{D} \rightarrow D$ be bijective, and $g : D \rightarrow \mathcal{Z}$. Assume that $\lim_{x \rightarrow x_0} f(x) = y_0$ and $\lim_{y \rightarrow y_0} f_{\text{inv}}(y) = x_0$. Then, $\lim_{x \rightarrow x_0} g(f(x)) = \lim_{y \rightarrow y_0} g(y)$ whenever one of the limits exists in $\mathcal{Z}$.

Proof We will prove the result by distinguishing two exhaustive cases:
 Case 1: $\lim_{y \rightarrow y_0} g(y) = z_0 \in \mathcal{Z}$; Case 2: $\lim_{x \rightarrow x_0} g(f(x)) = z_0 \in \mathcal{Z}$.

Case 1: $\lim_{y \rightarrow y_0} g(y) = z_0 \in \mathcal{Z}$. We will further distinguish three exhaustive and mutually exclusive subcases: Case 1a: $y_0 \notin D$; Case 1b: $y_0 \in D$ and $x_0 = f_{\text{inv}}(y_0)$; Case 1c: $y_0 \in D$ and $x_0 \neq \bar{x}_0 := f_{\text{inv}}(y_0)$. Case 1a: $y_0 \notin D$. Then, (i) of Proposition 3.76 is satisfied with $O_0 := \mathcal{X}$. By Proposition 3.76, we have $\lim_{x \rightarrow x_0} g(f(x)) = z_0 \in \mathcal{Z}$. Case 1b: $y_0 \in D$ and $x_0 = f_{\text{inv}}(y_0)$. Let $O_0 := \mathcal{X}$. $\forall x \in (O_0 \cap \bar{D}) \setminus \{x_0\}$, we have $D \ni f(x) \neq f(x_0) = y_0$ by f being bijective. This implies that $f(O_0 \setminus \{x_0\}) \subseteq D \setminus \{y_0\}$. Then, $\lim_{x \rightarrow x_0} g(f(x)) = z_0 \in \mathcal{Z}$ by Proposition 3.76. Case 1c: $y_0 \in D$ and $x_0 \neq \bar{x}_0 := f_{\text{inv}}(y_0)$. By $\mathcal{X}$ being Hausdorff, $\exists O_0 \in \mathcal{O}_{\mathcal{X}}$ such that $x_0 \in O_0$ and $\bar{x}_0 \notin O_0$. This leads to $\forall x \in (O_0 \cap \bar{D}) \setminus \{x_0\}$, we have $D \ni f(x) \neq f(\bar{x}_0) = y_0$ by f being bijective. This implies that $f(O_0 \setminus \{x_0\}) \subseteq D \setminus \{y_0\}$. Then, $\lim_{x \rightarrow x_0} g(f(x)) = z_0 \in \mathcal{Z}$ by Proposition 3.76. Hence, $\lim_{x \rightarrow x_0} g(f(x)) = \lim_{y \rightarrow y_0} g(y) = z_0 \in \mathcal{Z}$ in all three subcases. Hence, the result holds in this case.

Case 2: $\lim_{x \rightarrow x_0} g(f(x)) = z_0 \in \mathcal{Z}$. Define $h : \bar{D} \rightarrow \mathcal{Z}$ by $h(x) = g(f(x))$, $\forall x \in \bar{D}$. Then, $\lim_{x \rightarrow x_0} h(x) = z_0$. By Case 1, we have $\lim_{y \rightarrow y_0} h(f_{\text{inv}}(y)) = z_0 \in \mathcal{Z}$. Then, $\lim_{y \rightarrow y_0} g(f(f_{\text{inv}}(y))) = \lim_{y \rightarrow y_0} g(y) = z_0$. Hence, the result holds in this case.

This completes the proof of the proposition. $\square$

Example 3.78 Let $g : \mathbb{R} \rightarrow \mathbb{R}$. It is desired to calculate $\lim_{y \rightarrow +\infty} g(y)$. We will apply Proposition 3.77 to this calculation. Take $\mathcal{X} = \mathbb{R}$, $\mathcal{Y} = \mathbb{R}_e$, and $\mathcal{Z} = \mathbb{R}_e$. Let $\bar{D} := (-\infty, -1] \cup (0, +\infty) \subset \mathbb{R} = \mathcal{X}$, $x_0 = 0$, $D := \mathbb{R} \subset \mathbb{R}_e = \mathcal{Y}$, and $y_0 = +\infty$. Then, $g : D \rightarrow \mathcal{Z}$. Define $f : \bar{D} \rightarrow D$ by $f(x) = \begin{cases} 1/x & x > 0 \\ x+1 & x \leq -1 \end{cases}$, $\forall x \in \bar{D}$. Clearly, f is bijective with $f_{\text{inv}} : D \rightarrow \bar{D}$ given by $f_{\text{inv}}(y) = \begin{cases} 1/y & y > 0 \\ y-1 & y \leq 0 \end{cases}$, $\forall y \in D$. Clearly, $\lim_{x \rightarrow 0} f(x) = +\infty$ and $\lim_{y \rightarrow +\infty} f_{\text{inv}}(y) = 0$. Then, by Proposition 3.77, $\lim_{y \rightarrow +\infty} g(y) = \lim_{x \rightarrow 0} g(f(x))$ whenever one of the limits exists in $\mathcal{Z}$. $\diamond$

Proposition 3.79 Let $\mathcal{X} := (X, \mathcal{O}_X)$ and $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be topological spaces, $D \subseteq \mathcal{X}$, $x_0 \in \mathcal{X}$ be an accumulation point of D , $y_0 \in \mathcal{Y}$, and $f : D \rightarrow \mathcal{Y}$. Then, the following statements are equivalent.

- (i) y_0 is a limit point of $f(x)$ as $x \rightarrow x_0$.
- (ii) $\forall \text{ net } (x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D \setminus \{x_0\}$ with x_0 as a limit, we have y_0 is a limit point of the net $(f(x_\alpha))_{\alpha \in \mathcal{A}}$.

Proof (i) $\Rightarrow$ (ii). Fix any net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D \setminus \{x_0\}$ with x_0 as a limit. $\forall U \in \mathcal{O}_Y$ with $y_0 \in U$, by (i), $\exists V \in \mathcal{O}_X$ with $x_0 \in V$ such that $f(V \setminus \{x_0\}) \subseteq U$. $\exists \alpha_0 \in \mathcal{A}$ such that $\forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$, we have $x_\alpha \in V$. Then, $x_\alpha \in (D \cap V) \setminus \{x_0\}$ and $f(x_\alpha) \in U$. Hence, $(f(x_\alpha))_{\alpha \in \mathcal{A}}$ has a limit y_0 .

(ii) $\Rightarrow$ (i). Suppose y_0 is not a limit point of $f(x)$ as $x \rightarrow x_0$. Then, $\exists U_0 \in \mathcal{O}_Y$ with $y_0 \in U_0$, $\forall V \in \mathcal{O}_X$ with $x_0 \in V$, we have $f(V \setminus \{x_0\}) \not\subseteq U_0$. Then, $\exists x_V \in (D \cap V) \setminus \{x_0\}$ such that $f(x_V) \notin U_0$. Let $M := \{V \in \mathcal{O}_X \mid x_0 \in V\}$ and $\mathcal{A} := (M, \supseteq)$. Clearly, $\mathcal{A}$ is a directed system. By Axiom of Choice, we may construct a net $(x_V)_{V \in \mathcal{A}} \subseteq D \setminus \{x_0\}$. Clearly, x_0 is a limit point of this net. But, $\forall V \in \mathcal{A}$, $f(x_V) \notin U_0$. Hence, y_0 is not a limit point of the net $(f(x_V))_{V \in \mathcal{A}}$. This contradicts with (ii). Therefore, y_0 is a limit point of $f(x)$ as $x \rightarrow x_0$.

This completes the proof of the proposition. $\square$

Example 3.80 The extended real line is $\mathbb{R}_e := \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$. We assume that $\forall x \in \mathbb{R}$, $-\infty < x < \infty$. We define the usual operations: $\forall x \in \mathbb{R}$,

$$\begin{aligned} x + \infty &= \infty & x - \infty &= -\infty \\ x \cdot \infty &= \infty; & \text{if } x > 0 \\ x \cdot (-\infty) &= -\infty; & \text{if } x > 0 \\ \infty + \infty &= \infty & -\infty - \infty &= -\infty \\ \infty \cdot \infty &= \infty & \infty \cdot (-\infty) &= -\infty \end{aligned}$$

The operations $\infty - \infty$ and $0 \cdot \infty$ are undefined.

On $\mathbb{R}_e$, we introduce the countable collection of subsets of $\mathbb{R}_e$, $\mathcal{B}_{\mathbb{R}_e}$, as follows. $\emptyset, \mathbb{R}, \mathbb{R}_e \in \mathcal{B}_{\mathbb{R}_e}$. $\forall r_1, r_2 \in \mathbb{Q}$ with $r_1 < r_2$, $[-\infty, r_1), (r_1, r_2), (r_2, +\infty] \in \mathcal{B}_{\mathbb{R}_e}$.

By Proposition 3.18, it is easy to show that $\mathcal{B}_{\mathbb{R}_e}$ is a basis for a topology on $\mathbb{R}_e$. This topology is denoted $\mathcal{O}_{\mathbb{R}_e}$, which is the usual topology on $\mathbb{R}_e$. It is easy to show that $(\mathbb{R}_e, \mathcal{O}_{\mathbb{R}_e})$ is an arcwise connected second countable Hausdorff topological space.

An important property of $\mathbb{R}_e$ is that $\forall E \subseteq \mathbb{R}_e$, $\sup_{x \in E} x \in \mathbb{R}_e$ and $\inf_{x \in E} x \in \mathbb{R}_e$.

It is easy to see that $\mathbb{R}$ as a subset of $\mathbb{R}_e$ admits the subset topology $\mathcal{O}$ that equals to the usual topology $\mathcal{O}_{\mathbb{R}}$ on $\mathbb{R}$. $\diamond$

Proposition 3.81 Let X be a set, $f_1 : X \rightarrow \mathbb{R}$, $f_2 : X \rightarrow \mathbb{R}$, $g_1 : X \rightarrow \mathbb{R}_e$, and $g_2 : X \rightarrow \mathbb{R}_e$. Assume that $g_1(x) + g_2(x) \in \mathbb{R}_e$ is well defined, $\forall x \in X$. Then,

- (i) $\sup_{x \in X} g_1(x) \leq M \in \mathbb{R}_e$ if, and only if, $\forall x \in X$, $g_1(x) \leq M$;
- (ii) $\inf_{x \in X} g_1(x) \geq m \in \mathbb{R}_e$ if, and only if, $\forall x \in X$, $g_1(x) \geq m$;
- (iii) $\sup_{x \in X} g_1(x) > M \in \mathbb{R}_e$ if, and only if, $\exists x \in X$, $g_1(x) > M$;
- (iv) $\inf_{x \in X} g_1(x) < m \in \mathbb{R}_e$ if, and only if, $\exists x \in X$, $g_1(x) < m$;
- (v) $\sup_{x \in X} (-g_1)(x) = -\inf_{x \in X} g_1(x)$;
- (vi) $\sup_{x \in X} (f_1 + f_2)(x) \leq \sup_{x \in X} f_1(x) + \sup_{x \in X} f_2(x)$;

- (vii) $\forall \alpha \in (0, \infty) \subset \mathbb{R}$, $\sup_{x \in X}(\alpha g_1(x)) = \alpha \sup_{x \in X} g_1(x)$; $\forall \alpha \in [0, \infty) \subset \mathbb{R}$, $\sup_{x \in X}(\alpha f_1(x)) = \alpha \sup_{x \in X} f_1(x)$ when $\sup_{x \in X} f_1(x) \in \mathbb{R}$;
- (viii) $\sup_{x \in X}(g_1 + g_2)(x) \leq \sup_{x \in X} g_1(x) + \sup_{x \in X} g_2(x)$, when the right-hand-side is well defined.

Proof This is straightforward, and is therefore omitted. $\square$

Definition 3.82 Let $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathbb{R}_e$ be a net. The limit superior and limit inferior of the net are defined by

$$\begin{aligned} \limsup_{\alpha \in \mathcal{A}} x_\alpha &= \inf_{\alpha \in \mathcal{A}} \sup_{\beta \in \mathcal{A} \text{ with } \alpha \prec \beta} x_\beta \in \mathbb{R}_e \\ \liminf_{\alpha \in \mathcal{A}} x_\alpha &= \sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{A} \text{ with } \alpha \prec \beta} x_\beta \in \mathbb{R}_e \end{aligned}$$

Proposition 3.83 Let $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathbb{R}_e$, $(y_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathbb{R}$, and $(z_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathbb{R}$ be nets over the same directed system. Then, we have

- (i) $\liminf_{\alpha \in \mathcal{A}} x_\alpha \leq \limsup_{\alpha \in \mathcal{A}} x_\alpha$;
- (ii) $-\liminf_{\alpha \in \mathcal{A}} x_\alpha = \limsup_{\alpha \in \mathcal{A}} (-x_\alpha)$;
- (iii) $\liminf_{\alpha \in \mathcal{A}} x_\alpha = \limsup_{\alpha \in \mathcal{A}} x_\alpha = L \in \mathbb{R}_e$ if, and only if, $\lim_{\alpha \in \mathcal{A}} x_\alpha = L$;
- (iv) $\limsup_{\alpha \in \mathcal{A}} (y_\alpha + z_\alpha) \leq \limsup_{\alpha \in \mathcal{A}} y_\alpha + \limsup_{\alpha \in \mathcal{A}} z_\alpha$, when the right-hand-side makes sense;
- (v) if $\lim_{\alpha \in \mathcal{A}} y_\alpha = y \in \mathbb{R}$, then $\limsup_{\alpha \in \mathcal{A}} (y_\alpha + z_\alpha) = y + \limsup_{\alpha \in \mathcal{A}} z_\alpha$.

Proof Let $V_\alpha := \{\beta \in \mathcal{A} \mid \alpha \prec \beta\}$, $\forall \alpha \in \mathcal{A}$. Then, $V_\alpha \neq \emptyset$, $\forall \alpha \in \mathcal{A}$, since $\mathcal{A}$ is a directed system, and $V_\alpha \supseteq V_\beta$, $\forall \alpha, \beta \in \mathcal{A}$ with $\alpha \prec \beta$.

(i) Let $l := \liminf_{\alpha \in \mathcal{A}} x_\alpha \in \mathbb{R}_e$ and $L := \limsup_{\alpha \in \mathcal{A}} x_\alpha \in \mathbb{R}_e$. $\forall m \in \mathbb{R}$ with $m < l$, $\sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{A} \text{ with } \alpha \prec \beta} x_\beta > m$ implies that, by Proposition 3.81, $\exists \alpha_0 \in \mathcal{A}$ such that $\inf_{\beta \in V_{\alpha_0}} x_\beta > m$. Then, $\forall \alpha \in \mathcal{A}$, $\exists \alpha_1 \in \mathcal{A}$ such that $\alpha_0 \prec \alpha_1$ and $\alpha \prec \alpha_1$. Then, $m < \inf_{\beta \in V_{\alpha_0}} x_\beta \leq \inf_{\beta \in V_{\alpha_1}} x_\beta \leq \sup_{\beta \in V_{\alpha_1}} x_\beta \leq \sup_{\beta \in V_\alpha} x_\beta$. Hence, $L \geq m$. By the arbitrariness of m , we have $L \geq l$.

(ii) Note that, by Proposition 3.81,

$$\begin{aligned} \limsup_{\alpha \in \mathcal{A}} (-x_\alpha) &= \inf_{\alpha \in \mathcal{A}} \sup_{\beta \in V_\alpha} (-x_\beta) = \inf_{\alpha \in \mathcal{A}} \left(- \inf_{\beta \in V_\alpha} x_\beta \right) \\ &= - \sup_{\alpha \in \mathcal{A}} \inf_{\beta \in V_\alpha} x_\beta = - \liminf_{\alpha \in \mathcal{A}} x_\alpha \end{aligned}$$

(iii) “If” $\forall m \in \mathbb{R}$ with $m > L$, $\exists \alpha_0 \in \mathcal{A}$ such that $x_\alpha \in (-\infty, m)$, $\forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$. Then, $\sup_{\beta \in V_{\alpha_0}} x_\beta \leq m$ and $\limsup_{\alpha \in \mathcal{A}} x_\alpha \leq m$. By the arbitrariness of m , we have $\limsup_{\alpha \in \mathcal{A}} x_\alpha \leq L$. By (ii), we have

$-L = \lim_{\alpha \in \mathcal{A}} (-x_\alpha) \geq \limsup_{\alpha \in \mathcal{A}} (-x_\alpha) = -\liminf_{\alpha \in \mathcal{A}} x_\alpha$. Then, by (i), we have $L \leq \liminf_{\alpha \in \mathcal{A}} x_\alpha \leq \limsup_{\alpha \in \mathcal{A}} x_\alpha \leq L$.

“Only if” We will distinguish three exhaustive and mutually exclusive cases: Case 1: $L = -\infty$; Case 2: $L \in \mathbb{R}$; Case 3: $L = +\infty$. Case 1: $L = -\infty$. $\forall m \in \mathbb{R}$, $\limsup_{\alpha \in \mathcal{A}} x_\alpha < m$ implies that $\exists \alpha_0 \in \mathcal{A}$ such that $\sup_{\beta \in V_{\alpha_0}} x_\beta < m$. Then, $x_\beta \in (-\infty, m)$, $\forall \beta \in V_{\alpha_0}$. Hence, we have $\lim_{\alpha \in \mathcal{A}} x_\alpha = -\infty = L$. Case 2: $L \in \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $L = \liminf_{\alpha \in \mathcal{A}} x_\alpha$ implies that $\exists \alpha_1 \in \mathcal{A}$ such that $\inf_{\beta \in V_{\alpha_1}} x_\beta > L - \epsilon$. Then, $x_\beta \in (L - \epsilon, +\infty) \subset \mathbb{R}$, $\forall \beta \in V_{\alpha_1}$. $L = \limsup_{\alpha \in \mathcal{A}} x_\alpha$ implies that $\exists \alpha_2 \in \mathcal{A}$ such that $\sup_{\beta \in V_{\alpha_2}} x_\beta < L + \epsilon$. Then, $x_\beta \in (-\infty, L + \epsilon) \subset \mathbb{R}$, $\forall \beta \in V_{\alpha_2}$. Let $\alpha_0 \in \mathcal{A}$ with $\alpha_1 \prec \alpha_0$ and $\alpha_2 \prec \alpha_0$. Then, $x_\beta \in (L - \epsilon, L + \epsilon)$, $\forall \beta \in \mathcal{A}$ with $\alpha_0 \prec \beta$. Therefore, $\lim_{\alpha \in \mathcal{A}} x_\alpha = L$. Case 3: $L = +\infty$. $\forall M \in \mathbb{R}$, $\liminf_{\alpha \in \mathcal{A}} x_\alpha > M$ implies that $\exists \alpha_0 \in \mathcal{A}$ such that $\inf_{\beta \in V_{\alpha_0}} x_\beta > M$. Then, $x_\beta \in (M, \infty)$, $\forall \beta \in V_{\alpha_0}$. Hence, we have $\lim_{\alpha \in \mathcal{A}} x_\alpha = +\infty = L$.

(iv) Note that, $\forall \alpha \in \mathcal{A}$, by Proposition 3.81,

$$\sup_{\beta \in V_\alpha} (y_\beta + z_\beta) \leq \sup_{\beta \in V_\alpha} y_\beta + \sup_{\beta \in V_\alpha} z_\beta$$

Then, by Proposition 3.81, we have, $\forall \alpha \in \mathcal{A}$,

$$\limsup_{\alpha \in \mathcal{A}} (y_\alpha + z_\alpha) \leq \sup_{\beta \in V_\alpha} y_\beta + \sup_{\beta \in V_\alpha} z_\beta$$

$\forall \gamma \in \mathcal{A}$, $\exists \alpha_0 \in \mathcal{A}$ with $\alpha \prec \alpha_0$ and $\gamma \prec \alpha_0$. Then,

$$\limsup_{\alpha \in \mathcal{A}} (y_\alpha + z_\alpha) \leq \sup_{\beta \in V_{\alpha_0}} y_\beta + \sup_{\beta \in V_{\alpha_0}} z_\beta \leq \sup_{\beta \in V_\alpha} y_\beta + \sup_{\beta \in V_\gamma} z_\beta$$

Hence, we have $\limsup_{\alpha \in \mathcal{A}} (y_\alpha + z_\alpha) \leq \limsup_{\alpha \in \mathcal{A}} y_\alpha + \limsup_{\alpha \in \mathcal{A}} z_\alpha$, when the right-hand-side makes sense.

(v) By (iv) and (iii), we have $\limsup_{\alpha \in \mathcal{A}} (y_\alpha + z_\alpha) \leq y + \limsup_{\alpha \in \mathcal{A}} z_\alpha$. Note that $\limsup_{\alpha \in \mathcal{A}} z_\alpha = \limsup_{\alpha \in \mathcal{A}} (y_\alpha + z_\alpha - y_\alpha) \leq \limsup_{\alpha \in \mathcal{A}} (y_\alpha + z_\alpha) + \limsup_{\alpha \in \mathcal{A}} (-y_\alpha)$. Then, we have $\limsup_{\alpha \in \mathcal{A}} (y_\alpha + z_\alpha) \geq y + \limsup_{\alpha \in \mathcal{A}} z_\alpha$. Hence, we have $\limsup_{\alpha \in \mathcal{A}} (y_\alpha + z_\alpha) = y + \limsup_{\alpha \in \mathcal{A}} z_\alpha$.

This completes the proof of the proposition. $\square$

Definition 3.84 Let $\mathcal{X} := (X, \mathcal{O}_X)$ be topological spaces, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathbb{R}$, and $x_0 \in \mathcal{X}$ be an accumulation point of D . Then the limit superior and limit inferior of $f(x)$ as $x \rightarrow x_0$ are defined by

$$\begin{aligned} \limsup_{x \rightarrow x_0} f(x) &= \inf_{O \in \mathcal{O}} \sup_{\text{with } x_0 \in O} \sup_{x \in (D \cap O) \setminus \{x_0\}} f(x) \in \mathbb{R}_e \\ \liminf_{x \rightarrow x_0} f(x) &= \sup_{O \in \mathcal{O}} \inf_{\text{with } x_0 \in O} \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) \in \mathbb{R}_e \end{aligned}$$

Proposition 3.85 Let $\mathcal{X}$ be a topological space, $D \subseteq \mathcal{X}$, $x_0 \in \mathcal{X}$ be an accumulation point of D , $f : D \rightarrow \mathbb{R}$, and $g : D \rightarrow \mathbb{R}$. Then, we have

- (i) $\liminf_{x \rightarrow x_0} f(x) \leq \limsup_{x \rightarrow x_0} f(x)$;
- (ii) $-\liminf_{x \rightarrow x_0} f(x) = \limsup_{x \rightarrow x_0} (-f)(x)$;
- (iii) $\liminf_{x \rightarrow x_0} f(x) = \limsup_{x \rightarrow x_0} f(x) = L \in \mathbb{R}_e$ if, and only if, $\lim_{x \rightarrow x_0} f(x) = L$;
- (iv) $\limsup_{x \rightarrow x_0} (f + g)(x) \leq \limsup_{x \rightarrow x_0} f(x) + \limsup_{x \rightarrow x_0} g(x)$, when the right-hand-side makes sense;
- (v) if $\lim_{x \rightarrow x_0} f(x) = y \in \mathbb{R}$, then $\limsup_{x \rightarrow x_0} (f + g)(x) = y + \limsup_{x \rightarrow x_0} g(x)$.

Proof (i) Let $l := \liminf_{x \rightarrow x_0} f(x) \in \mathbb{R}_e$ and $L := \limsup_{x \rightarrow x_0} f(x) \in \mathbb{R}_e$. $\forall m \in \mathbb{R}$ with $m < l$, $\sup_{O \in \mathcal{O} \text{ with } x_0 \in O} \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) > m$ implies that $\exists U \in \mathcal{O}$ with $x_0 \in U$ such that $\inf_{x \in (D \cap U) \setminus \{x_0\}} f(x) > m$. $\forall O \in \mathcal{O}$ with $x_0 \in O$, we have $x_0 \in V := O \cap U \in \mathcal{O}$ and $(D \cap V) \setminus \{x_0\} \neq \emptyset$, since x_0 is an accumulation point of D . Then, $m < \inf_{x \in (D \cap U) \setminus \{x_0\}} f(x) \leq \inf_{x \in (D \cap V) \setminus \{x_0\}} f(x) \leq \sup_{x \in (D \cap V) \setminus \{x_0\}} f(x) \leq \sup_{x \in (D \cap O) \setminus \{x_0\}} f(x)$. Hence, $L \geq m$. By the arbitrariness of m , we have $L \geq l$.

(ii) Note that, by Proposition 3.81,

$$\begin{aligned}
 \limsup_{x \rightarrow x_0} (-f)(x) &= \inf_{O \in \mathcal{O} \text{ with } x_0 \in O} \sup_{x \in (D \cap O) \setminus \{x_0\}} (-f(x)) \\
 &= \inf_{O \in \mathcal{O} \text{ with } x_0 \in O} \left(- \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) \right) \\
 &= - \sup_{O \in \mathcal{O} \text{ with } x_0 \in O} \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) \\
 &= - \liminf_{x \rightarrow x_0} f(x)
 \end{aligned}$$

(iii) “If” $\forall m \in \mathbb{R}$ with $m > L$, $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $f(V \setminus \{x_0\}) \subseteq (-\infty, m)$. Then, $\sup_{x \in (D \cap V) \setminus \{x_0\}} f(x) \leq m$ and $\limsup_{x \rightarrow x_0} f(x) \leq m$.

By the arbitrariness of m , we have $\limsup_{x \rightarrow x_0} f(x) \leq L$. By (ii), we have $-L = \lim_{x \rightarrow x_0} (-f)(x) \geq \limsup_{x \rightarrow x_0} (-f)(x) = - \liminf_{x \rightarrow x_0} f(x)$. Then, by (i), we have $L \leq \liminf_{x \rightarrow x_0} f(x) \leq \limsup_{x \rightarrow x_0} f(x) \leq L$.

“Only if” We will distinguish three exhaustive and mutually exclusive cases: Case 1: $L = -\infty$; Case 2: $L \in \mathbb{R}$; Case 3: $L = +\infty$. Case 1: $L = -\infty$. $\forall m \in \mathbb{R}$, $\limsup_{x \rightarrow x_0} f(x) < m$ implies that $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $\sup_{x \in (D \cap V) \setminus \{x_0\}} f(x) < m$. Then, $f(V \setminus \{x_0\}) \subseteq (-\infty, m)$. Hence, we have $\lim_{x \rightarrow x_0} f(x) = -\infty = L$. Case 2: $L \in \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $L = \liminf_{x \rightarrow x_0} f(x)$ implies that $\exists V_1 \in \mathcal{O}$ with $x_0 \in V_1$ such that $\inf_{x \in (D \cap V_1) \setminus \{x_0\}} f(x) > L - \epsilon$. Then, $f(V_1 \setminus \{x_0\}) \subseteq (L - \epsilon, +\infty) \subseteq \mathbb{R}$. $L = \limsup_{x \rightarrow x_0} f(x)$ implies that $\exists V_2 \in \mathcal{O}$ with $x_0 \in V_2$ such that $\sup_{x \in (D \cap V_2) \setminus \{x_0\}} f(x) < L + \epsilon$. Then, $f(V_2 \setminus \{x_0\}) \subseteq (-\infty, L + \epsilon)$.

Let $V := V_1 \cap V_2 \in \mathcal{O}$. Clearly, $x_0 \in V$ and, by Proposition 2.5, $f(V \setminus \{x_0\}) \subseteq f(V_1 \setminus \{x_0\}) \cap f(V_2 \setminus \{x_0\}) \subseteq (L - \epsilon, L + \epsilon)$. Therefore, $\lim_{x \rightarrow x_0} f(x) = L$. Case 3: $L = +\infty$. $\forall M \in \mathbb{R}$, $\liminf_{x \rightarrow x_0} f(x) > M$ implies that $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $\inf_{x \in (D \cap V) \setminus \{x_0\}} f(x) > M$. Then, $f(V \setminus \{x_0\}) \subseteq (M, +\infty)$. Hence, we have $\lim_{x \rightarrow x_0} f(x) = +\infty = L$.

(iv) Note that, $\forall O \in \mathcal{O}$ with $x_0 \in O$, by Proposition 3.81,

$$\sup_{x \in (D \cap O) \setminus \{x_0\}} (f + g)(x) \leq \left(\sup_{x \in (D \cap O) \setminus \{x_0\}} f(x) \right) + \left(\sup_{x \in (D \cap O) \setminus \{x_0\}} g(x) \right)$$

Then, by Proposition 3.81, we have, $\forall O \in \mathcal{O}$ with $x_0 \in O$,

$$\limsup_{x \rightarrow x_0} (f + g)(x) \leq \left(\sup_{x \in (D \cap O) \setminus \{x_0\}} f(x) \right) + \left(\sup_{x \in (D \cap O) \setminus \{x_0\}} g(x) \right)$$

$\forall U \in \mathcal{O}$ with $x_0 \in U$, we have $x_0 \in V := U \cap O \in \mathcal{O}$. Then,

$$\begin{aligned} \limsup_{x \rightarrow x_0} (f + g)(x) &\leq \left(\sup_{x \in (D \cap V) \setminus \{x_0\}} f(x) \right) + \left(\sup_{x \in (D \cap V) \setminus \{x_0\}} g(x) \right) \\ &\leq \left(\sup_{x \in (D \cap O) \setminus \{x_0\}} f(x) \right) + \left(\sup_{x \in (D \cap U) \setminus \{x_0\}} g(x) \right) \end{aligned}$$

Hence, we have $\limsup_{x \rightarrow x_0} (f + g)(x) \leq \limsup_{x \rightarrow x_0} f(x) + \limsup_{x \rightarrow x_0} g(x)$, when the right-hand-side makes sense.

(v) By (iv) and (iii), $\limsup_{x \rightarrow x_0} (f + g)(x) \leq y + \limsup_{x \rightarrow x_0} g(x)$. Note that $\limsup_{x \rightarrow x_0} g(x) = \limsup_{x \rightarrow x_0} (f + g - f)(x) \leq \limsup_{x \rightarrow x_0} (f + g)(x) + \limsup_{x \rightarrow x_0} (-f)(x)$. Then, we have $\limsup_{x \rightarrow x_0} (f + g)(x) \geq y + \limsup_{x \rightarrow x_0} g(x)$. Hence, we have $\limsup_{x \rightarrow x_0} (f + g)(x) = y + \limsup_{x \rightarrow x_0} g(x)$.

This completes the proof of the proposition. $\square$

Proposition 3.86 *Let $\mathcal{X} := (X, \mathcal{O}_X)$ be a topological space, $D \subseteq X$ with the subset topology $\mathcal{O}_D$, $f : D \rightarrow \mathbb{R}$, and $x_0 \in D$. Then, the following statements are equivalent.*

(i) *f is upper semicontinuous at x_0 .*

(ii) *If x_0 is an accumulation point of D , then $\limsup_{x \rightarrow x_0} f(x) \leq f(x_0)$.*

Proof (i) $\Rightarrow$ (ii). Let x_0 be an accumulation point of D . By the upper semicontinuity of f at x_0 , $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists O \in \mathcal{O}_X$ with $x_0 \in O$ such that $f(x) < f(x_0) + \epsilon$, $\forall x \in O \cap D$. Then, $\sup_{x \in (O \cap D) \setminus \{x_0\}} f(x) \leq f(x_0) + \epsilon$, and $\limsup_{x \rightarrow x_0} f(x) \leq f(x_0) + \epsilon$. By the arbitrariness of ϵ , we have $\limsup_{x \rightarrow x_0} f(x) \leq f(x_0)$.

(ii) $\Rightarrow$ (i). We will distinguish two exhaustive and mutually exclusive cases: Case 1: x_0 is not an accumulation point of D ; Case 2: x_0 is an accumulation point of D . Case 1: x_0 is not an accumulation point of D . $\exists V \in \mathcal{O}_X$ with $x_0 \in V$ such that $V \cap D = \{x_0\}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, we have

$f(x) = f(x_0) < f(x_0) + \epsilon$, $\forall x \in V \cap D$. Hence, f is upper semicontinuous at x_0 .

Case 2: x_0 is an accumulation point of D . $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\limsup_{x \rightarrow x_0} f(x) < f(x_0) + \epsilon$ implies that $\exists O \in \mathcal{O}$ with $x_0 \in O$ such that $\sup_{x \in (O \cap D) \setminus \{x_0\}} f(x) < f(x_0) + \epsilon$. Then, $f(x) < f(x_0) + \epsilon$, $\forall x \in O \cap D$. Hence, f is upper semicontinuous at x_0 .

In both cases, f is upper semicontinuous at x_0 .

This completes the proof of the proposition. $\square$

Proposition 3.87 *Let $\mathcal{X}$ be a topological space, $D \subseteq \mathcal{X}$, $x_0 \in \mathcal{X}$ be an accumulation point of D , and $f : D \rightarrow \mathbb{R}$. Assume there exists a net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D \setminus \{x_0\}$ such that $\lim_{\alpha \in \mathcal{A}} x_\alpha = x_0$ and $\liminf_{\alpha \in \mathcal{A}} f(x_\alpha) = c \in \mathbb{R}_e$. Then, $\liminf_{x \rightarrow x_0} f(x) \leq c$.*

Proof By Definitions 3.84 and 3.82, we have

$$\begin{aligned} \liminf_{x \rightarrow x_0} f(x) &= \sup_{O \in \mathcal{O} \text{ with } x_0 \in O} \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) \\ \liminf_{\alpha \in \mathcal{A}} f(x_\alpha) &= \sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{A} \text{ with } \alpha \prec \beta} f(x_\beta) = c \end{aligned}$$

$\forall O \in \mathcal{O}$ with $x_0 \in O$, since $\lim_{\alpha \in \mathcal{A}} x_\alpha = x_0$, then $\exists \alpha_0 \in \mathcal{A}$ such that $\forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$, we have $x_\alpha \in O$. Then, $x_\alpha \in (O \cap D) \setminus \{x_0\}$. This leads to

$$\inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) \leq \inf_{\beta \in \mathcal{A} \text{ with } \alpha_0 \prec \beta} f(x_\beta) \leq c$$

Hence, we have $\liminf_{x \rightarrow x_0} f(x) \leq c$. This completes the proof of the proposition. $\square$

Definition 3.88 *Let $\mathcal{A}_i := (A_i, \prec_i)$ be a directed system, $i = 1, 2$, and $\mathcal{X} := (X, \mathcal{O})$ be a topological space. Define a relation $\prec$ on $A_1 \times A_2$ by, $\forall (\alpha_1, \beta_1), (\alpha_2, \beta_2) \in A_1 \times A_2$, we say $(\alpha_1, \beta_1) \prec (\alpha_2, \beta_2)$ if $\alpha_1 \prec_1 \alpha_2$ and $\beta_1 \prec_2 \beta_2$. It is easy to verify that $\mathcal{A} := (A_1 \times A_2, \prec) =: \mathcal{A}_1 \times \mathcal{A}_2$ is a directed system. A joint net is a mapping of the directed system $\mathcal{A}_1 \times \mathcal{A}_2$ to $\mathcal{X}$, denoted by $(x_{\alpha, \beta})_{(\alpha, \beta) \in \mathcal{A}_1 \times \mathcal{A}_2}$. The joint net is said to admit a (joint) limit point $\hat{x} \in \mathcal{X}$ if it admits a limit point $\hat{x}$ when viewed as a net over the directed system $\mathcal{A}$. When $\mathcal{X}$ is Hausdorff, by Proposition 3.65, the joint net admits at most one joint limit point, which will be denoted by $\lim_{(\alpha, \beta) \in \mathcal{A}_1 \times \mathcal{A}_2} x_{\alpha, \beta} \in \mathcal{X}$ if it exists.*

Proposition 3.89 *Let $X := [0, \infty] \subset \mathbb{R}_e$ with subset topology $\mathcal{O}$, $\mathcal{X} := (X, \mathcal{O})$, and $c \in (0, \infty) \subset \mathbb{R}$. Then, $+$: $\mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ and $c \cdot$: $\mathcal{X} \rightarrow \mathcal{X}$ are continuous.*

Proof We will prove this using Proposition 3.9. $\forall (x_1, x_2) \in X \times X$. We will distinguish four exhaustive and mutually exclusive cases. Case 1: $x_1 < \infty$ and $x_2 < \infty$. Then, $x_1 + x_2 < \infty$. $\forall$ basis open set $U := (r_1, r_2) \cap X \in \mathcal{O}$

with $x_1 + x_2 \in U$, take $V_1 := (x_1 - (x_1 + x_2 - r_1)/2, x_1 + (r_2 - x_1 - x_2)/2) \cap X \in \mathcal{O}$ and $V_2 := (x_2 - (x_1 + x_2 - r_1)/2, x_2 + (r_2 - x_1 - x_2)/2) \cap X \in \mathcal{O}$. Then, $V_1 \times V_2$ is an open set in $\mathcal{X} \times \mathcal{X}$ with $(x_1, x_2) \in V_1 \times V_2$ and $\forall (\bar{x}_1, \bar{x}_2) \in V_1 \times V_2$, we have $\bar{x}_1 + \bar{x}_2 \in U$. Hence, $+: \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ is continuous at (x_1, x_2) .

Case 2: $x_1 < \infty$ and $x_2 = \infty$. Then, $x_1 + x_2 = \infty$. $\forall$ basis open set $U := (r_1, \infty] \cap X \in \mathcal{O}$ with $x_1 + x_2 \in U$, take $V_1 := (x_1 - 1, x_1 + 1) \cap X \in \mathcal{O}$ and $V_2 := (r_1 - x_1 + 1, \infty] \cap X \in \mathcal{O}$. Then, $V_1 \times V_2$ is an open set in $\mathcal{X} \times \mathcal{X}$ with $(x_1, x_2) \in V_1 \times V_2$ and $\forall (\bar{x}_1, \bar{x}_2) \in V_1 \times V_2$, we have $\bar{x}_1 + \bar{x}_2 \in U$. Hence, $+: \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ is continuous at (x_1, x_2) .

Case 3: $x_1 = \infty$ and $x_2 < \infty$. By Case 2 and symmetry, $+: \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ is continuous at (x_1, x_2) .

Case 4: $x_1 = \infty$ and $x_2 = \infty$. Then, $x_1 + x_2 = \infty$. $\forall$ basis open set $U := (r_1, \infty] \cap X \in \mathcal{O}$ with $x_1 + x_2 \in U$, take $V_1 := (r_1/2, \infty] \cap X \in \mathcal{O}$ and $V_2 := (r_1/2, \infty] \cap X \in \mathcal{O}$. Then, $V_1 \times V_2$ is an open set in $\mathcal{X} \times \mathcal{X}$ with $(x_1, x_2) \in V_1 \times V_2$ and $\forall (\bar{x}_1, \bar{x}_2) \in V_1 \times V_2$, we have $\bar{x}_1 + \bar{x}_2 \in U$. Hence, $+: \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ is continuous at (x_1, x_2) .

In all cases, we have $+: \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ is continuous at (x_1, x_2) . By Proposition 3.9, $+: \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ is continuous.

$\forall x \in X$. We will distinguish two exhaustive and mutually exclusive cases. Case 1: $x < \infty$. Then, $cx < \infty$. $\forall$ basis open set $U := (r_1, r_2) \cap X \in \mathcal{O}$ with $cx \in U$, take $V := (r_1/c, r_2/c) \cap X \in \mathcal{O}$. Then, $x \in V$ and $\forall \bar{x} \in V$, we have $c\bar{x} \in U$. Hence, $c: \mathcal{X} \rightarrow \mathcal{X}$ is continuous at x .

Case 2: $x = \infty$. Then, $cx = \infty$. $\forall$ basis open set $U := (r_1, \infty] \cap X \in \mathcal{O}$ with $cx \in U$, take $V := (r_1/c, \infty] \cap X \in \mathcal{O}$. Then, $x \in V$ and $\forall \bar{x} \in V$, we have $c\bar{x} \in U$. Hence, $c: \mathcal{X} \rightarrow \mathcal{X}$ is continuous at x .

In both cases, we have $c: \mathcal{X} \rightarrow \mathcal{X}$ is continuous at x . By Proposition 3.9, $c: \mathcal{X} \rightarrow \mathcal{X}$ is continuous. This completes the proof of the proposition. $\square$

Chapter 4

Metric Spaces

4.1 Fundamental Notions

Definition 4.1 A metric space (X, ρ) is a set X together with a metric $\rho : X \times X \rightarrow \mathbb{R}$ such that, $\forall x, y, z \in X$,

- (i) $\rho(x, y) \geq 0$;
- (ii) $\rho(x, y) = 0 \Leftrightarrow x = y$;
- (iii) $\rho(x, y) = \rho(y, x)$;
- (iv) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$.

Let $S \subseteq X$. Then, $(S, \rho|_{S \times S})$ is also a metric space.

Example 4.2 $(\mathbb{R}, \rho)$, with $\rho(x, y) = |x - y|$, $\forall x, y \in \mathbb{R}$, is a metric space. $(\mathbb{R}^n, \rho)$, with $\rho(x, y) = (\sum_{i=1}^n |x_i - y_i|^2)^{1/2}$, $\forall x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{R}^n$, is a metric space, where $n \in \mathbb{N}$. $(\mathbb{R}^n, \rho)$, with $\rho(x, y) = \sum_{i=1}^n |x_i - y_i|$, $\forall x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{R}^n$, is a metric space, where $n \in \mathbb{N}$. $\diamond$

Proposition 4.3 Let $\mathcal{X} := (X, \rho)$ be a metric space, an open ball centered at $x_0 \in X$ with radius $r \in (0, \infty) \subset \mathbb{R}$ is defined by $\mathcal{B}_{\mathcal{X}}(x_0, r) := \{x \in X \mid \rho(x, x_0) < r\}$.

The metric space generates a natural topology $\mathcal{O}$ on X with the basis collection $\mathcal{B}$ given by $\mathcal{B} := \{\mathcal{B}_{\mathcal{X}}(x_0, r) \mid x_0 \in X, r \in (0, \infty) \subset \mathbb{R}\}$.

The closed ball centered at $x_0 \in X$ with radius $r \in [0, \infty) \subset \mathbb{R}$ is defined by $\overline{\mathcal{B}}_{\mathcal{X}}(x_0, r) := \{x \in X \mid \rho(x, x_0) \leq r\}$, which is closed.

Proof We will show that $\mathcal{B}$ is the basis for its generated topology by Proposition 3.18. (i) $\forall x \in X, x \in \mathcal{B}_{\mathcal{X}}(x, 1)$. (ii) $\forall \mathcal{B}_{\mathcal{X}}(x_1, r_1), \mathcal{B}_{\mathcal{X}}(x_2, r_2) \in \mathcal{B}$, let $x \in \mathcal{B}_{\mathcal{X}}(x_1, r_1) \cap \mathcal{B}_{\mathcal{X}}(x_2, r_2)$. Then, we have $\rho(x, x_1) < r_1$ and $\rho(x, x_2) < r_2$. Let $r := \min\{r_1 - \rho(x, x_1), r_2 - \rho(x, x_2)\} \in (0, \infty) \subset \mathbb{R}$.

Then, $x \in \mathcal{B}_{\mathcal{X}}(x, r) \in \mathcal{B}$. $\forall x_3 \in \mathcal{B}_{\mathcal{X}}(x, r)$, we have $\rho(x_3, x_1) \leq \rho(x_3, x) + \rho(x, x_1) < r_1$ and $\rho(x_3, x_2) \leq \rho(x_3, x) + \rho(x, x_2) < r_2$. Hence, we have $x_3 \in \mathcal{B}_{\mathcal{X}}(x_1, r_1) \cap \mathcal{B}_{\mathcal{X}}(x_2, r_2)$. Therefore, we have $\mathcal{B}_{\mathcal{X}}(x, r) \subseteq \mathcal{B}_{\mathcal{X}}(x_1, r_1) \cap \mathcal{B}_{\mathcal{X}}(x_2, r_2)$. Hence, the assumptions of Proposition 3.18 are satisfied, and then $\mathcal{B}$ is a basis for $\mathcal{O}$.

Next, we show that $\overline{\mathcal{B}_{\mathcal{X}}}(x_0, r)$ is a closed set. $\forall x \in (\overline{\mathcal{B}_{\mathcal{X}}}(x_0, r))^{\sim}$, we have $\rho(x, x_0) > r$. Let $r_1 := \rho(x, x_0) - r \in (0, \infty) \subset \mathbb{R}$. $\forall x_1 \in \mathcal{B}_{\mathcal{X}}(x, r_1)$, we have $\rho(x_1, x_0) \geq \rho(x, x_0) - \rho(x, x_1) > r$. Hence, we have $x \in \mathcal{B}_{\mathcal{X}}(x, r_1) \subseteq (\overline{\mathcal{B}_{\mathcal{X}}}(x_0, r))^{\sim}$. Therefore, $(\overline{\mathcal{B}_{\mathcal{X}}}(x_0, r))^{\sim}$ is open and $\overline{\mathcal{B}_{\mathcal{X}}}(x_0, r)$ is closed.

This completes the proof of the proposition. $\square$

We will some times talk about a metric space $\mathcal{X} := (X, \rho)$ without referring the the components of $\mathcal{X}$, where the metric is understood to be $\rho_{\mathcal{X}}$ the natural topology is understood to be $\mathcal{O}_{\mathcal{X}}$. When it is clear from the context, we will also neglect the subscript $\mathcal{X}$. We will abuse the notation and say $x \in \mathcal{X}$ and $A \subseteq \mathcal{X}$ when $x \in X$ and $A \subseteq X$.

On a metric space, we can talk about open and closed sets, and all those concepts defined in Chapter 3, all with respect to the natural topology.

A metric space is clearly first countable, where a countable basis at $x_0 \in \mathcal{X}$ is $\{\mathcal{B}(x_0, r) \mid r \in \mathbb{Q}, r > 0\}$.

Proposition 4.4 *A metric space (X, ρ) is separable if, and only if, it is second countable.*

Proof “Only if” Let $D \subseteq X$ be a countable dense set. Let $\mathcal{M} := \{\mathcal{B}(x, r) \mid x \in D, r \in \mathbb{Q}, \text{ and } r > 0\}$. Clearly, $\mathcal{M}$ is countable and $\mathcal{M} \subseteq \mathcal{O}$. $\forall O \in \mathcal{O}$, $\forall x \in O$. $\exists r \in (0, \infty) \cap \mathbb{Q}$ such that $\mathcal{B}(x, r) \subseteq O$. Since D is dense, then $\exists x_1 \in \mathcal{B}(x, r/2) \cap D$. Let $r_1 = r/2 \in (0, \infty) \cap \mathbb{Q}$. Then, we have $x \in \mathcal{B}(x_1, r_1) \subseteq \mathcal{B}(x, r) \subseteq O$ and $\mathcal{B}(x_1, r_1) \in \mathcal{M}$. Hence, $\mathcal{M}$ is a basis for $\mathcal{O}$. Hence, $(X, \mathcal{O})$ is second countable.

“If” Let $\mathcal{O}$ has a countable basis $\mathcal{B}$. By Axiom of Choice, we may assign a $x_B \in B$, $\forall B \in \mathcal{B}$ with $B \neq \emptyset$. Let $D = \{x_B \in X \mid B \in \mathcal{B}, B \neq \emptyset\}$. Then, D is countable. $\forall x \in X$, $\forall O \in \mathcal{O}$ with $x \in O$, $\exists B \in \mathcal{B}$ such that $x \in B \subseteq O$. Then, $x_B \in D \cap B \subseteq D \cap O \neq \emptyset$. Hence, by Proposition 3.3, we have $x \in \overline{D}$. Therefore, by the arbitrariness of x , we have D is dense. Hence, $(X, \mathcal{O})$ is separable.

This completes the proof of the proposition. $\square$

Proposition 4.5 *Let $\mathcal{X}$ be a topological space and $\mathcal{Y}$ and $\mathcal{Z}$ be metric spaces. Let $f : \mathcal{X} \rightarrow \mathcal{Y}$, $g : \mathcal{Y} \rightarrow \mathcal{Z}$, $h : \mathcal{Y} \rightarrow \mathcal{X}$, $x_0 \in \mathcal{X}$, and $y_0 \in \mathcal{Y}$. Then, the following statements hold.*

1. *f is continuous at x_0 if, and only if, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x_0 \in U$ such that $\rho_{\mathcal{Y}}(f(x), f(x_0)) < \epsilon$, $\forall x \in U$.*
2. *g is continuous at y_0 if, and only if, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\rho_{\mathcal{Z}}(g(y), g(y_0)) < \epsilon$, $\forall y \in \mathcal{B}_{\mathcal{Y}}(y_0, \delta)$.*

3. h is continuous at y_0 if, and only if, $\forall U \in \mathcal{O}_{\mathcal{X}}$ with $h(y_0) \in U$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $h(y) \in U$, $\forall y \in \mathcal{B}_{\mathcal{Y}}(y_0, \delta)$.

Proof The proof is straightforward, and is therefore omitted. $\square$

Definition 4.6 Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be a homeomorphism. f is said to be an isometry between $\mathcal{X}$ and $\mathcal{Y}$ if $\rho_{\mathcal{Y}}(f(x_1), f(x_2)) = \rho_{\mathcal{X}}(x_1, x_2)$, $\forall x_1, x_2 \in \mathcal{X}$. Then, the two metric spaces are said to be isometric.

Definition 4.7 Let X be a set and ρ_1 and ρ_2 be two metrics on X . ρ_1 and ρ_2 are said to be equivalent if the identity map from (X, ρ_1) to (X, ρ_2) is a homeomorphism.

When two metrics are equivalent, then the natural topologies generated by them are equal to each other.

Clearly, a metric space is Hausdorff.

4.2 Convergence and Completeness

Proposition 4.8 Let $\mathcal{X}$ be a metric space and $(x_{\alpha})_{\alpha \in \mathcal{A}} \subseteq \mathcal{X}$ be a net. Then, $\lim_{\alpha \in \mathcal{A}} x_{\alpha} = x \in \mathcal{X}$ if, and only if, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \alpha_0 \in \mathcal{A}$, $\forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$, we have $\rho(x_{\alpha}, x) < \epsilon$.

Proof This is straightforward, and is omitted. $\square$

Since metric spaces are Hausdorff, then the limit is unique if it exists.

Definition 4.9 Let $\mathcal{X}$ be a metric space, $x_0 \in \mathcal{X}$, and $S \subseteq \mathcal{X}$. The distance from x_0 to S is $\text{dist}(x_0, S) := \inf_{s \in S} \rho(x_0, s) \in [0, \infty] \subset \mathbb{R}_e$.

$\text{dist}(x_0, S) = \infty$ if, and only if, $S = \emptyset$.

Proposition 4.10 Let $\mathcal{X}$ be a metric space, $x_0 \in \mathcal{X}$, $S \subseteq \mathcal{X}$, and S is closed. Then, $x_0 \in S$ if, and only if, $\text{dist}(x_0, S) = 0$.

Proof “Only if” This is obvious.

“If” By the fact that $\text{dist}(x_0, S) = 0$, $\forall n \in \mathbb{N}$, $\exists x_n \in S$ such that $\rho(x_0, x_n) < 1/n$. Then, $\lim_{n \in \mathbb{N}} x_n = x_0$. By Proposition 3.68, $x_0 \in \overline{S}$. Since S is closed, then, by Proposition 3.3, $S = \overline{S}$. Hence, $x_0 \in S$.

This completes the proof of the proposition. $\square$

Proposition 4.11 A metric space with its natural topology is normal.

Proof Let $\mathcal{X}$ be the metric space. Clearly, $\mathcal{X}$ is Hausdorff. $\forall$ closed sets $F_1, F_2 \subseteq \mathcal{X}$ with $F_1 \cap F_2 = \emptyset$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $F_1 = \emptyset$ or $F_2 = \emptyset$; Case 2: $F_1 \neq \emptyset$ and $F_2 \neq \emptyset$. Case 1: $F_1 = \emptyset$ or $F_2 = \emptyset$. Without loss

of generality, assume $F_1 = \emptyset$. Take $O_1 = \emptyset \in \mathcal{O}$ and $O_2 = \mathcal{X} \in \mathcal{O}$, then $F_1 \subseteq O_1$, $F_2 \subseteq O_2$, $O_1 \cap O_2 = \emptyset$. Case 2: $F_1 \neq \emptyset$ and $F_2 \neq \emptyset$. $\forall x \in F_1$, $\text{dist}(x, F_2) \in (0, \infty) \subset \mathbb{R}$, by Proposition 4.10. Define $O_1 \in \mathcal{O}$ by $O_1 := \bigcup_{x \in F_1} \mathcal{B}(x, \text{dist}(x, F_2)/3)$. $\forall x \in F_2$, $\text{dist}(x, F_1) \in (0, \infty) \subset \mathbb{R}$ by Proposition 4.10. Define $O_2 \in \mathcal{O}$ by $O_2 := \bigcup_{x \in F_2} \mathcal{B}(x, \text{dist}(x, F_1)/3)$. Clearly, $F_1 \subseteq O_1$ and $F_2 \subseteq O_2$. Note that $O_1 \cap O_2 = \emptyset$, since otherwise, $\exists x_0 \in O_1 \cap O_2$, $\exists x_1 \in F_1$ such that $x_0 \in \mathcal{B}(x_1, \text{dist}(x_1, F_2)/3)$, $\exists x_2 \in F_2$ such that $x_0 \in \mathcal{B}(x_2, \text{dist}(x_2, F_1)/3)$, without loss of generality, assume $\text{dist}(x_1, F_2) \leq \text{dist}(x_2, F_1)$, then $\text{dist}(x_2, F_1) \leq \rho(x_2, x_1) \leq \rho(x_2, x_0) + \rho(x_0, x_1) < \text{dist}(x_2, F_1)/3 + \text{dist}(x_1, F_2)/3 \leq 2 \text{dist}(x_2, F_1)/3$, which is a contradiction.

Hence, in both cases, $\exists O_1, O_2 \in \mathcal{O}$ such that $F_1 \subseteq O_1$, $F_2 \subseteq O_2$, $O_1 \cap O_2 = \emptyset$. Hence, $\mathcal{X}$ is normal. This completes the proof of the proposition. $\square$

Definition 4.12 Let $\mathcal{X}$ be a metric space and $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$. The sequence is said to be a Cauchy sequence if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$, $\forall n, m \geq N$, $\rho(x_n, x_m) < \epsilon$.

Clearly, every convergent sequence in a metric space is a Cauchy sequence.

Proposition 4.13 Let $\mathcal{X}$ be a metric space, $E \subseteq \mathcal{X}$, and $x_0 \in \mathcal{X}$. Then, $x_0 \in \overline{E}$ if, and only if, $\exists (x_n)_{n=1}^\infty \subseteq E$ such that $\lim_{n \in \mathbb{N}} x_n = x_0$.

Proof “Only if” $\forall n \in \mathbb{N}$, since $x_0 \in \overline{E}$, then, by Proposition 3.3, $\exists x_n \in E \cap \mathcal{B}(x_0, 1/n)$. Clearly, $(x_n)_{n=1}^\infty \subseteq E$ and $\lim_{n \in \mathbb{N}} x_n = x_0$.

“If” This is immediate by Proposition 3.68.

This completes the proof of the proposition. $\square$

Definition 4.14 A metric space is said to be complete if every Cauchy sequence in the metric space converges to a point in the space.

Proposition 4.15 Let $\mathcal{X} := (X, \rho)$ be a metric space, $\mathcal{Y} := (Y, \mathcal{O})$ be a topological space, $f : \mathcal{X} \rightarrow \mathcal{Y}$, and $x_0 \in \mathcal{X}$. Then, the following statements are equivalent.

- (i) f is continuous at x_0 ;
- (ii) if x_0 is an accumulation point of $\mathcal{X}$, then $f(x)$ converges to $f(x_0)$ as $x \rightarrow x_0$;
- (iii) $\forall (x_n)_{n=1}^\infty \subseteq \mathcal{X}$ with $\lim_{n \in \mathbb{N}} x_n = x_0$, we have $(f(x_n))_{n=1}^\infty \subseteq \mathcal{Y}$ converges to $f(x_0)$;
- (iv) $\forall (x_n)_{n=1}^\infty \subseteq \mathcal{X}$ with x_0 as a cluster point, we have that $(f(x_n))_{n=1}^\infty \subseteq \mathcal{Y}$ admits a cluster point $f(x_0)$.

Proof (i) $\Leftrightarrow$ (ii). This follows from Proposition 3.74.

(i) $\Rightarrow$ (iii). This follows from Proposition 3.66.

(iii) $\Rightarrow$ (iv). $\forall (x_n)_{n=1}^\infty \subseteq \mathcal{X}$ with x_0 as a cluster point. Since $\mathcal{X}$ is a metric space and therefore first countable, then $\exists$ a subsequence $(x_{n_i})_{i=1}^\infty$ of $(x_n)_{n=1}^\infty$ such that $\lim_{i \in \mathbb{N}} x_{n_i} = x_0$. Then, by (iii), $(f(x_{n_i}))_{i=1}^\infty$ converges to $f(x_0)$. Then, $f(x_0)$ is a cluster point of $(f(x_{n_i}))_{i=1}^\infty$ and therefore a cluster point of $(f(x_n))_{n=1}^\infty$.

(iv) $\Rightarrow$ (i). Suppose f is not continuous at x_0 . $\exists O_{Y0} \in \mathcal{O}_Y$ with $f(x_0) \in O_{Y0}$ such that $\forall n \in \mathbb{N}$, we have $f(\mathcal{B}(x_0, 1/n)) \not\subseteq O_{Y0}$. Then, $\exists x_n \in \mathcal{B}(x_0, 1/n)$ such that $f(x_n) \in \widetilde{O_{Y0}}$. Consider the sequence $(x_n)_{n=1}^\infty$. Clearly, $x_0 = \lim_{n \in \mathbb{N}} x_n$ and therefore is a cluster point of the sequence. Consider the sequence $(f(x_n))_{n=1}^\infty$. For the open set $O_{Y0} \ni f(x_0)$, $\forall n \in \mathbb{N}$, $f(x_n) \in \widetilde{O_{Y0}}$. Then, $f(x_0)$ is not a cluster point of $(f(x_n))_{n=1}^\infty$. This contradicts with the assumption. Therefore, f must be continuous at x_0 .

This completes the proof of the proposition. $\square$

Proposition 4.16 Let $\mathcal{X} := (X, \rho)$ be a metric space, $\mathcal{Y} := (Y, \mathcal{O})$ be a topological space, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, $x_0 \in \mathcal{X}$ be an accumulation point of D , and $y_0 \in \mathcal{Y}$. Then, the following statements are equivalent.

(i) $f(x)$ converges to y_0 as $x \rightarrow x_0$.

(ii) $\forall (x_n)_{n=1}^\infty \subseteq D \setminus \{x_0\}$ with $\lim_{n \in \mathbb{N}} x_n = x_0$, we have that y_0 is a limit point of $(f(x_n))_{n=1}^\infty$.

Proof (i) $\Rightarrow$ (ii). By (i), $\forall O \in \mathcal{O}$ with $y_0 \in O$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\forall x \in (D \cap \mathcal{B}_X(x_0, \delta)) \setminus \{x_0\}$, $f(x) \in O$. $\forall (x_n)_{n=1}^\infty \subseteq D \setminus \{x_0\}$ with $\lim_{n \in \mathbb{N}} x_n = x_0$, $\exists N \in \mathbb{N}$ such that $\forall n \geq N$, $x_n \in \mathcal{B}_X(x_0, \delta)$. Then, $x_n \in (D \cap \mathcal{B}_X(x_0, \delta)) \setminus \{x_0\}$ and $f(x_n) \in O$. Hence, y_0 is a limit point of $(f(x_n))_{n=1}^\infty$.

(ii) $\Rightarrow$ (i). We will show this by an argument of contradiction. Suppose (i) does not hold. Then, $\exists O_0 \in \mathcal{O}$ with $y_0 \in O_0$, $\forall n \in \mathbb{N}$, $\exists x_n \in (D \cap \mathcal{B}_X(x_0, 1/n)) \setminus \{x_0\}$ such that $f(x_n) \in \mathcal{X} \setminus O_0$. Clearly, the sequence $(x_n)_{n=1}^\infty \subseteq D \setminus \{x_0\}$ and $\lim_{n \in \mathbb{N}} x_n = x_0$. But, $f(x_n) \notin O_0$, $\forall n \in \mathbb{N}$. Then, y_0 is not a limit point of $(f(x_n))_{n=1}^\infty$. This contradicts (ii). Hence, (i) must hold.

This completes the proof of the proposition. $\square$

Proposition 4.17 Let $\mathcal{X}$ be a metric space, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathbb{R}$, and $x_0 \in \mathcal{X}$ be an accumulation point of D . Then, we have

$$\begin{aligned} \limsup_{x \rightarrow x_0} f(x) &= \inf_{\epsilon \in (0, \infty) \subset \mathbb{R}} \sup_{x \in (D \cap \mathcal{B}(x_0, \epsilon)) \setminus \{x_0\}} f(x) \\ \liminf_{x \rightarrow x_0} f(x) &= \sup_{\epsilon \in (0, \infty) \subset \mathbb{R}} \inf_{x \in (D \cap \mathcal{B}(x_0, \epsilon)) \setminus \{x_0\}} f(x) \end{aligned}$$

Proof Let $L = \limsup_{x \rightarrow x_0} f(x) \in \mathbb{R}_e$ and $\bar{L} := \inf_{\epsilon \in (0, \infty) \subset \mathbb{R}} \sup_{x \in (D \cap \mathcal{B}(x_0, \epsilon)) \setminus \{x_0\}} f(x) \in \mathbb{R}_e$. $\forall m \in \mathbb{R}$ with $m < L$, we have $m < \sup_{x \in (D \cap V) \setminus \{x_0\}} f(x)$, $\forall V \in \mathcal{O}_X$ with $x_0 \in V$. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\sup_{x \in (D \cap \mathcal{B}(x_0, \epsilon)) \setminus \{x_0\}} f(x) > m$. Hence, $m \leq \bar{L}$. By the arbitrariness of m , we have $L \leq \bar{L}$. $\forall m \in \mathbb{R}$ with $m > L$, $\exists V \in \mathcal{O}_X$ with $x_0 \in V$ such that $\sup_{x \in (D \cap V) \setminus \{x_0\}} f(x) < m$. Then, $\exists \epsilon \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(x_0, \epsilon) \subseteq V$. Then, $\sup_{x \in (D \cap \mathcal{B}(x_0, \epsilon)) \setminus \{x_0\}} f(x) \leq \sup_{x \in (D \cap V) \setminus \{x_0\}} f(x) < m$. Then, we have $\bar{L} < m$. By the arbitrariness of m , we have $\bar{L} \leq L$. Hence, $L = \bar{L}$.
Note that, by Propositions 3.85 and 3.81,

$$\begin{aligned} \liminf_{x \rightarrow x_0} f(x) &= -\limsup_{x \rightarrow x_0} (-f)(x) = -\inf_{\epsilon \in (0, \infty) \subset \mathbb{R}} \sup_{x \in (D \cap \mathcal{B}(x_0, \epsilon)) \setminus \{x_0\}} (-f)(x) \\ &= \sup_{\epsilon \in (0, \infty) \subset \mathbb{R}} \inf_{x \in (D \cap \mathcal{B}(x_0, \epsilon)) \setminus \{x_0\}} f(x) \end{aligned}$$

This completes the proof of the proposition. $\square$

4.3 Uniform Continuity and Uniformity

Definition 4.18 Let X and Y be metric spaces and $f : X \rightarrow Y$. f is said to be uniformly continuous if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\rho_Y(f(x_1), f(x_2)) < \epsilon$, $\forall x_1, x_2 \in X$ with $\rho_X(x_1, x_2) < \delta$.

Clearly, a function f is uniformly continuous implies that it is continuous.

Definition 4.19 Let X and Y be metric spaces and $f : X \rightarrow Y$. f is said to be a uniform homeomorphism if f is bijective and both f and f_{inv} are uniformly continuous.

Properties preserved under uniform homeomorphisms are called uniform properties. These includes Cauchy sequences, completeness, uniform continuity, and total boundedness.

Definition 4.20 Let X be a set and ρ_1 and ρ_2 be two metrics defined on X . Then, the two metrics are said to be uniformly equivalent if the identity map from (X, ρ_1) to (X, ρ_2) is a uniform homeomorphism.

Proposition 4.21 Let (X, ρ) be a metric space. Define $\sigma : X \times X \rightarrow \mathbb{R}$ by $\sigma(x_1, x_2) = \frac{\rho(x_1, x_2)}{1 + \rho(x_1, x_2)}$, $\forall x_1, x_2 \in X$. Then, σ is a metric on X and ρ and σ are uniformly equivalent.

Proof $\forall x_1, x_2, x_3 \in X$, $\sigma(x_1, x_2) \in [0, 1)$; $\sigma(x_1, x_2) = 0 \Leftrightarrow \rho(x_1, x_2) = 0 \Leftrightarrow x_1 = x_2$; $\sigma(x_1, x_2) = \sigma(x_2, x_1)$; by the monotonicity of $\frac{s}{1+s}$ on $s > -1$,

we have

$$\begin{aligned}\sigma(x_1, x_2) &\leq \frac{\rho(x_1, x_3) + \rho(x_3, x_2)}{1 + \rho(x_1, x_3) + \rho(x_3, x_2)} \leq \frac{\rho(x_1, x_3)}{1 + \rho(x_1, x_3)} + \frac{\rho(x_3, x_2)}{1 + \rho(x_3, x_2)} \\ &= \sigma(x_1, x_3) + \sigma(x_3, x_2)\end{aligned}$$

Hence, σ defines a metric on X .

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x_1, x_2 \in X$ with $\rho(x_1, x_2) < \epsilon$, we have $\sigma(x_1, x_2) \leq \rho(x_1, x_2) < \epsilon$. Hence, $\text{id}_X : (X, \rho) \rightarrow (X, \sigma)$ is uniformly continuous.

On the other hand, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x_1, x_2 \in X$ with $\sigma(x_1, x_2) < \frac{\epsilon}{1+\epsilon}$, we have $\rho(x_1, x_2) < \epsilon$. Hence, $\text{id}_X : (X, \sigma) \rightarrow (X, \rho)$ is uniformly continuous.

Therefore, ρ and σ are uniformly equivalent. $\square$

Definition 4.22 A metric space $\mathcal{X}$ is said to be totally bounded if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exist finitely many open balls with radius ϵ that cover $\mathcal{X}$.

Proposition 4.23 Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{W}$, and $\mathcal{Z}$ be metric spaces, $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$ be a Cauchy sequence, $f : \mathcal{X} \rightarrow \mathcal{Y}$ and $g : \mathcal{Y} \rightarrow \mathcal{Z}$ be uniformly continuous functions, and $h : \mathcal{X} \rightarrow \mathcal{W}$ be a uniform homeomorphism. Then, the following statements hold.

- (i) $(f(x_n))_{n=1}^\infty$ is a Cauchy sequence.
- (ii) $g \circ f$ is uniformly continuous.
- (iii) If $\mathcal{X}$ is complete, then $\mathcal{W}$ is complete.
- (iv) If $\mathcal{X}$ is totally bounded and f is surjective, then $\mathcal{Y}$ is totally bounded.

Proof (i) $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the uniform continuity of f , $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\rho_{\mathcal{Y}}(f(x_a), f(x_b)) < \epsilon$, $\forall x_a, x_b \in \mathcal{X}$ with $\rho_{\mathcal{X}}(x_a, x_b) < \delta$. Since $(x_i)_{i=1}^\infty$ is Cauchy, then $\exists N \in \mathbb{N}$ such that $\rho_{\mathcal{X}}(x_n, x_m) < \delta$, $\forall n, m \geq N$. Then, $\rho_{\mathcal{Y}}(f(x_n), f(x_m)) < \epsilon$. Hence, $(f(x_i))_{i=1}^\infty$ is a Cauchy sequence.

(ii) $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the uniform continuity of g , $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $\rho_{\mathcal{Z}}(g(y_1), g(y_2)) < \epsilon$, $\forall y_1, y_2 \in \mathcal{Y}$ with $\rho_{\mathcal{Y}}(y_1, y_2) < \delta_1$. By the uniform continuity of f , $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\rho_{\mathcal{Y}}(f(x_a), f(x_b)) < \delta_1$, $\forall x_a, x_b \in \mathcal{X}$ with $\rho_{\mathcal{X}}(x_a, x_b) < \delta$. Then, we have $\rho_{\mathcal{Z}}(g(f(x_a)), g(f(x_b))) < \epsilon$. Hence, $g \circ f$ is uniformly continuous.

(iii) $\forall$ Cauchy sequence $(w_i)_{i=1}^\infty \subseteq \mathcal{W}$. By (i), $(h_{\text{inv}}(w_i))_{i=1}^\infty \subseteq \mathcal{X}$ is a Cauchy sequence. Since $\mathcal{X}$ is complete, then $\lim_{i \in \mathbb{N}} h_{\text{inv}}(w_i) = x_0 \in \mathcal{X}$. By Proposition 3.66, we have $\lim_{i \in \mathbb{N}} w_i = \lim_{i \in \mathbb{N}} h(h_{\text{inv}}(w_i)) = h(x_0) \in \mathcal{W}$. Hence, $\mathcal{W}$ is complete.

(iv) $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the uniform continuity of f , $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\rho_{\mathcal{Y}}(f(x_a), f(x_b)) < \epsilon$, $\forall x_a, x_b \in \mathcal{X}$ with $\rho_{\mathcal{X}}(x_a, x_b) < \delta$. By the total boundedness of $\mathcal{X}$, there exists a finite set $X_N \subseteq \mathcal{X}$, such that

$\bigcup_{x \in X_N} \mathcal{B}_X(x, \delta) = \mathcal{X}$. Then, by the surjectiveness of f and Proposition 2.5, we have $\bigcup_{x \in X_N} f(\mathcal{B}_X(x, \delta)) = \mathcal{Y}$. Note that $f(\mathcal{B}_X(x, \delta)) \subseteq \mathcal{B}_Y(f(x), \epsilon)$, $\forall x \in \mathcal{X}$. Then, we have $\bigcup_{x \in X_N} \mathcal{B}_Y(f(x), \epsilon) = \mathcal{Y}$. Hence, $\mathcal{Y}$ is totally bounded.

This completes the proof of the proposition. $\square$

Definition 4.24 Let X be a set and $\mathcal{Y} := (Y, \rho)$ be a metric space. Let $(f_\alpha)_{\alpha \in \mathcal{A}}$ be a net of functions of X to $\mathcal{Y}$. Then, the net is said to converge uniformly to a function $f : X \rightarrow \mathcal{Y}$ if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \alpha_0 \in \mathcal{A}$, $\forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$, we have $\rho(f_\alpha(x), f(x)) < \epsilon$, $\forall x \in X$.

Definition 4.25 Let X be a set and $\mathcal{Y} := (Y, \rho)$ be a metric space. Let $(f_n)_{n=1}^\infty$ be a sequence of functions of X to $\mathcal{Y}$. Then, the sequence is said to be a uniform Cauchy sequence if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$, $\forall n, m \in \mathbb{N}$ with $n, m \geq N$, we have $\rho(f_n(x), f_m(x)) < \epsilon$, $\forall x \in X$.

A uniformly convergent sequence $(f_n)_{n=1}^\infty$ is a uniform Cauchy sequence. A uniform Cauchy sequence in a complete metric space is uniformly convergent.

Proposition 4.26 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $\mathcal{Y} := (Y, \rho)$ be a metric space. Let $(f_n)_{n=1}^\infty$ be a uniformly convergent sequence of functions of $\mathcal{X}$ to $\mathcal{Y}$ whose limit is $f : \mathcal{X} \rightarrow \mathcal{Y}$. Assume that, $\forall n \in \mathbb{N}$, f_n is continuous at $x_0 \in \mathcal{X}$. Then, f is continuous at x_0 .

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$, we have $\rho(f_N(x), f(x)) < \epsilon/3$, $\forall x \in \mathcal{X}$. Since f_N is continuous at x_0 , then $\exists U \in \mathcal{O}$ with $x_0 \in U$, $\forall x \in U$, we have $\rho(f_N(x), f_N(x_0)) < \epsilon/3$. Then, $\forall x \in U$, we have

$$\begin{aligned} \rho(f(x), f(x_0)) &\leq \rho(f(x), f_N(x)) + \rho(f_N(x), f_N(x_0)) + \rho(f_N(x_0), f(x_0)) \\ &< \epsilon \end{aligned}$$

Therefore, f is continuous at x_0 .

This completes the proof of the proposition. $\square$

Definition 4.27 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $\mathcal{Y} := (Y, \rho)$ be a metric space. Let $\mathcal{F}$ be a family of continuous functions of $\mathcal{X}$ to $\mathcal{Y}$. $\mathcal{F}$ is said to be equicontinuous at $x_0 \in \mathcal{X}$ if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists U \in \mathcal{O}$ with $x_0 \in U$, $\forall x \in U$, we have $\rho(f(x), f(x_0)) < \epsilon$, $\forall f \in \mathcal{F}$. $\mathcal{F}$ is said to be equicontinuous if it is equicontinuous at x_0 , $\forall x_0 \in \mathcal{X}$.

4.4 Product Metric Spaces

Definition 4.28 Let (X, ρ_X) and (Y, ρ_Y) be metric spaces. Then, the product metric space $(X \times Y, \rho) := (X, \rho_X) \times (Y, \rho_Y)$ is defined by

$$\rho((x_1, y_1), (x_2, y_2)) := ((\rho_X(x_1, x_2))^2 + (\rho_Y(y_1, y_2))^2)^{1/2}$$

$\forall x_1, x_2 \in X$, $\forall y_1, y_2 \in Y$. ρ is called the Cartesian metric.

In the above definition, it is straightforward to show that ρ defines a metric on $X \times Y$.

Proposition 4.29 *Let $\mathcal{X} := (X, \rho_X)$ be a metric space with natural topology $\mathcal{O}_X$. Let $\mathcal{Y} := (Y, \rho_Y)$ be a metric space with natural topology $\mathcal{O}_Y$. Let $\mathcal{X} \times \mathcal{Y} := (X \times Y, \rho)$ be the product metric space with the natural topology $\mathcal{O}_m$. Let $(X \times Y, \mathcal{O})$ be the product topological space $(X, \mathcal{O}_X) \times (Y, \mathcal{O}_Y)$. Then, $\mathcal{O} = \mathcal{O}_m$.*

Proof A basis for $\mathcal{O}$ is $\mathcal{M} := \{O_X \times O_Y \mid O_X \in \mathcal{O}_X, O_Y \in \mathcal{O}_Y\}$. A basis for $\mathcal{O}_m$ is $\mathcal{M}_m := \{\mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x, y), r) \mid x \in X, y \in Y, r \in (0, \infty) \subset \mathbb{R}\}$.

$\forall \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), r) \in \mathcal{M}_m$. $\forall (x, y) \in \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), r)$. Let $d := r - \rho((x, y), (x_0, y_0)) \in (0, \infty) \subset \mathbb{R}$. $\forall (\bar{x}, \bar{y}) \in \mathcal{B}_{\mathcal{X}}(x, d/\sqrt{2}) \times \mathcal{B}_{\mathcal{Y}}(y, d/\sqrt{2}) \in \mathcal{M}$, we have $\rho((\bar{x}, \bar{y}), (x_0, y_0)) \leq \rho((\bar{x}, \bar{y}), (x, y)) + \rho((x, y), (x_0, y_0)) = ((\rho_X(\bar{x}, x))^2 + (\rho_Y(\bar{y}, y))^2)^{1/2} + r - d < r$. Then, $(\bar{x}, \bar{y}) \in \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), r)$. Then, $\mathcal{B}_{\mathcal{X}}(x, d/\sqrt{2}) \times \mathcal{B}_{\mathcal{Y}}(y, d/\sqrt{2}) \subseteq \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), r)$. This implies that $\mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), r) \in \mathcal{O}$. Then, we have $\mathcal{O}_m \subseteq \mathcal{O}$.

On the other hand, $\forall O_X \times O_Y \in \mathcal{M}$. $\forall (x, y) \in O_X \times O_Y$, we have $x \in O_X$ and $y \in O_Y$. Then, $\exists d_1 \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{X}}(x, d_1) \subseteq O_X$ and $\exists d_2 \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{Y}}(y, d_2) \subseteq O_Y$. Let $d = \min\{d_1, d_2\} \in (0, \infty) \subset \mathbb{R}$. $\forall (\bar{x}, \bar{y}) \in \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x, y), d) \in \mathcal{M}_m$, we have $\rho_X(\bar{x}, x) < d$ and $\rho_Y(\bar{y}, y) < d$, which implies that $\bar{x} \in \mathcal{B}_{\mathcal{X}}(x, d_1)$ and $\bar{y} \in \mathcal{B}_{\mathcal{Y}}(y, d_2)$. Then, $\mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x, y), d) \subseteq O_X \times O_Y$. Hence, we have $O_X \times O_Y \in \mathcal{O}_m$. Then, $\mathcal{O} \subseteq \mathcal{O}_m$.

Hence, $\mathcal{O} = \mathcal{O}_m$. This completes the proof of the proposition. $\square$

The above proposition shows that the natural topology induced by the product metric defined in Definition 4.28 is the product topology on $X \times Y$.

Proposition 4.30 *Let $\mathcal{X}$ be a metric space. Then the metric $\rho_{\mathcal{X}}$ is a uniformly continuous function on $\mathcal{X} \times \mathcal{X}$.*

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$. Let $\delta = \epsilon/\sqrt{2} \in (0, \infty) \subset \mathbb{R}$. $\forall (x_1, x_2), (y_1, y_2) \in \mathcal{X} \times \mathcal{X}$ with $\rho_{\mathcal{X} \times \mathcal{X}}((x_1, x_2), (y_1, y_2)) < \delta$. Then, we have $\delta > ((\rho_{\mathcal{X}}(x_1, y_1))^2 + (\rho_{\mathcal{X}}(x_2, y_2))^2)^{1/2} \geq (\rho_{\mathcal{X}}(x_1, y_1) + \rho_{\mathcal{X}}(x_2, y_2))/\sqrt{2}$. Note that

$$\begin{aligned} \rho_{\mathcal{X}}(x_1, x_2) &\leq \rho_{\mathcal{X}}(x_1, y_1) + \rho_{\mathcal{X}}(y_1, x_2) \\ \rho_{\mathcal{X}}(y_1, x_2) &\leq \rho_{\mathcal{X}}(x_1, x_2) + \rho_{\mathcal{X}}(x_1, y_1) \\ \rho_{\mathcal{X}}(y_1, x_2) &\leq \rho_{\mathcal{X}}(y_1, y_2) + \rho_{\mathcal{X}}(x_2, y_2) \\ \rho_{\mathcal{X}}(y_1, y_2) &\leq \rho_{\mathcal{X}}(x_2, y_1) + \rho_{\mathcal{X}}(y_2, x_2) \end{aligned}$$

This implies that

$$\begin{aligned} -\epsilon &= -\sqrt{2}\delta < -\rho_{\mathcal{X}}(x_1, y_1) - \rho_{\mathcal{X}}(x_2, y_2) \leq \rho_{\mathcal{X}}(x_1, x_2) - \rho_{\mathcal{X}}(x_2, y_1) \\ &\quad + \rho_{\mathcal{X}}(x_2, y_1) - \rho_{\mathcal{X}}(y_1, y_2) = \rho_{\mathcal{X}}(x_1, x_2) - \rho_{\mathcal{X}}(y_1, y_2) \leq \rho_{\mathcal{X}}(x_1, y_1) \end{aligned}$$

$$+\rho_{\mathcal{X}}(x_2, y_1) + \rho_{\mathcal{X}}(x_2, y_2) - \rho_{\mathcal{X}}(x_2, y_1) = \rho_{\mathcal{X}}(x_1, y_1) + \rho_{\mathcal{X}}(x_2, y_2) < \epsilon$$

Hence, we have $|\rho_{\mathcal{X}}(x_1, x_2) - \rho_{\mathcal{X}}(y_1, y_2)| < \epsilon$. Hence, $\rho_{\mathcal{X}}$ is uniformly continuous on $\mathcal{X} \times \mathcal{X}$. This completes the proof of the proposition. $\square$

Proposition 4.31 *Let $\mathcal{X}$ and $\mathcal{Y}$ be complete metric spaces and $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$ be the product metric space with the cartesian metric ρ . Then, $\mathcal{Z}$ is complete.*

Proof Fix any Cauchy sequence $((x_n, y_n))_{n=1}^{\infty} \subseteq \mathcal{Z}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $\rho((x_n, y_n), (x_m, y_m)) < \epsilon$, $\forall n, m \geq N$. Then, we have $\rho_{\mathcal{X}}(x_n, x_m) < \epsilon$ and $\rho_{\mathcal{Y}}(y_n, y_m) < \epsilon$. Hence, $(x_n)_{n=1}^{\infty} \subseteq \mathcal{X}$ and $(y_n)_{n=1}^{\infty} \subseteq \mathcal{Y}$ are Cauchy sequences. By the completeness of $\mathcal{X}$ and $\mathcal{Y}$, $\exists x_0 \in \mathcal{X}$ and $\exists y_0 \in \mathcal{Y}$ such that $\lim_{n \in \mathbb{N}} x_n = x_0$ and $\lim_{n \in \mathbb{N}} y_n = y_0$. By Proposition 3.67, we have $\lim_{n \in \mathbb{N}} (x_n, y_n) = (x_0, y_0) \in \mathcal{Z}$. Hence, $\mathcal{Z}$ is complete. This completes the proof of the proposition. $\square$

Clearly, Definition 4.28 and Propositions 4.29 and 4.31 may be easily generalized to the case of $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$, where $n \in \mathbb{N}$ and $\mathcal{X}_i$ are metric spaces, $i = 1, \dots, n$. When $n = 0$, it should be noted that $\prod_{\alpha \in \emptyset} \mathcal{X}_{\alpha} = (\{\emptyset\}, \rho)$, where $\rho(\emptyset, \emptyset) = 0$.

Proposition 4.32 *Let $\mathcal{X}_{\alpha}$ be a metric space, $\alpha \in \Lambda$, where Λ is a finite set. Let $\Lambda = \bigcup_{\beta \in \Gamma} \Lambda_{\beta}$, where Λ_{β} 's are pairwise disjoint and finite and Γ is also finite. $\forall \beta \in \Gamma$, let $\mathcal{X}^{(\beta)} := \prod_{\alpha \in \Lambda_{\beta}} \mathcal{X}_{\alpha}$ be the product metric space. Let $\mathcal{X}^{(\Gamma)} := \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} \mathcal{X}_{\alpha}$ be the product metric space of product metric spaces, and $\mathcal{X} := \prod_{\alpha \in \Lambda} \mathcal{X}_{\alpha}$ be the product metric space. Then, $\mathcal{X}$ and $\mathcal{X}^{(\Gamma)}$ are isometric.*

Proof Define $E : \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} \mathcal{X}_{\alpha} \rightarrow \prod_{\alpha \in \Lambda} \mathcal{X}_{\alpha}$ by, $\forall x \in \mathcal{X}^{(\Gamma)}$, $\forall \alpha \in \Lambda$, $\exists! \beta_{\alpha} \in \Gamma \ni \alpha \in \Lambda_{\beta_{\alpha}}$, $\pi_{\alpha}(E(x)) = \pi_{\alpha}^{(\beta_{\alpha})}(\pi_{\beta_{\alpha}}^{(\Gamma)}(x))$. By Proposition 3.30, E is a homeomorphism. $\forall x, y \in \mathcal{X}^{(\Gamma)}$, we have

$$\begin{aligned} \rho(E(x), E(y)) &= \left(\sum_{\alpha \in \Lambda} (\rho_{\alpha}(\pi_{\alpha}(E(x)), \pi_{\alpha}(E(y))))^2 \right)^{1/2} \\ &= \left(\sum_{\beta \in \Gamma} \sum_{\alpha \in \Lambda_{\beta}} (\rho_{\alpha}(\pi_{\alpha}(E(x)), \pi_{\alpha}(E(y))))^2 \right)^{1/2} \\ &= \left(\sum_{\beta \in \Gamma} \sum_{\alpha \in \Lambda_{\beta}} (\rho_{\alpha}(\pi_{\alpha}^{(\beta)}(\pi_{\beta}^{(\Gamma)}(x)), \pi_{\alpha}^{(\beta)}(\pi_{\beta}^{(\Gamma)}(y))))^2 \right)^{1/2} \\ &= \left(\sum_{\beta \in \Gamma} \left(\left(\sum_{\alpha \in \Lambda_{\beta}} (\rho_{\alpha}(\pi_{\alpha}^{(\beta)}(\pi_{\beta}^{(\Gamma)}(x)), \pi_{\alpha}^{(\beta)}(\pi_{\beta}^{(\Gamma)}(y))))^2 \right)^{1/2} \right)^2 \right)^{1/2} \\ &= \left(\sum_{\beta \in \Gamma} (\rho^{(\beta)}(\pi_{\beta}^{(\Gamma)}(x), \pi_{\beta}^{(\Gamma)}(y)))^2 \right)^{1/2} = \rho^{(\Gamma)}(x, y) \end{aligned}$$

Hence, E is an isometry. This completes the proof of the proposition. $\square$

Proposition 4.33 *Let $\mathcal{X}_\alpha := (X_\alpha, \rho_{X_\alpha})$ and $\mathcal{Y}_\alpha := (Y_\alpha, \rho_{Y_\alpha})$ be uniformly homeomorphic metric spaces, $\forall \alpha \in \Lambda$, where Λ is a finite index set. Define the product metric spaces $\mathcal{X} := (X, \rho_X) := \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha$ and $\mathcal{Y} := (Y, \rho_Y) := \prod_{\alpha \in \Lambda} \mathcal{Y}_\alpha$, where ρ_X and ρ_Y are the Cartesian metric. Then, $\mathcal{X}$ and $\mathcal{Y}$ are uniformly homeomorphic.*

Proof Let $F_\alpha : \mathcal{X}_\alpha \rightarrow \mathcal{Y}_\alpha$ be a uniform homeomorphism, $\forall \alpha \in \Lambda$. Define $F : \mathcal{X} \rightarrow \mathcal{Y}$ by, $\forall x \in \mathcal{X}$, $\pi_\alpha^{\langle Y \rangle}(F(x)) = F_\alpha(\pi_\alpha^{\langle X \rangle}(x))$, $\forall \alpha \in \Lambda$. By Propositions 4.29 and 3.31, F is a homeomorphism between $\mathcal{X}$ and $\mathcal{Y}$. We need only to show that F and F_{inv} are uniformly continuous.

Let $m \in \mathbb{Z}_+$ be the number of elements in Λ . $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall \alpha \in \Lambda$, $\exists \delta_\alpha \in (0, \infty) \subset \mathbb{R}$ such that, $\forall x_{\alpha 1}, x_{\alpha 2} \in \mathcal{X}_\alpha$ with $\rho_{X_\alpha}(x_{\alpha 1}, x_{\alpha 2}) < \delta_\alpha$, we have $\rho_{Y_\alpha}(F_\alpha(x_{\alpha 1}), F_\alpha(x_{\alpha 2})) < \epsilon/\sqrt{1+m}$, by the uniform continuity of F_α . Let $\delta = \min\{\inf_{\alpha \in \Lambda} \delta_\alpha, 1\} \in (0, \infty) \subset \mathbb{R}$. $\forall x_1, x_2 \in \mathcal{X}$ with $\rho_X(x_1, x_2) < \delta$, we have, $\forall \alpha \in \Lambda$,

$$\rho_{X_\alpha}(\pi_\alpha^{\langle X \rangle}(x_1), \pi_\alpha^{\langle X \rangle}(x_2)) < \delta \leq \delta_\alpha$$

This implies that $\rho_{Y_\alpha}(F_\alpha(\pi_\alpha^{\langle X \rangle}(x_1)), F_\alpha(\pi_\alpha^{\langle X \rangle}(x_2))) < \epsilon/\sqrt{1+m}$. Hence, we have

$$\begin{aligned} \rho_Y(F(x_1), F(x_2)) &= \left(\sum_{\alpha \in \Lambda} (\rho_{Y_\alpha}(\pi_\alpha^{\langle Y \rangle}(F(x_1)), \pi_\alpha^{\langle Y \rangle}(F(x_2))))^2 \right)^{1/2} \\ &= \left(\sum_{\alpha \in \Lambda} (\rho_{Y_\alpha}(F_\alpha(\pi_\alpha^{\langle X \rangle}(x_1)), F_\alpha(\pi_\alpha^{\langle X \rangle}(x_2))))^2 \right)^{1/2} < \epsilon \end{aligned}$$

Hence, F is uniformly continuous.

It is easy to see that $F_{\text{inv}} : \mathcal{Y} \rightarrow \mathcal{X}$ is given by, $\forall y \in \mathcal{Y}$, $\pi_\alpha^{\langle X \rangle}(F_{\text{inv}}(y)) = F_{\alpha \text{inv}}(\pi_\alpha^{\langle Y \rangle}(y))$, $\forall \alpha \in \Lambda$. Then, it is straightforward to apply a similar argument as above to show that F_{inv} is uniformly continuous.

This completes the proof of the proposition. $\square$

Proposition 4.34 *Let $\mathcal{X}_i := (X_i, \rho_i)$ be a metric space with natural topology $\mathcal{O}_i$, $\forall i \in \mathbb{N}$. On $X := \prod_{i=1}^\infty X_i$ define the metric, $\forall x, y \in X$,*

$$\rho(x, y) = \sum_{i=1}^\infty 2^{-i} \frac{\rho_i(\pi_i(x), \pi_i(y))}{1 + \rho_i(\pi_i(x), \pi_i(y))}$$

Then, the natural topology, $\mathcal{O}_m$, on $\mathcal{X} := (X, \rho)$ induced by the metric ρ equals to the product topology $\mathcal{O}$, where $(X, \mathcal{O}) = \prod_{i=1}^\infty (X_i, \mathcal{O}_i)$.

Proof We will first show that ρ defines a metric on X . $\forall x, y, z \in X$. Clearly, we have $0 \leq \rho(x, y) < 1$ and $\rho(x, y) = \rho(y, x)$. When $x = y$, we have $\rho(x, y) = 0$. If $\rho(x, y) = 0$, then $\rho_i(\pi_i(x), \pi_i(y)) = 0$, $\forall i \in \mathbb{N}$,

which implies that $\pi_i(x) = \pi_i(y)$, $\forall i \in \mathbb{N}$, and hence $x = y$. Note that the function $\frac{s}{1+s} = 1 - \frac{1}{1+s}$ is strictly increasing on $s > -1$. Then, we have

$$\begin{aligned} \rho(x, y) &\leq \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_i(\pi_i(x), \pi_i(z)) + \rho_i(\pi_i(z), \pi_i(y))}{1 + \rho_i(\pi_i(x), \pi_i(z)) + \rho_i(\pi_i(z), \pi_i(y))} \\ &\leq \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_i(\pi_i(x), \pi_i(z))}{1 + \rho_i(\pi_i(x), \pi_i(z))} + \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_i(\pi_i(z), \pi_i(y))}{1 + \rho_i(\pi_i(z), \pi_i(y))} \\ &= \rho(x, z) + \rho(z, y) \end{aligned}$$

Hence, ρ is a metric on X .

Fix any basis open set $B = \mathcal{B}_{\mathcal{X}}(x, r) \in \mathcal{O}_m$, where $x \in X$ and $r \in (0, \infty) \subset \mathbb{R}$. $\forall y \in B$. Let $\delta := r - \rho(x, y) > 0$ and $N \in \mathbb{N}$ be such that $2^{-N} < \delta/2$. Consider the set $C_y = \prod_{i=1}^{\infty} C_{yi} \subseteq X$ given by $C_{yi} = \mathcal{B}_{\mathcal{X}_i}(\pi_i(y), \delta/2)$, $i = 1, \dots, N$, and $C_{yi} = X_i$, $i = N+1, N+2, \dots$. Clearly, $y \in C_y \in \mathcal{O}$. $\forall z \in C_y$, we have $\rho(y, z) < \sum_{i=1}^N 2^{-i-1}\delta + \sum_{i=N+1}^{\infty} 2^{-i} < \delta/2 + 2^{-N} < \delta$. Then, $\rho(x, z) \leq \rho(x, y) + \rho(y, z) < r$. Hence, $z \in B$. Then, $C_y \subseteq B$. This shows that $B \in \mathcal{O}$. Hence, $\mathcal{O}_m \subseteq \mathcal{O}$.

Fix any basis open set $C \in \mathcal{O}$. Then, $C = \prod_{i=1}^{\infty} C_i$, where $C_i \in \mathcal{O}_i$, $i = 1, \dots, N$, and $C_i = X_i$, $i = N+1, N+2, \dots$, for some $N \in \mathbb{N}$. $\forall x \in C$, we have $\pi_i(x) \in C_i$, $\forall i \in \mathbb{N}$. Then, $\forall i = 1, \dots, N$, $\exists \delta_i \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{X}_i}(\pi_i(x), \delta_i) \subseteq C_i$. Let $\delta_x := \min_{1 \leq i \leq N} 2^{-i} \frac{\delta_i}{1+\delta_i} > 0$. Let $B = \mathcal{B}_{\mathcal{X}}(x, \delta_x) \in \mathcal{O}_m$. $\forall y \in B$, $\rho(x, y) < \delta_x$ implies that $2^{-i} \frac{\rho_i(\pi_i(x), \pi_i(y))}{1+\rho_i(\pi_i(x), \pi_i(y))} < 2^{-i} \frac{\delta_i}{1+\delta_i}$ and $\rho_i(\pi_i(x), \pi_i(y)) < \delta_i$, $i = 1, \dots, N$. Hence, $y \in C$. Then, we have $x \in B \subseteq C$. This shows that $C \in \mathcal{O}_m$. Hence, $\mathcal{O} \subseteq \mathcal{O}_m$.

Hence, $\mathcal{O} = \mathcal{O}_m$. This completes the proof of the proposition. $\square$

Proposition 4.35 *Let $\mathcal{X}_i := (X_i, \rho_i)$ be a complete metric space, $\forall i \in \mathbb{N}$. On $X := \prod_{i=1}^{\infty} X_i$ define the metric ρ as in Proposition 4.34. Then, the product metric space $\mathcal{X} := (X, \rho)$ is complete.*

Proof $\forall$ Cauchy sequence $(x_n)_{n=1}^{\infty} \subseteq \mathcal{X}$, clearly, $(\pi_i(x_n))_{n=1}^{\infty} \subseteq \mathcal{X}_i$ is a Cauchy sequence, $\forall i \in \mathbb{N}$. By the completeness of $\mathcal{X}_i$, $\exists y_i \in \mathcal{X}_i$ such that $\lim_{n \in \mathbb{N}} \pi_i(x_n) = y_i$, $\forall i \in \mathbb{N}$. Define $y \in \mathcal{X}$ by $\pi_i(y) = y_i$, $\forall i \in \mathbb{N}$. By Propositions 3.67 and 4.34, we have y is the limit of $(x_n)_{n=1}^{\infty}$. Hence, $\mathcal{X}$ is complete. This completes the proof of the proposition. $\square$

Proposition 4.36 *Let $\mathcal{X}_i := (X_i, \rho_{X_i})$ and $\mathcal{Y}_i := (Y_i, \rho_{Y_i})$ be uniformly homeomorphic metric spaces, $\forall i \in \mathbb{N}$. Define the infinite product metric spaces $\mathcal{X} := (X, \rho_X) := (\prod_{i=1}^{\infty} X_i, \rho_X)$ and $\mathcal{Y} := (Y, \rho_Y) := (\prod_{i=1}^{\infty} Y_i, \rho_Y)$, where ρ_X and ρ_Y are defined as in Proposition 4.34. Then, $\mathcal{X}$ and $\mathcal{Y}$ are uniformly homeomorphic.*

Proof Let $F_i : \mathcal{X}_i \rightarrow \mathcal{Y}_i$ be a uniform homeomorphism, $\forall i \in \mathbb{N}$. Define $F : \mathcal{X} \rightarrow \mathcal{Y}$ by, $\forall x \in \mathcal{X}$, $\pi_i^{\langle Y \rangle}(F(x)) = F_i(\pi_i^{\langle X \rangle}(x))$, $\forall i \in \mathbb{N}$. By

Propositions 4.34 and 3.31, F is a homeomorphism between $\mathcal{X}$ and $\mathcal{Y}$. We need only to show that F and F_{inv} are uniformly continuous.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $2^{-N} < \epsilon/2$. $\forall i = 1, \dots, N$, $\exists \delta_i \in (0, \infty) \subset \mathbb{R}$ such that, $\forall x_{i1}, x_{i2} \in X_i$ with $\rho_{X_i}(x_{i1}, x_{i2}) < \delta_i$, we have $\rho_{Y_i}(F_i(x_{i1}), F_i(x_{i2})) < \epsilon/2$, by the uniform continuity of F_i . Let $\delta = \min_{1 \leq i \leq N} 2^{-i} \frac{\delta_i}{1 + \delta_i} > 0$. $\forall x_1, x_2 \in \mathcal{X}$ with $\rho_X(x_1, x_2) < \delta$, we have, $\forall i = 1, \dots, N$,

$$\frac{\rho_{X_i}(\pi_i^{\langle X \rangle}(x_1), \pi_i^{\langle X \rangle}(x_2))}{1 + \rho_{X_i}(\pi_i^{\langle X \rangle}(x_1), \pi_i^{\langle X \rangle}(x_2))} < 2^i \delta \leq \frac{\delta_i}{1 + \delta_i}$$

This implies that $\rho_{X_i}(\pi_i^{\langle X \rangle}(x_1), \pi_i^{\langle X \rangle}(x_2)) < \delta_i$, which further implies that $\rho_{Y_i}(F_i(\pi_i^{\langle X \rangle}(x_1)), F_i(\pi_i^{\langle X \rangle}(x_2))) < \epsilon/2$. Hence, we have

$$\begin{aligned} \rho_Y(F(x_1), F(x_2)) &= \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_{Y_i}(\pi_i^{\langle Y \rangle}(F(x_1)), \pi_i^{\langle Y \rangle}(F(x_2)))}{1 + \rho_{Y_i}(\pi_i^{\langle Y \rangle}(F(x_1)), \pi_i^{\langle Y \rangle}(F(x_2)))} \\ &= \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_{Y_i}(F_i(\pi_i^{\langle X \rangle}(x_1)), F_i(\pi_i^{\langle X \rangle}(x_2)))}{1 + \rho_{Y_i}(F_i(\pi_i^{\langle X \rangle}(x_1)), F_i(\pi_i^{\langle X \rangle}(x_2)))} \\ &< \sum_{i=1}^N 2^{-i} \frac{\epsilon}{2} + 2^{-N} < \epsilon \end{aligned}$$

Hence, F is uniformly continuous.

It is easy to see that $F_{\text{inv}} : \mathcal{Y} \rightarrow \mathcal{X}$ is given by, $\forall y \in \mathcal{Y}$, $\pi_i^{\langle X \rangle}(F_{\text{inv}}(y)) = F_{\text{inv}}(\pi_i^{\langle Y \rangle}(y))$, $\forall i \in \mathbb{N}$. Then, it is straightforward to apply a similar argument as above to show that F_{inv} is uniformly continuous.

This completes the proof of the proposition. $\square$

4.5 Subspaces

Proposition 4.37 *Let $\mathcal{X} := (X, \rho)$ be a metric space with natural topology $\mathcal{O}$ and $A \subseteq \mathcal{X}$ be a subset. $\mathcal{A} := (A, \rho)$ is a metric space with natural topology $\mathcal{O}_A$. A may also be endowed with the subset topology $\mathcal{O}_s$ with respect to $(X, \mathcal{O})$. Then, $\mathcal{O}_A = \mathcal{O}_s$.*

Proof A basis for the natural topology $\mathcal{O}_A$ is

$$\begin{aligned} \mathcal{B} &:= \{ \mathcal{B}_A(x, r) \mid x \in A, r \in (0, \infty) \subseteq \mathbb{R} \} \\ &= \{ \mathcal{B}_X(x, r) \cap A \mid x \in A, r \in (0, \infty) \subseteq \mathbb{R} \} \end{aligned}$$

$\mathcal{O}_s = \{ O \cap A \mid O \in \mathcal{O} \}$. Clearly, we have $\mathcal{B} \subseteq \mathcal{O}_s$. Then, $\mathcal{O}_A \subseteq \mathcal{O}_s$. On the other hand, $\forall O_s \in \mathcal{O}_s$, $O_s = O \cap A$ with $O \in \mathcal{O}$. $\forall x \in O_s$, $x \in A$ and $\exists r \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_X(x, r) \subseteq O$. Then, $\mathcal{B}_A(x, r) \subseteq O_s$. Then, $O_s \in \mathcal{O}_A$. Then, $\mathcal{O}_s \subseteq \mathcal{O}_A$. Hence, $\mathcal{O}_A = \mathcal{O}_s$.

This completes the proof of the proposition. $\square$

Proposition 4.38 *Let $\mathcal{X}$ be a metric space and $S \subseteq \mathcal{X}$. If $\mathcal{X}$ is separable then $\mathcal{S} := (S, \rho)$ is separable. If S is separable, then $\overline{S}$ is separable.*

Proof Since $\mathcal{X}$ is separable, by Proposition 4.4, there exists a countable basis $\mathcal{B}$ for $\mathcal{X}$. Let $\mathcal{B}_S := \{B \cap S \mid B \in \mathcal{B}\}$, which is a countable basis for $\mathcal{S}$. By Proposition 4.4, $\mathcal{S}$ is separable.

Let S be separable. Then, there exists a countable dense subset $D \subseteq S$. The closure of D in $\mathcal{X}$ is $\overline{D}$. By Proposition 3.5, $S \subseteq \overline{D}$. By Proposition 3.3, $\overline{D} \subseteq \overline{S} \subseteq \overline{D}$. Hence, $\overline{D} = \overline{S}$ and D is dense in $\overline{S}$. Hence, $\overline{S}$ is separable. This completes the proof of the proposition. $\square$

Proposition 4.39 *Let $\mathcal{X} := (X, \rho)$ be a metric space and $S \subseteq X$.*

(i) *If (S, ρ) is complete then S is closed in $\mathcal{X}$.*

(ii) *If $\mathcal{X}$ is complete and S is closed in $\mathcal{X}$, then (S, ρ) is complete.*

Proof (i) Any convergent sequence is Cauchy. Then, by Proposition 4.13, (S, ρ) is complete implies that any point of closure of S is in S . Then, $S = \overline{S}$ and S is closed. (ii) Any Cauchy sequence in S will converge in $\mathcal{X}$ since $\mathcal{X}$ is complete. Then, by Proposition 4.13, the limit lies in S since S is closed. Hence, (S, ρ) is complete. This completes the proof of the proposition. $\square$

4.6 Baire Category

Theorem 4.40 (Baire) *Let $\mathcal{X}$ be a complete metric space and $(O_k)_{k=1}^\infty$ be a sequence of open dense sets in $\mathcal{X}$. Then, $\bigcap_{k=1}^\infty O_k$ is dense in $\mathcal{X}$.*

Proof $\forall U \in \mathcal{O}$ with $U \neq \emptyset$. Since O_1 is dense, then $\exists x_1 \in O_1 \cap U$. By the openness of U and O_1 , $\exists r_1 \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(x_1, r_1) \subseteq O_1 \cap U$. $\forall n \in \mathbb{N}$ with $n > 1$, since O_n is dense, then $\exists x_n \in O_n \cap \mathcal{B}(x_{n-1}, r_{n-1}/2)$. Since $O_n, \mathcal{B}(x_{n-1}, r_{n-1}) \in \mathcal{O}$, then $\exists r_n \in (0, r_{n-1}/2] \subseteq \mathbb{R}$ such that $\mathcal{B}(x_n, r_n) \subseteq O_n \cap \mathcal{B}(x_{n-1}, r_{n-1})$.

Note that $\forall n \in \mathbb{N}$, $r_n \leq 2^{1-n}r_1$ and, $\forall m > n$,

$$\begin{aligned} \rho(x_n, x_m) &\leq \rho(x_n, x_{n+1}) + \cdots + \rho(x_{m-1}, x_m) < \frac{1}{2}(r_n + \cdots + r_{m-1}) \\ &\leq 2^{-n}r_1 + \cdots + 2^{1-m}r_1 < 2^{1-n}r_1 \end{aligned}$$

Hence, $(x_n)_{n=1}^\infty$ is a Cauchy sequence, which converges to $x_0 \in \mathcal{X}$, by the completeness of $\mathcal{X}$. By Propositions 3.66, 3.67, and 4.30, we have, $\forall n \in \mathbb{N}$,

$$\rho(x_n, x_0) = \lim_{m \in \mathbb{N}} \rho(x_n, x_m) < \frac{1}{2}(r_n + 2^{-1}r_n + 2^{-2}r_n + \cdots) = r_n$$

Then, $x_0 \in \mathcal{B}(x_n, r_n) \subseteq O_n$. Hence, $x_0 \in (\bigcap_{k=1}^\infty O_k) \cap U \neq \emptyset$. Hence, $\bigcap_{k=1}^\infty O_k$ is dense in $\mathcal{X}$. This completes the proof of the theorem. $\square$

Theorem 4.41 (Baire Category Theorem) *Let $\mathcal{X}$ be a complete metric space, then it is second category everywhere, that is no nonempty open subset of $\mathcal{X}$ is of first category.*

Proof We will prove this theorem by using Proposition 3.38. Let $(O_\alpha)_{\alpha \in \Lambda}$ be a countable collection of open dense sets in $\mathcal{X}$. When Λ is finite, we may add the open dense set $\mathcal{X}$ into the collection to make it infinite and countable. By Baire's Theorem, we then have $\bigcap_{\alpha \in \Lambda} O_\alpha$ is dense in $\mathcal{X}$. Then, by Proposition 3.38, $\mathcal{X}$ is second category everywhere. This completes the proof of the theorem. $\square$

4.7 Completion of Metric Spaces

Proposition 4.42 *Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $\mathcal{Y} := (Y, \rho)$ be a complete metric space. Let $(f_n)_{n=1}^\infty$ be a uniform Cauchy sequence of continuous functions of $\mathcal{X}$ to $\mathcal{Y}$. Then, there exists a continuous function $f : \mathcal{X} \rightarrow \mathcal{Y}$ such that $(f_n)_{n=1}^\infty$ converges uniformly to f .*

Proof $\forall x \in \mathcal{X}$, $(f_n(x))_{n=1}^\infty$ is a Cauchy sequence in $\mathcal{Y}$. By the completeness of $\mathcal{Y}$ and Proposition 3.65, $\exists! y \in \mathcal{Y}$ such that $\lim_{n \in \mathbb{N}} f_n(x) = y$. This defines a function $f : \mathcal{X} \rightarrow \mathcal{Y}$.

Since $(f_n)_{n=1}^\infty$ is a uniform Cauchy sequence, then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$, $\forall n, m \in \mathbb{N}$ with $n, m \geq N$, we have $\rho(f_n(x), f_m(x)) < \epsilon$, $\forall x \in \mathcal{X}$. By Propositions 4.30, 3.66, and 3.67, we have $\rho(f_n(x), f(x)) = \lim_{m \in \mathbb{N}} \rho(f_n(x), f_m(x)) \leq \epsilon$, $\forall x \in \mathcal{X}$. Hence, $(f_n)_{n=1}^\infty$ converges uniformly to f . By Proposition 4.26, f is continuous.

This completes the proof of the proposition. $\square$

Definition 4.43 *Let $\mathcal{X}$ be a metric space. A net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathcal{X}$ is said to be Cauchy if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \alpha_0 \in \mathcal{A}$, $\forall \alpha_1, \alpha_2 \in \mathcal{A}$ with $\alpha_0 \prec \alpha_1$ and $\alpha_0 \prec \alpha_2$, we have $\rho(x_{\alpha_1}, x_{\alpha_2}) < \epsilon$.*

Any convergent net in a metric space is Cauchy.

Proposition 4.44 *Let $\mathcal{X}$ be a metric space. Then, $\mathcal{X}$ is complete if, and only if, every Cauchy net has a limit in $\mathcal{X}$.*

Proof “If” A Cauchy sequence is a Cauchy net and admits a limit in $\mathcal{X}$. Hence, $\mathcal{X}$ is complete.

“Only if” Let $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathcal{X}$ be a Cauchy net. $\exists \alpha_1 \in \mathcal{A} \ni \forall \bar{\alpha}_1, \bar{\alpha}_2 \in \mathcal{A}$ with $\alpha_1 \prec \bar{\alpha}_1$ and $\alpha_1 \prec \bar{\alpha}_2$, we have $\rho(x_{\bar{\alpha}_1}, x_{\bar{\alpha}_2}) < 1$. $\forall n \in \mathbb{N}$ with $n > 1$, $\exists \alpha_n \in \mathcal{A}$ with $\alpha_{n-1} \prec \alpha_n$, $\forall \bar{\alpha}_1, \bar{\alpha}_2 \in \mathcal{A}$ with $\alpha_n \prec \bar{\alpha}_1$ and $\alpha_n \prec \bar{\alpha}_2$, we have $\rho(x_{\bar{\alpha}_1}, x_{\bar{\alpha}_2}) < 1/n$. $\forall n \in \mathbb{N}$, $\forall m_1, m_2 \in \mathbb{N}$ with $m_1, m_2 > n$, we have $\alpha_n \prec \alpha_{n+1} \prec \dots \prec \alpha_{m_1}$ and $\alpha_n \prec \alpha_{n+1} \prec \dots \prec \alpha_{m_2}$. Then, $\rho(x_{\alpha_{m_1}}, x_{\alpha_{m_2}}) < 1/n$. Hence, $(x_{\alpha_n})_{n=1}^\infty \subseteq \mathcal{X}$ is a Cauchy sequence. By the completeness of $\mathcal{X}$, $\exists x_0 \in \mathcal{X}$ such that $\lim_{n \in \mathbb{N}} x_{\alpha_n} = x_0$. $\forall n \in \mathbb{N}$, $\forall \alpha \in \mathcal{A}$

with $\alpha_n \prec \alpha$, $\forall m \in \mathbb{N}$ with $m > n$, we have $\alpha_n \prec \alpha_{n+1} \prec \cdots \prec \alpha_m$ and $\rho(x_\alpha, x_{\alpha_m}) < 1/n$. Then, by Propositions 4.30, 3.66, and 3.67, we have $\rho(x_\alpha, x_0) = \lim_{m \in \mathbb{N}} \rho(x_\alpha, x_{\alpha_m}) \leq 1/n$. Hence, we have $\lim_{\alpha \in \mathcal{A}} x_\alpha = x_0$.

This completes the proof of the proposition. $\square$

Proposition 4.45 *Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $\mathcal{Y} := (Y, \rho)$ be a complete metric space, $D \subseteq X$, $x_0 \in X$ be an accumulation point of D , and $f : D \rightarrow \mathcal{Y}$. Then, $\lim_{x \rightarrow x_0} f(x) \in \mathcal{Y}$ if, and only if, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists O \in \mathcal{O}$ with $x_0 \in O$, $\forall \bar{x}, \hat{x} \in (D \cap O) \setminus \{x_0\}$, we have $\rho(f(\bar{x}), f(\hat{x})) < \epsilon$.*

Proof “Sufficiency” Assume that $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists O \in \mathcal{O}$ with $x_0 \in O$, $\forall \bar{x}, \hat{x} \in (D \cap O) \setminus \{x_0\}$, we have $\rho(f(\bar{x}), f(\hat{x})) < \epsilon$. Define $\mathcal{M} := \{O \in \mathcal{O} \mid x_0 \in O\}$. Clearly, $X \in \mathcal{M}$ and $\mathcal{M} \neq \emptyset$. It is easy to see that $\mathcal{A} := (\mathcal{M}, \supseteq)$ is a directed system. Since x_0 is an accumulation point of D , then $\forall O \in \mathcal{A}$, $(D \cap O) \setminus \{x_0\} \neq \emptyset$. By Axiom of Choice, $\exists$ a net $(x_O)_{O \in \mathcal{A}} \subseteq X$ such that $x_O \in (D \cap O) \setminus \{x_0\}$, $\forall O \in \mathcal{A}$. This also defines a net $(f(x_O))_{O \in \mathcal{A}} \subseteq Y$ by Axiom of Replacement. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the assumption, $\exists \hat{O} \in \mathcal{A}$, $\forall \bar{x}, \hat{x} \in (D \cap \hat{O}) \setminus \{x_0\}$, we have $\rho(f(\bar{x}), f(\hat{x})) < \epsilon$. $\forall O_1, O_2 \in \mathcal{A}$ with $\hat{O} \supseteq O_1$ and $\hat{O} \supseteq O_2$, $x_{O_i} \in (D \cap O_i) \setminus \{x_0\} \subseteq (D \cap \hat{O}) \setminus \{x_0\}$, $i = 1, 2$. This implies that $\rho(f(x_{O_1}), f(x_{O_2})) < \epsilon$. This shows that the net $(f(x_O))_{O \in \mathcal{A}} \subseteq Y$ is Cauchy. By Proposition 4.44, $\lim_{O \in \mathcal{A}} f(x_O) = y_0 \in Y$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists O_1 \in \mathcal{A}$, $\forall O \in \mathcal{A}$ with $O_1 \supseteq O$, we have $\rho(y_0, f(x_O)) < \epsilon/2$. By the assumption, $\exists O_2 \in \mathcal{A}$, $\forall \bar{x}, \hat{x} \in (D \cap O_2) \setminus \{x_0\}$, we have $\rho(f(\bar{x}), f(\hat{x})) < \epsilon/2$. Let $O_3 := O_1 \cap O_2 \in \mathcal{A}$. Then, $O_1 \supseteq O_3$ and $x_{O_3} \in (D \cap O_3) \setminus \{x_0\} \subseteq (D \cap O_2) \setminus \{x_0\}$. $\forall x \in (D \cap O_3) \setminus \{x_0\}$, we have $\rho(y_0, f(x)) \leq \rho(y_0, f(x_{O_3})) + \rho(f(x_{O_3}), f(x)) < \epsilon$. Hence, we have $\lim_{x \rightarrow x_0} f(x) = y_0 \in \mathcal{Y}$.

“Necessity” Let $\lim_{x \rightarrow x_0} f(x) = y_0 \in \mathcal{Y}$. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists O \in \mathcal{O}$ with $x_0 \in O$, $\forall x \in (D \cap O) \setminus \{x_0\}$, we have $\rho(f(x), y_0) < \epsilon/2$. Hence, $\forall \bar{x}, \hat{x} \in (D \cap O) \setminus \{x_0\}$, $\rho(f(\bar{x}), f(\hat{x})) \leq \rho(f(\bar{x}), y_0) + \rho(y_0, f(\hat{x})) < \epsilon$.

This completes the proof of the proposition. $\square$

Proposition 4.46 *Let (X, ρ_X) and (Y, ρ_Y) be metric spaces, (Y, ρ_Y) be complete, $E \subseteq X$, and $f : E \rightarrow Y$ be uniformly continuous. Then, there is a unique continuous extension $g : \bar{E} \rightarrow Y$. Furthermore, g is uniformly continuous.*

Proof $\forall (x_i)_{i=1}^\infty \subseteq E$ with $\lim_{i \in \mathbb{N}} x_i = x_0 \in X$, by Proposition 4.23 and the uniform continuity of f , $(f(x_i))_{i=1}^\infty$ is a Cauchy sequence in Y . Since (Y, ρ_Y) is complete, then $\exists y_0 \in Y$ such that $\lim_{i \in \mathbb{N}} f(x_i) = y_0$. Let $(\bar{x}_i)_{i=1}^\infty \subseteq E$ be any other sequence with $\lim_{i \in \mathbb{N}} \bar{x}_i = x_0$. Then, the sequence $(x_1, \bar{x}_1, x_2, \bar{x}_2, \dots)$ converges to x_0 . By Proposition 4.23 and the uniform continuity of f , $(f(x_1), f(\bar{x}_1), f(x_2), f(\bar{x}_2), \dots)$ is a Cauchy sequence in Y , which converges since (Y, ρ_Y) is complete. The limit for this Cauchy sequence must be y_0 by Proposition 3.70. Hence, we have $\lim_{i \in \mathbb{N}} f(\bar{x}_i) = y_0$. Hence, y_0 is dependent only on x_0 but not on the sequence $(x_i)_{i=1}^\infty$.

By Proposition 4.13, we may define a function $g : \overline{E} \rightarrow Y$ by $g(x_0) = y_0$, $\forall x_0 \in \overline{E}$. $\forall x_0 \in E$, choose a sequence $(x_0, x_0, \dots)$, which converges to x_0 , then $y_0 = f(x_0)$. Hence, we have $g|_E = f$.

Next, we show that g is uniformly continuous. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the uniform continuity of f on E , $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\rho_Y(f(x_1), f(x_2)) < \epsilon$, $\forall x_1, x_2 \in E$ with $\rho_X(x_1, x_2) < \delta$. $\forall x_1, x_2 \in \overline{E}$ with $\rho_X(x_1, x_2) < \delta$. By Proposition 4.13, $\exists \left(x_i^{(1)}\right)_{i=1}^{\infty}, \left(x_i^{(2)}\right)_{i=1}^{\infty} \subseteq E$ such that $\lim_{i \in \mathbb{N}} x_i^{(j)} = x_j$, $j = 1, 2$. By the definition of g , we have $\lim_{i \in \mathbb{N}} f(x_i^{(j)}) = g(x_j)$, $j = 1, 2$. By Proposition 3.67, we have $\lim_{i \in \mathbb{N}} (f(x_i^{(1)}), f(x_i^{(2)})) = (g(x_1), g(x_2))$. By Propositions 3.66 and 4.30, we have $\lim_{i \in \mathbb{N}} \rho_Y(f(x_i^{(1)}), f(x_i^{(2)})) = \rho_Y(g(x_1), g(x_2))$. Hence, we have $\rho_Y(g(x_1), g(x_2)) \leq \epsilon$. Hence, g is uniformly continuous.

Let $h : \overline{E} \rightarrow Y$ be any continuous mapping such that $h|_E = f$. $\forall \bar{x} \in \overline{E}$, by Proposition 4.13, $\exists (x_i)_{i=1}^{\infty} \subseteq E$ such that $\lim_{i \in \mathbb{N}} x_i = \bar{x}$. By continuity of h and g and Proposition 3.66, we have $h(\bar{x}) = \lim_{i \in \mathbb{N}} h(x_i) = \lim_{i \in \mathbb{N}} f(x_i) = \lim_{i \in \mathbb{N}} g(x_i) = g(\bar{x})$. Then, $h = g$. This shows that g is unique in the class of continuous functions that extends f .

This completes the proof of the proposition. $\square$

Definition 4.47 A pseudo-metric ρ on a set X satisfies (i), (iii), and (iv) in Definition 4.1 and $\rho(x, x) = 0$, $\forall x \in X$, but not necessarily (ii).

Lemma 4.48 Let X be a set and ρ be a pseudo-metric on X . Define an equivalence relation $\equiv$ on X by $x_1 \equiv x_2$ if $\rho(x_1, x_2) = 0$, $\forall x_1, x_2 \in X$. Let $\mathcal{Y}$ be the quotient set $X/\equiv$, that is $\mathcal{Y} := \{F \subseteq X \mid F \neq \emptyset, \forall x_1, x_2 \in F, x_1 \equiv x_2, \forall x_3 \in X \setminus F, x_1 \not\equiv x_3\}$. Define a mapping $\rho_Y : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}$ by $\rho_Y(Y_1, Y_2) = \rho(x_1, x_2)$, $\forall Y_1, Y_2 \in \mathcal{Y}$ and $x_1 \in Y_1$ and $x_2 \in Y_2$. Then, $(\mathcal{Y}, \rho_Y)$ is a metric space and said to be the quotient space of (X, ρ) modulo ρ .

Proof We first show that $\equiv$ is an equivalence relation. Clearly, $\equiv$ is reflexive and symmetric.

Claim 4.48.1 $\forall x, y, z \in X$, if $\rho(y, z) = 0$, then $\rho(x, y) = \rho(x, z)$.

Proof of claim: $\rho(x, z) \leq \rho(x, y) + \rho(y, z) = \rho(x, y) \leq \rho(x, z) + \rho(y, z) = \rho(x, z)$. This completes the proof of the claim. $\square$

By Claim 4.48.1, $\equiv$ is transitive. Hence, $\equiv$ is a equivalence relation.

Again by Claim 4.48.1, the mapping ρ_Y is well-defined.

To show that ρ_Y is a metric on $\mathcal{Y}$, we note that, $\forall Y_1, Y_2, Y_3 \in \mathcal{Y}$, (i) $0 \leq \rho_Y(Y_1, Y_2) < \infty$; (ii) $\rho_Y(Y_1, Y_2) = 0 \Leftrightarrow \forall x_1 \in Y_1, \forall x_2 \in Y_2, \rho(x_1, x_2) = 0$ and $x_1 \equiv x_2 \Leftrightarrow Y_1 = Y_2$; (iii) $\rho_Y(Y_1, Y_2) = \rho(x_1, x_2) = \rho(x_2, x_1) = \rho_Y(Y_2, Y_1)$, for some $x_1 \in Y_1$ and $x_2 \in Y_2$; (iv) $\rho_Y(Y_1, Y_2) = \rho(x_1, x_2) \leq \rho(x_1, x_3) + \rho(x_3, x_2) = \rho_Y(Y_1, Y_3) + \rho_Y(Y_3, Y_2)$, for some $x_1 \in Y_1, x_2 \in Y_2$, and $x_3 \in Y_3$. Hence, ρ_Y is a metric on $\mathcal{Y}$.

This completes the proof of the lemma. $\square$

Theorem 4.49 *Let (X, ρ) be an metric space. Then,*

- (i) *there exists a complete metric space $(\bar{X}, \bar{\rho})$ such that X is isometrically embedded as a dense subset in $\bar{X}$;*
- (ii) *let (Y, ρ_Y) be any complete metric space with $X \subseteq Y$ and $\rho_Y|_{X \times X} = \rho$. Then, $(\bar{X}, \bar{\rho})$ is isometric with $\bar{X}$ in Y .*

Proof (i) Let $\mathcal{X} := \{(x_n)_{n=1}^\infty \subseteq X \mid (x_n)_{n=1}^\infty \text{ is a Cauchy sequence}\}$. Define $\sigma : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ by $\sigma\left(\left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(2)}\right)_{n=1}^\infty\right) = \lim_{n \in \mathbb{N}} \rho(x_n^{(1)}, x_n^{(2)})$, $\forall \left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(2)}\right)_{n=1}^\infty \in \mathcal{X}$. The mapping σ is well defined by Propositions 4.30 and 4.23 and the fact that $\left(x_n^{(1)}, x_n^{(2)}\right)_{n=1}^\infty$ is a Cauchy sequence.

$\forall \left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(2)}\right)_{n=1}^\infty, \left(x_n^{(3)}\right)_{n=1}^\infty \in \mathcal{X}$,
 $\sigma\left(\left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(2)}\right)_{n=1}^\infty\right) \in [0, \infty) \subset \mathbb{R}$; $\sigma\left(\left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(2)}\right)_{n=1}^\infty\right) = \lim_{n \in \mathbb{N}} \rho(x_n^{(1)}, x_n^{(2)}) = \lim_{n \in \mathbb{N}} \rho(x_n^{(2)}, x_n^{(1)}) = \sigma\left(\left(x_n^{(2)}\right)_{n=1}^\infty, \left(x_n^{(1)}\right)_{n=1}^\infty\right)$;
 $\sigma\left(\left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(2)}\right)_{n=1}^\infty\right) = \lim_{n \in \mathbb{N}} \rho(x_n^{(1)}, x_n^{(2)}) \leq \lim_{n \in \mathbb{N}} (\rho(x_n^{(1)}, x_n^{(3)}) + \rho(x_n^{(3)}, x_n^{(2)})) = \sigma\left(\left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(3)}\right)_{n=1}^\infty\right) + \sigma\left(\left(x_n^{(3)}\right)_{n=1}^\infty, \left(x_n^{(2)}\right)_{n=1}^\infty\right)$;
 and also $0 = \lim_{n \in \mathbb{N}} \rho(x_n^{(1)}, x_n^{(1)}) = \sigma\left(\left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(1)}\right)_{n=1}^\infty\right)$. Hence, σ defines a pseudo-metric on $\mathcal{X}$.

By Lemma 4.48, we can define the equivalence relation $\equiv$ on $\mathcal{X}$ by $\left(x_n^{(1)}\right)_{n=1}^\infty \equiv \left(x_n^{(2)}\right)_{n=1}^\infty$ if $\lim_{n \in \mathbb{N}} \rho(x_n^{(1)}, x_n^{(2)}) = 0$, $\forall \left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(2)}\right)_{n=1}^\infty \in \mathcal{X}$. Then, define the set $\bar{X} := \mathcal{X} / \equiv := \left\{ F \subseteq \mathcal{X} \mid F \neq \emptyset, \forall \left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(2)}\right)_{n=1}^\infty \in F, \left(x_n^{(1)}\right)_{n=1}^\infty \equiv \left(x_n^{(2)}\right)_{n=1}^\infty, \forall \left(x_n^{(3)}\right)_{n=1}^\infty \in \mathcal{X} \setminus F, \left(x_n^{(1)}\right)_{n=1}^\infty \not\equiv \left(x_n^{(3)}\right)_{n=1}^\infty \right\}$. A metric on $\bar{X}$ is $\bar{\rho}$, defined by $\bar{\rho}(F_1, F_2) = \sigma\left(\left(x_n^{(1)}\right)_{n=1}^\infty, \left(x_n^{(2)}\right)_{n=1}^\infty\right)$, $\forall F_1, F_2 \in \bar{X}$, for some $\left(x_n^{(1)}\right)_{n=1}^\infty \in F_1$ and $\left(x_n^{(2)}\right)_{n=1}^\infty \in F_2$.

Define mapping $T : X \rightarrow \bar{X}$ by, $\forall x \in X$, $T(x) = F \in \bar{X}$ such that the Cauchy sequence $(x, x, \dots) \in F$. Let X_I be the image of X under T . Then, $T : X \rightarrow X_I$ is surjective. $\forall x_1, x_2 \in X$ with $T(x_1) = T(x_2)$, we have $(x_1, x_1, \dots) \equiv (x_2, x_2, \dots)$, which implies that $\sigma((x_1, x_1, \dots), (x_2, x_2, \dots)) = \rho(x_1, x_2) = 0$, and hence $x_1 = x_2$. Therefore, T is injective. Hence, $T : X \rightarrow X_I$ is bijective and admits an inverse $T_{\text{inv}} : X_I \rightarrow X$. $\forall x_1, x_2 \in X$, $\bar{\rho}(T(x_1), T(x_2)) = \sigma((x_1, x_1, \dots), (x_2, x_2, \dots)) = \rho(x_1, x_2)$. Hence, T is metric preserving. Then, both T and T_{inv} are uniformly continuous. Hence, T is an isometry between X and X_I .

Therefore, (X, ρ) is isometrically embedded in $(\bar{X}, \bar{\rho})$.

Next, we show that $(\bar{X}, \bar{\rho})$ is complete. Fix a Cauchy sequence $(F_n)_{n=1}^\infty \subseteq \bar{X}$. $\forall n \in \mathbb{N}$, let $(x_i^{(n)})_{i=1}^\infty \in F_n$. Note that $(x_i^{(n)})_{i=1}^\infty \subseteq X$ is a Cauchy sequence. $\forall n, m \in \mathbb{N}$, $\bar{\rho}(F_n, F_m) = \sigma\left((x_i^{(n)})_{i=1}^\infty, (x_i^{(m)})_{i=1}^\infty\right) = \lim_{i \in \mathbb{N}} \rho(x_i^{(n)}, x_i^{(m)})$. Define a sequence $(x_i^{(0)})_{i=1}^\infty \subseteq X$ as following. Set $n_0 = 0$. $\forall i \in \mathbb{N}$, choose $n_i \in \mathbb{N}$ with $n_i > n_{i-1}$ such that $\forall m_1, m_2 \geq n_i$, we have $\rho(x_{m_1}^{(i)}, x_{m_2}^{(i)}) < 1/i$; and set $x_i^{(0)} = x_{n_i}^{(i)}$.

Claim 4.49.1 $(x_i^{(0)})_{i=1}^\infty \in \mathcal{X}$.

Proof of claim: $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N_1 \in \mathbb{N}$ such that $\forall n, m \geq N_1$, we have $\bar{\rho}(F_n, F_m) < \epsilon$. This is valid since $(F_n)_{n=1}^\infty$ is Cauchy in $\bar{X}$. Let $N_2 \in \mathbb{N}$ be such that $1/N_2 < \epsilon$. Let $N = \max\{N_1, N_2\} \in \mathbb{N}$. $\forall i, j \geq N$, $\rho(x_i^{(0)}, x_j^{(0)}) = \rho(x_{n_i}^{(i)}, x_{n_j}^{(j)})$. Since $\bar{\rho}(F_i, F_j) = \lim_{l \in \mathbb{N}} \rho(x_l^{(i)}, x_l^{(j)}) < \epsilon$, then $\exists i_0 \in \mathbb{N}$ with $i_0 \geq \max\{n_i, n_j\}$ such that $0 \leq \rho(x_{i_0}^{(i)}, x_{i_0}^{(j)}) < \bar{\rho}(F_i, F_j) + \epsilon < 2\epsilon$. Then, we have

$$\begin{aligned} \rho(x_i^{(0)}, x_j^{(0)}) &\leq \rho(x_{n_i}^{(i)}, x_{i_0}^{(i)}) + \rho(x_{i_0}^{(i)}, x_{i_0}^{(j)}) + \rho(x_{i_0}^{(j)}, x_{n_j}^{(j)}) \\ &< 1/i + 2\epsilon + 1/j < 4\epsilon \end{aligned}$$

Hence, $(x_i^{(0)})_{i=1}^\infty \in \mathcal{X}$. This completes the proof of the claim. $\square$

Let $F_0 \in \bar{X}$ be such that $(x_i^{(0)})_{i=1}^\infty \in F_0$.

Claim 4.49.2 $\lim_{n \in \mathbb{N}} F_n = F_0$.

Proof of claim: $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N_1 \in \mathbb{N}$ such that $\forall n, m \geq N_1$, we have $\bar{\rho}(F_n, F_m) < \epsilon$. Let $N_2 \in \mathbb{N}$ be such that $1/N_2 < \epsilon$. Let $N := \max\{N_1, N_2\} \in \mathbb{N}$. $\forall m \geq N$, fix $i_0 := n_m \geq m \geq N$. $\forall i \geq i_0$, $\exists i_1 > n_i \geq i \geq n_m$ such that $0 \leq \rho(x_{i_1}^{(m)}, x_{i_1}^{(i)}) < \bar{\rho}(F_m, F_i) + \epsilon < 2\epsilon$. Then,

$$\begin{aligned} 0 &\leq \rho(x_i^{(m)}, x_i^{(0)}) = \rho(x_{i_1}^{(m)}, x_{n_i}^{(i)}) \\ &\leq \rho(x_{i_1}^{(m)}, x_{i_1}^{(i)}) + \rho(x_{i_1}^{(i)}, x_{i_1}^{(i)}) + \rho(x_{i_1}^{(i)}, x_{n_i}^{(i)}) \\ &< 1/m + 2\epsilon + 1/i < 4\epsilon \end{aligned}$$

Hence, $0 \leq \bar{\rho}(F_m, F_0) = \lim_{i \in \mathbb{N}} \rho(x_i^{(m)}, x_i^{(0)}) \leq 4\epsilon$. Therefore, we have $\lim_{n \in \mathbb{N}} \bar{\rho}(F_n, F_0) = 0$. This completes the proof of the claim. $\square$

Hence, we have shown that any Cauchy sequence $(F_n)_{n=1}^\infty \subseteq \bar{X}$ admits a limit $F_0 \in \bar{X}$. Hence, $(\bar{X}, \bar{\rho})$ is complete.

To complete the proof of (i), we need only to show that X_I is dense in $\bar{X}$. $\forall F \in \bar{X}$, let $(x_n)_{n=1}^\infty \in F$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$

with $n \geq N$, we have $\rho(x_N, x_n) < \epsilon/2$. Let $F_N := T(x_N) \in X_I$. Then, by Propositions 4.30, 3.66, and 3.67, $0 \leq \bar{\rho}(F_N, F) = \lim_{n \in \mathbb{N}} \rho(x_N, x_n) \leq \epsilon/2 < \epsilon$. Hence, $F_N \in \mathcal{B}_{(\bar{X}, \bar{\rho})}(F, \epsilon) \cap X_I \neq \emptyset$. Hence, X_I is dense in $\bar{X}$, by Proposition 3.3.

(ii) Let X_c be the closure of X in (Y, ρ_Y) . Since T is an isometry between X and X_I , then $T : X \rightarrow \bar{X}$ is uniformly continuous. By Proposition 4.46, $\exists G : X_c \rightarrow \bar{X}$ such that $G|_X = T$ and G is uniformly continuous. Note that $T_{\text{inv}} : X_I \rightarrow X \subseteq Y$ is also uniformly continuous, since T_{inv} is metric preserving. Then, by Propositions 4.39 and 4.46, $\exists H : \bar{X} \rightarrow X_c$ such that $H|_{X_I} = T_{\text{inv}}$ and H is uniformly continuous. $G \circ H : \bar{X} \rightarrow \bar{X}$ is uniformly continuous, by Proposition 4.23. Note that $(G \circ H)|_{X_I} = \text{id}_{\bar{X}}|_{X_I}$. By Proposition 3.56, we have $G \circ H = \text{id}_{\bar{X}}$. $H \circ G : X_c \rightarrow X_c$ is uniformly continuous, by Proposition 4.23. Note that $(H \circ G)|_X = \text{id}_{X_c}|_X$. By Proposition 3.56, we have $H \circ G = \text{id}_{X_c}$. Therefore, $G : X_c \rightarrow \bar{X}$ is bijective with inverse $H : \bar{X} \rightarrow X_c$, by Proposition 2.4. Hence, X_c and $\bar{X}$ are uniformly homeomorphic.

$\forall y_1, y_2 \in X_c$, $\exists (x_i^{(1)})_{i=1}^\infty, (x_i^{(2)})_{i=1}^\infty \in \mathcal{X}$ such that $\lim_{i \in \mathbb{N}} x_i^{(j)} = y_j$, $j = 1, 2$, by Proposition 4.13. Then, by Proposition 3.66, we have $G(y_j) = \lim_{i \in \mathbb{N}} G(x_i^{(j)}) = \lim_{i \in \mathbb{N}} T(x_i^{(j)})$, $j = 1, 2$. By Propositions 3.66, 3.67, and 4.30 and the fact that T is metric preserving on X , we have $\bar{\rho}(G(y_1), G(y_2)) = \lim_{i \in \mathbb{N}} \bar{\rho}(T(x_i^{(1)}), T(x_i^{(2)})) = \lim_{i \in \mathbb{N}} \rho(x_i^{(1)}, x_i^{(2)}) = \lim_{i \in \mathbb{N}} \rho_Y(x_i^{(1)}, x_i^{(2)}) = \rho_Y(y_1, y_2)$. Hence, $G : X_c \rightarrow \bar{X}$ is also metric preserving. Hence, G is an isometry between X_c and $\bar{X}$.

This completes the proof of the theorem. $\square$

4.8 Metrization of Topological Spaces

Definition 4.50 A topological space $(X, \mathcal{O})$ is said to be metrizable if there is a metric ρ on X whose natural topology is exactly the same as $\mathcal{O}$.

Proposition 4.51 Let $\mathcal{X}_\alpha := (X_\alpha, \mathcal{O}_\alpha)$ be metrizable topological spaces, $\forall \alpha \in \Lambda$, where Λ is a countable set. Then, the product topological space $\mathcal{X} = (X, \mathcal{O}) := \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha$ is metrizable.

Proof $\forall \alpha \in \Lambda$, let ρ_α be a metric on X_α whose natural topology is exactly $\mathcal{O}_\alpha$. We will distinguish three exhaustive and mutually exclusive cases: Case 1: $\Lambda = \emptyset$; Case 2: $\Lambda \neq \emptyset$ and is finite; Case 3: Λ is countably infinite.

Case 1: $\Lambda = \emptyset$. Then, $(X, \mathcal{O}) = (\{\emptyset\}, \{\emptyset, \{\emptyset\}\})$. It is clearly metrizable, where the metric ρ may be defined by $\rho(\emptyset, \emptyset) = 0$.

Case 2: $\Lambda \neq \emptyset$ and is finite. We may define a metric ρ on X be to the Cartesian metric as in Definition 4.28. By Proposition 4.29, the natural topology induced by ρ is exactly the same as the product topology $\mathcal{O}$. Hence, $\mathcal{X}$ is metrizable.

Case 3: Λ is countable infinite. The result follows by Proposition 4.34. This completes the proof of the proposition. $\square$

Definition 4.52 Let $I := [0, 1] \subset \mathbb{R}$ and A be a set. I^A is called a cube. Associate with I the subset topology of $\mathbb{R}$. Denote $\mathcal{X}^{\mathbb{N}}$ the countably infinite product space of $\mathcal{X}$. The cube $I^{\mathbb{N}}$ is metrizable, by Proposition 4.51, and is called a Hilbert cube.

Theorem 4.53 (Urysohn Metrization Theorem) Every normal topological space satisfying the second axiom of countability is metrizable.

Proof Let $\mathcal{X} := (X, \mathcal{O})$ be a second countable normal topological space. Let $(B_\alpha)_{\alpha \in \Lambda}$ be a countable basis for the topology $\mathcal{O}$. $\forall \alpha_1, \alpha_2 \in \Lambda$, if $\widetilde{B_{\alpha_1}} \cap \overline{B_{\alpha_2}} = \emptyset$, then, by Urysohn's Lemma, $\exists$ a continuous function $f_{\alpha_1 \alpha_2} : \mathcal{X} \rightarrow [0, 1] =: I \subset \mathbb{R}$ such that $f_{\alpha_1 \alpha_2}|_{\widetilde{B_{\alpha_1}}} = 0$ and $f_{\alpha_1 \alpha_2}|_{\overline{B_{\alpha_2}}} = 1$.

Define $\mathcal{F} := \left\{ f_{\alpha_1 \alpha_2} : \mathcal{X} \rightarrow I \mid \alpha_1, \alpha_2 \in \Lambda, \widetilde{B_{\alpha_1}} \cap \overline{B_{\alpha_2}} = \emptyset \right\}$. Clearly, $\mathcal{F}$ is a countable set. $\forall x_0 \in \mathcal{X}, \forall$ closed set $F \subseteq \mathcal{X}$ with $x_0 \notin F$, we have $\exists \alpha_1 \in \Lambda$ such that $x_0 \in B_{\alpha_1} \subseteq \widetilde{F}$. By Definition 3.33 and Propositions 3.34 and 3.35, $\exists U \in \mathcal{O}$ such that $\{x_0\} \subseteq U \subseteq \overline{U} \subseteq B_{\alpha_1}$. Then, $\exists \alpha_2 \in \Lambda$ such that $x_0 \in B_{\alpha_2} \subseteq U$. Then, we have $\{x_0\} \subseteq B_{\alpha_2} \subseteq \overline{B_{\alpha_2}} \subseteq \overline{U} \subseteq B_{\alpha_1}$, and hence $\widetilde{B_{\alpha_1}} \cap \overline{B_{\alpha_2}} = \emptyset$. Then, $\exists f_{\alpha_1 \alpha_2} \in \mathcal{F}$ such that $f_{\alpha_1 \alpha_2}|_{\widetilde{B_{\alpha_1}}} = 0$ and $f_{\alpha_1 \alpha_2}|_{\overline{B_{\alpha_2}}} = 1$, which implies that $f_{\alpha_1 \alpha_2}|_F = 0$ and $f_{\alpha_1 \alpha_2}(x_0) = 1$. Hence, $\mathcal{F}$ is a countable collection of continuous $[0, 1]$ -valued functions and satisfies $\forall x_0 \in \mathcal{X}, \forall$ closed set $F \subseteq \mathcal{X}$ with $x_0 \notin F$, $\exists f \in \mathcal{F} \ni \cdot f(x_0) = 1$ and $f|_F = 0$. Since $\mathcal{X}$ is normal, then it is Tychonoff and any singleton subset is closed. $\forall x, y \in \mathcal{X}$ with $x \neq y$, we have $\{y\}$ is closed and therefore, $\exists f \in \mathcal{F} \ni \cdot f(x) = 1 \neq 0 = f(y)$.

By Proposition 3.59, the equivalence map $E : \mathcal{X} \rightarrow I^{\mathcal{F}}$ is a homeomorphism between $\mathcal{X}$ and $E(\mathcal{X})$. Since I is metrizable and $\mathcal{F}$ is countable, then, by Proposition 4.51, $I^{\mathcal{F}}$ is metrizable with a metric $\rho_{I^{\mathcal{F}}}$. By Proposition 4.37, the metric $\rho_{I^{\mathcal{F}}}$ generates the subset topology on $E(\mathcal{X})$. Since E is a homeomorphism between $\mathcal{X}$ and $E(\mathcal{X})$, then $\mathcal{X}$ is metrizable with a metric $\rho_{\mathcal{X}} : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ given by $\rho_{\mathcal{X}}(x, y) := \rho_{I^{\mathcal{F}}}(E(x), E(y)), \forall x, y \in \mathcal{X}$.

This completes the proof of the theorem. $\square$

4.9 Interchange Limits

Theorem 4.54 (Joint Limit Theorem) Let $\mathcal{Y} := (Y, \rho)$ be a metric space and $(y_{\alpha, \beta})_{(\alpha, \beta) \in \mathcal{A}_1 \times \mathcal{A}_2} \subseteq \mathcal{Y}$ be a joint net. Assume that $\lim_{(\alpha, \beta) \in \mathcal{A}_1 \times \mathcal{A}_2} y_{\alpha, \beta} = \hat{y} \in \mathcal{Y}$ and, $\forall \alpha \in \mathcal{A}_1$, $\lim_{\beta \in \mathcal{A}_2} y_{\alpha, \beta} = \bar{y}_\alpha \in \mathcal{Y}$. Then, $\lim_{\alpha \in \mathcal{A}_1} \bar{y}_\alpha = \lim_{\alpha \in \mathcal{A}_1} \lim_{\beta \in \mathcal{A}_2} y_{\alpha, \beta} = \lim_{(\alpha, \beta) \in \mathcal{A}_1 \times \mathcal{A}_2} y_{\alpha, \beta}$.

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Definition 3.88 and Proposition 4.8, $\exists (\alpha_0, \beta_0) \in \mathcal{A}_1 \times \mathcal{A}_2, \forall (\alpha, \beta) \in \mathcal{A}_1 \times \mathcal{A}_2$ with $(\alpha_0, \beta_0) \prec (\alpha, \beta)$, we have

$\rho(y_{\alpha,\beta}, \hat{y}) < \epsilon/2$. Fix any $\alpha \in \mathcal{A}_1$ with $\alpha_0 \prec_1 \alpha$. By the fact that $\lim_{\beta \in \mathcal{A}_2} y_{\alpha,\beta} = \bar{y}_\alpha \in \mathcal{Y}$, $\exists \beta_\alpha \in \mathcal{A}_2$, $\forall \beta \in \mathcal{A}_2$ with $\beta_\alpha \prec_2 \beta$, we have $\rho(y_{\alpha,\beta}, \bar{y}_\alpha) < \epsilon/2$. Since $\mathcal{A}_2$ is a directed system, then $\exists \beta_1 \in \mathcal{A}_2$ such that $\beta_0 \prec_2 \beta_1$ and $\beta_\alpha \prec_2 \beta_1$. Then, we have $\rho(\bar{y}_\alpha, \hat{y}) \leq \rho(\bar{y}_\alpha, y_{\alpha,\beta_1}) + \rho(y_{\alpha,\beta_1}, \hat{y}) < \epsilon$. This shows that $\lim_{\alpha \in \mathcal{A}_1} \bar{y}_\alpha = \hat{y}$. Hence, we have $\lim_{\alpha \in \mathcal{A}_1} \bar{y}_\alpha = \lim_{\alpha \in \mathcal{A}_1} \lim_{\beta \in \mathcal{A}_2} y_{\alpha,\beta} = \lim_{(\alpha,\beta) \in \mathcal{A}_1 \times \mathcal{A}_2} y_{\alpha,\beta}$. This completes the proof of the theorem. $\square$

Corollary 4.55 *Let $\mathcal{Y} := (Y, \rho)$ be a metric space and $(y_{\alpha,\beta})_{(\alpha,\beta) \in \mathcal{A}_1 \times \mathcal{A}_2} \subseteq \mathcal{Y}$ be a joint net. Assume that $\lim_{(\alpha,\beta) \in \mathcal{A}_1 \times \mathcal{A}_2} y_{\alpha,\beta} = \hat{y} \in \mathcal{Y}$, $\forall \alpha \in \mathcal{A}_1$, $\lim_{\beta \in \mathcal{A}_2} y_{\alpha,\beta} = \bar{y}_\alpha \in \mathcal{Y}$, and, $\forall \beta \in \mathcal{A}_2$, $\lim_{\alpha \in \mathcal{A}_1} y_{\alpha,\beta} = \tilde{y}_\beta \in \mathcal{Y}$. Then, $\lim_{\alpha \in \mathcal{A}_1} \bar{y}_\alpha = \lim_{\alpha \in \mathcal{A}_1} \lim_{\beta \in \mathcal{A}_2} y_{\alpha,\beta} = \lim_{(\alpha,\beta) \in \mathcal{A}_1 \times \mathcal{A}_2} y_{\alpha,\beta} = \lim_{\beta \in \mathcal{A}_2} \tilde{y}_\beta = \lim_{\beta \in \mathcal{A}_2} \lim_{\alpha \in \mathcal{A}_1} y_{\alpha,\beta}$.*

Proof This is straightforward by Joint Limit Theorem 4.54. $\square$

Theorem 4.56 (Iterated Limit Theorem) *Let $\mathcal{Y} := (Y, \rho)$ be a complete metric space and $(y_{\alpha,\beta})_{(\alpha,\beta) \in \mathcal{A}_1 \times \mathcal{A}_2} \subseteq \mathcal{Y}$ be a joint net. Assume that*

- (i) $\forall \alpha \in \mathcal{A}_1$, $\lim_{\beta \in \mathcal{A}_2} y_{\alpha,\beta} = \bar{y}_\alpha \in \mathcal{Y}$;
- (ii) *the nets $(y_{\alpha,\beta})_{\alpha \in \mathcal{A}_1}$ converge uniformly to $\tilde{y}_\beta \in \mathcal{Y}$, $\forall \beta \in \mathcal{A}_2$.*

Then, $\lim_{\alpha \in \mathcal{A}_1} \bar{y}_\alpha = \lim_{\alpha \in \mathcal{A}_1} \lim_{\beta \in \mathcal{A}_2} y_{\alpha,\beta} = \lim_{(\alpha,\beta) \in \mathcal{A}_1 \times \mathcal{A}_2} y_{\alpha,\beta} = \lim_{\beta \in \mathcal{A}_2} \tilde{y}_\beta = \lim_{\beta \in \mathcal{A}_2} \lim_{\alpha \in \mathcal{A}_1} y_{\alpha,\beta}$.

Proof By (ii), $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \alpha_0 \in \mathcal{A}_1$, $\forall \alpha \in \mathcal{A}_1$ with $\alpha_0 \prec_1 \alpha$, $\forall \beta \in \mathcal{A}_2$, we have $\rho(y_{\alpha,\beta}, \tilde{y}_\beta) < \epsilon/4$. Fix an $\alpha_1 \in \mathcal{A}_1$ with $\alpha_0 \prec_1 \alpha_1$. Since $\lim_{\beta \in \mathcal{A}_2} y_{\alpha_1,\beta} = \bar{y}_{\alpha_1}$, then $\exists \beta_0 \in \mathcal{A}_2$, $\forall \beta \in \mathcal{A}_2$ with $\beta_0 \prec_2 \beta$, we have $\rho(y_{\alpha_1,\beta}, \bar{y}_{\alpha_1}) < \epsilon/4$. Then, $\forall \beta_1, \beta_2 \in \mathcal{A}_2$ with $\beta_0 \prec_2 \beta_1$ and $\beta_0 \prec_2 \beta_2$, we have $\rho(\tilde{y}_{\beta_1}, \tilde{y}_{\beta_2}) \leq \rho(\tilde{y}_{\beta_1}, y_{\alpha_1,\beta_1}) + \rho(y_{\alpha_1,\beta_1}, \bar{y}_{\alpha_1}) + \rho(\bar{y}_{\alpha_1}, y_{\alpha_1,\beta_2}) + \rho(y_{\alpha_1,\beta_2}, \tilde{y}_{\beta_2}) < \epsilon$. Hence, the net $(\tilde{y}_\beta)_{\beta \in \mathcal{A}_2}$ is a Cauchy net. By Proposition 4.44, $\lim_{\beta \in \mathcal{A}_2} \tilde{y}_\beta = \hat{y} \in \mathcal{Y}$.

Then, $\exists \bar{\beta} \in \mathcal{A}_2$, $\forall \beta \in \mathcal{A}_2$ with $\bar{\beta} \prec_2 \beta$, we have $\rho(\tilde{y}_\beta, \hat{y}) < \epsilon/2$. Note that $\forall (\alpha, \beta) \in \mathcal{A}_1 \times \mathcal{A}_2$ with $(\alpha_0, \bar{\beta}) \prec (\alpha, \beta)$, we have $\rho(y_{\alpha,\beta}, \hat{y}) \leq \rho(y_{\alpha,\beta}, \tilde{y}_\beta) + \rho(\tilde{y}_\beta, \hat{y}) < \epsilon$. Hence, $\lim_{(\alpha,\beta) \in \mathcal{A}_1 \times \mathcal{A}_2} y_{\alpha,\beta} = \hat{y}$. Finally, by Joint Limit Theorem 4.54, we have $\lim_{\alpha \in \mathcal{A}_1} \lim_{\beta \in \mathcal{A}_2} y_{\alpha,\beta} = \lim_{\alpha \in \mathcal{A}_1} \bar{y}_\alpha = \hat{y}$.

This completes the proof of the theorem. $\square$

Proposition 4.57 *Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $\mathcal{Y} := (Y, \rho)$ be a complete metric space, $D \subseteq \mathcal{X}$, $x_0 \in \mathcal{X} \setminus D$ be an accumulation point of D , and $(f_\alpha)_{\alpha \in \mathcal{A}}$ be a net of functions of D to $\mathcal{Y}$. Assume that $(f_\alpha)_{\alpha \in \mathcal{A}}$ converges uniformly to function $f : D \rightarrow \mathcal{Y}$ and $\lim_{x \rightarrow x_0} f_\alpha(x) = y_\alpha \in \mathcal{Y}$, $\forall \alpha \in \mathcal{A}$. Then, $\lim_{\alpha \in \mathcal{A}} y_\alpha = \lim_{\alpha \in \mathcal{A}} \lim_{x \rightarrow x_0} f_\alpha(x) = \lim_{x \rightarrow x_0} \lim_{\alpha \in \mathcal{A}} f_\alpha(x) = \lim_{x \rightarrow x_0} f(x) \in \mathcal{Y}$.*

Proof Fix any net $(x_\beta)_{\beta \in \hat{\mathcal{A}}} \subseteq D$ with x_0 as a limit. Such net exists by Proposition 3.68. By the assumption that $(f_\alpha)_{\alpha \in \mathcal{A}}$ converges uniformly to f , then, for the joint net $(f_\alpha(x_\beta))_{(\alpha, \beta) \in \mathcal{A} \times \hat{\mathcal{A}}}$, we have the nets $(f_\alpha(x_\beta))_{\alpha \in \mathcal{A}}$ converge uniformly to $f(x_\beta) \in \mathcal{Y}$, $\forall \beta \in \hat{\mathcal{A}}$. By Proposition 3.79, we have $\lim_{\beta \in \hat{\mathcal{A}}} f_\alpha(x_\beta) = y_\alpha$, $\forall \alpha \in \mathcal{A}$. By Iterated Limit Theorem 4.56, we have $\lim_{\alpha \in \mathcal{A}} y_\alpha = \lim_{\alpha \in \mathcal{A}} \lim_{\beta \in \hat{\mathcal{A}}} f_\alpha(x_\beta) = \lim_{(\alpha, \beta) \in \mathcal{A} \times \hat{\mathcal{A}}} f_\alpha(x_\beta) = \lim_{\beta \in \hat{\mathcal{A}}} \lim_{\alpha \in \mathcal{A}} f_\alpha(x_\beta) = \lim_{\beta \in \hat{\mathcal{A}}} f(x_\beta) = \hat{y} \in \mathcal{Y}$. By the arbitrariness of $(x_\beta)_{\beta \in \hat{\mathcal{A}}}$ and Proposition 3.79, we have $\hat{y} = \lim_{x \rightarrow x_0} f(x)$. Hence, we have $\lim_{\alpha \in \mathcal{A}} y_\alpha = \lim_{\alpha \in \mathcal{A}} \lim_{x \rightarrow x_0} f_\alpha(x) = \hat{y} = \lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} \lim_{\alpha \in \mathcal{A}} f_\alpha(x) \in \mathcal{Y}$. This completes the proof of the proposition. $\square$

Chapter 5

Compact and Locally Compact Spaces

5.1 Compact Spaces

Definition 5.1 A topological space $(X, \mathcal{O})$ is called compact if every open covering has a finite subcovering, that is $\forall (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}$ with $X \subseteq \bigcup_{\alpha \in \Lambda} O_\alpha$, where Λ is an index set, $\exists$ a finite set $\Lambda_N \subseteq \Lambda$ such that $X \subseteq \bigcup_{\alpha \in \Lambda_N} O_\alpha$.

A subset $K \subseteq X$ is called compact if it is compact in its subset topology. This is equivalent to any open covering of K in X has a finite subcovering of K .

Definition 5.2 Let X be a set and $\mathcal{F}$ be a collection of subsets in X . $\mathcal{F}$ is said to admit the finite intersection property if any finite subcollection of $\mathcal{F}$ has a nonempty intersection.

Proposition 5.3 A topological space $\mathcal{X}$ is compact if, and only if, any collection of closed sets $\mathcal{F}$ with the finite intersection property has a nonempty intersection.

Proof “If” Suppose $\mathcal{X}$ is not compact. Then, there exists an open covering $(O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}$ of $\mathcal{X}$, which does not have any finite subcovering. Then, $(\widetilde{O}_\alpha)_{\alpha \in \Lambda}$ is a collection of closed sets with the finite intersection property. Yet, $\bigcap_{\alpha \in \Lambda} \widetilde{O}_\alpha = (\bigcup_{\alpha \in \Lambda} O_\alpha)^\sim = \emptyset$. This contradicts the assumption. Hence, $\mathcal{X}$ is compact.

“Only if” Suppose that the result does not hold. Then, there exists a collection of closed sets $\mathcal{F}$ with finite intersection property such that $\bigcap_{F \in \mathcal{F}} F = \emptyset$. Then, we have $\bigcup_{F \in \mathcal{F}} \widetilde{F} = (\bigcap_{F \in \mathcal{F}} F)^\sim = \mathcal{X}$. By the compactness of $\mathcal{X}$, there exists a finite collection $\mathcal{F}_N \subseteq \mathcal{F}$ such that

$\bigcup_{F \in \mathcal{F}_N} \tilde{F} = \mathcal{X}$. Then, $\bigcap_{F \in \mathcal{F}_N} F = \emptyset$. This contradicts with the assumption that $\mathcal{F}$ has finite intersection property. Hence, the result holds.

This completes the proof of the proposition. $\square$

Proposition 5.4 *A topological space $\mathcal{X}$ is compact if, and only if, $\forall$ net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathcal{X}$ has a cluster point in $\mathcal{X}$.*

Proof “Only if” Fix any net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathcal{X}$. $\forall \alpha \in \mathcal{A}$, let $F_\alpha := \{x_\beta \mid \beta \in \mathcal{A}, \alpha \prec \beta\}$. Let $\overline{F_\alpha}$ be the closure of F_α . For any finite set $A \subseteq \mathcal{A}$, $\exists \alpha_0 \in \mathcal{A}$ such that $\alpha \prec \alpha_0, \forall \alpha \in A$, since $\mathcal{A}$ is a directed system. Then, $x_{\alpha_0} \in F_{\alpha_0}, \forall \alpha \in A$. Hence, we have $x_{\alpha_0} \in \bigcap_{\alpha \in A} \overline{F_\alpha} \neq \emptyset$. Hence, $(\overline{F_\alpha})_{\alpha \in \mathcal{A}}$ is a collection of closed sets in $\mathcal{X}$ with the finite intersection property. By Proposition 5.3 and the compactness of $\mathcal{X}$, we have $\bigcap_{\alpha \in \mathcal{A}} \overline{F_\alpha} \neq \emptyset$. Fix an $x \in \bigcap_{\alpha \in \mathcal{A}} \overline{F_\alpha}$. $\forall O \in \mathcal{O}$ with $x \in O, \forall \alpha_0 \in \mathcal{A}$, we have $x \in \overline{F_{\alpha_0}}$ and $O \cap F_{\alpha_0} \neq \emptyset$. Hence, $\exists \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$ such that $x_\alpha \in O \cap F_{\alpha_0} \subseteq O$. Hence, x is a cluster point of the net $(x_\alpha)_{\alpha \in \mathcal{A}}$.

“If” Suppose $\mathcal{X}$ is not compact. Then, there exists an open covering $(O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}$ of $\mathcal{X}$, where Λ is an index set, such that $\mathcal{X}$ is not covered by any finite subcollection of sets in $(O_\alpha)_{\alpha \in \Lambda}$. Let $\Gamma := \{B \subseteq \Lambda \mid B \text{ is a finite set}\}$. Define a relation $\prec$ on Γ by $\forall B_1, B_2 \in \Gamma, B_1 \prec B_2$ if $B_1 \subseteq B_2$. Clearly, $\mathcal{A} := (\Gamma, \prec)$ is a directed system. $\forall B \in \Gamma$, let $x_B \in (\bigcup_{\alpha \in B} O_\alpha)^\sim$. This constructs a net $(x_B)_{B \in \mathcal{A}} \subseteq \mathcal{X}$. By the assumption, $\exists x_0 \in \mathcal{X}$ such that x_0 is a cluster point of the net. $\forall \alpha_0 \in \Lambda, \{\alpha_0\} \in \Gamma$. $\forall O \in \mathcal{O}$ with $x_0 \in O, \exists B \in \Gamma$ with $\{\alpha_0\} \prec B$ such that $x_B \in O$. Then, $x_B \in O \cap (\bigcup_{\alpha \in B} O_\alpha)^\sim \subseteq O \cap \widetilde{O_{\alpha_0}} \neq \emptyset$. Hence, by the arbitrariness of O and Proposition 3.3, $x_0 \in \widetilde{\widetilde{O_{\alpha_0}}} = \widetilde{O_{\alpha_0}}$. Then, we have $x_0 \in \bigcap_{\alpha_0 \in \Lambda} \widetilde{O_{\alpha_0}} = (\bigcup_{\alpha \in \Lambda} O_\alpha)^\sim = \emptyset$. This is a contradiction. Hence, $\mathcal{X}$ must be compact.

This completes the proof of the proposition. $\square$

Proposition 5.5 *A closed subset of a compact space is compact. A compact subset of a Hausdorff space is closed.*

Proof Let $(X, \mathcal{O})$ be a compact topological space and $K \subseteq X$ be closed. $\forall$ open covering $(O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}$ of K , where Λ is an index set. We have $X = (\bigcup_{\alpha \in \Lambda} O_\alpha) \cup \tilde{K}$, which forms an open covering of X . By the compactness of X , there exists a finite set $\Lambda_N \subseteq \Lambda$ such that $X = (\bigcup_{\alpha \in \Lambda_N} O_\alpha) \cup \tilde{K}$. Then, we have $K \subseteq \bigcup_{\alpha \in \Lambda_N} O_\alpha$. Therefore, K is compact.

Let $(X, \mathcal{O})$ be a Hausdorff space and $K \subseteq X$ be compact. Suppose that K is not closed. Then, by Proposition 3.3, $\exists x_0 \in \overline{K} \setminus K$. $\forall x \in K$, we have $x \neq x_0$. Since $(X, \mathcal{O})$ is Hausdorff, $\exists O_1^{(x)}, O_2^{(x)} \in \mathcal{O}$ such that $x_0 \in O_1^{(x)}, x \in O_2^{(x)}$, and $O_1^{(x)} \cap O_2^{(x)} = \emptyset$. Then, $(O_2^{(x)})_{x \in K}$ forms an open covering of K . By the compactness of K , $\exists$ a finite set $K_N \subseteq K$ such that $(O_2^{(x)})_{x \in K_N}$

forms an open covering of K . Then, $x_0 \in O := \bigcap_{x \in K_N} O_1^{(x)} \in \mathcal{O}$, and $O \cap K \subseteq O \cap \left(\bigcup_{x \in K_N} O_2^{(x)} \right) = \emptyset$. This contradicts with $x_0 \in \overline{K}$, by Proposition 3.3. Hence, K is closed.

This completes the proof of the proposition. $\square$

Theorem 5.6 (Heine-Borel) *Let A be a subset of $\mathbb{R}$. A is compact if, and only if, A is closed and bounded.*

Proof “If” Consider first the special case $A = [a, b]$, where $a, b \in \mathbb{R}$ and $a < b$. Let $(O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}_{\mathbb{R}}$ be an arbitrary open covering of A , where Λ is an index set. Let $B := \{x \in A \mid \text{The interval } [a, x] \text{ can be covered by finitely many sets in } (O_\alpha)_{\alpha \in \Lambda}\}$. Clearly, $a \in B$. Let $c = \sup B$. Then, $c \in A$. $\exists \alpha_0 \in \Lambda$ such that $c \in O_{\alpha_0}$. $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $[c - \delta, c + \delta] \subseteq O_{\alpha_0}$. By the definition of c , $\exists d \in B$ such that $d \in [c - \delta, c]$. By $d \in B$, $[a, d]$ can be covered by finitely many sets in $(O_\alpha)_{\alpha \in \Lambda}$. Now, adding O_{α_0} to this finitely many sets, then $[a, c + \delta]$ is covered by finitely many sets in $(O_\alpha)_{\alpha \in \Lambda}$. Hence, $c \in B$. Note that we must have $c = b$ since otherwise $c < b$ and $\min\{c + \delta, b\} \in B$ which contradicts the definition of c . Therefore, there exists a finite subcovering of $[a, b] = A$. This shows that A is compact.

Now, let A be any closed and bounded subset of $\mathbb{R}$. Let $(O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}_{\mathbb{R}}$ be an arbitrary open covering of A , where Λ is an index set. Since A is bounded, then $A \subseteq [a, b]$, for some $a, b \in \mathbb{R}$ with $a < b$. Note that $(\bigcup_{\alpha \in \Lambda} O_\alpha) \cup \tilde{A} \supseteq [a, b]$. By the compactness of $[a, b]$, there exists a finite set $\Lambda_N \subseteq \Lambda$ such that $[a, b] \subseteq (\bigcup_{\alpha \in \Lambda_N} O_\alpha) \cup \tilde{A}$. Then, $A \subseteq \bigcup_{\alpha \in \Lambda_N} O_\alpha$. Hence, A is compact.

“Only if” Since $\mathbb{R}$ is Hausdorff, by Proposition 5.5, A is closed. Let $I_n := (-n, n) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$. Then, $\bigcup_{n \in \mathbb{N}} I_n \supseteq A$. By the compactness of A , $\exists N \in \mathbb{N}$ such that $A \subseteq \bigcup_{1 \leq n \leq N} I_n = I_N$. Hence, A is bounded.

This completes the proof of the theorem. $\square$

Proposition 5.7 *Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces, $\mathcal{X}$ be compact, and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous. Then, $f(\mathcal{X}) \subseteq \mathcal{Y}$ is compact.*

Proof Let $(O_{Y_\alpha})_{\alpha \in \Lambda} \subseteq \mathcal{O}_{\mathcal{Y}}$ be an arbitrary open covering of $f(\mathcal{X})$, where Λ is an index set. Then, $(f_{\text{inv}}(O_{Y_\alpha}))_{\alpha \in \Lambda} \subseteq \mathcal{O}_{\mathcal{X}}$ is an open covering of $\mathcal{X}$, by the continuity of f . By the compactness of $\mathcal{X}$, there exists a finite set $\Lambda_N \subseteq \Lambda$ such that $(f_{\text{inv}}(O_{Y_\alpha}))_{\alpha \in \Lambda_N}$ is a subcovering of $\mathcal{X}$. Then, by Proposition 2.5, $(O_{Y_\alpha})_{\alpha \in \Lambda_N}$ is a finite subcovering of $f(\mathcal{X})$. Hence, $f(\mathcal{X})$ is compact. This completes the proof of the proposition. $\square$

Proposition 5.8 *Let $\mathcal{X}$ be a compact space, $\mathcal{Y}$ be a Hausdorff topological space, and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be a bijective continuous function. Then, f is a homeomorphism.*

Proof By Proposition 5.7, $\mathcal{Y}$ is compact. By the assumption of the proposition, f is invertible with inverse $f_{\text{inv}} : \mathcal{Y} \rightarrow \mathcal{X}$. $\forall$ closed set $F \subseteq \mathcal{X}$.

By Proposition 5.5, F is compact. By Proposition 5.7, $f(F)$ is compact. By Proposition 5.5, $f(F)$ is closed. By Proposition 3.10, f_{inv} is continuous. Hence, f is a homeomorphism. This completes the proof of the proposition. $\square$

Definition 5.9 Let $(X, \mathcal{O})$ be a topological space and $\mathcal{U} \subseteq \mathcal{O}$ be an open covering of X . $\mathcal{V} \subseteq \mathcal{O}$ is said to be an open refinement of $\mathcal{U}$ (or to refine $\mathcal{U}$) if $\mathcal{V}$ is an open covering of X and $\forall V \in \mathcal{V}, \exists U \in \mathcal{U}$ such that $V \subseteq U$.

Proposition 5.10 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space. $\mathcal{X}$ is compact if, and only if, any open covering $\mathcal{U} \subseteq \mathcal{O}$ has a finite open refinement $\mathcal{V}$.

Proof “Only if” Let $\mathcal{U} \subseteq \mathcal{O}$ be an open covering of X . By the compactness of $\mathcal{X}$, there exists a finite set $\mathcal{V} \subseteq \mathcal{U}$ that covers X . Then, $\mathcal{V}$ is a finite open refinement of $\mathcal{U}$.

“If” Let $\mathcal{U} \subseteq \mathcal{O}$ be an arbitrary open covering of X . Then, there exists a finite open refinement $\mathcal{V}$ of $\mathcal{U}$. $\forall V \in \mathcal{V}, \exists U_V \in \mathcal{U}$ such that $V \subseteq U_V$. By Axiom of Choice and Axiom of Replacement, we may define a set $\bar{\mathcal{U}} := \{U_V \in \mathcal{U} \mid V \in \mathcal{V}\}$. Clearly, $\bar{\mathcal{U}} \subseteq \mathcal{U}$ is a finite subcovering of X . Hence, $\mathcal{X}$ is compact.

This completes the proof of the proposition. $\square$

Proposition 5.11 Let $(X, \mathcal{O})$ be a Hausdorff topological space and $(K_n)_{n=1}^{\infty}$ be a sequence of compact subsets of X with $K_{n+1} \subseteq K_n, \forall n \in \mathbb{N}$. Let $O \in \mathcal{O}$ and $\bigcap_{n \in \mathbb{N}} K_n \subseteq O$. Then, $\exists n_0 \in \mathbb{N}$ such that $K_{n_0} \subseteq O$.

Proof By Proposition 5.5, K_n is closed, $\forall n \in \mathbb{N}$. Let $O_n := O \cup \widetilde{K_n}$, $\forall n \in \mathbb{N}$. Then, we have

$$\bigcup_{n=1}^{\infty} O_n = O \cup \left(\bigcup_{n=1}^{\infty} \widetilde{K_n} \right) = O \cup \left(\bigcap_{n=1}^{\infty} K_n \right)^{\sim} = X \supseteq K_1$$

By the compactness of K_1 , $\exists N \in \mathbb{N}$ such that $K_1 \subseteq \bigcup_{n=1}^N O_n = O \cup \left(\bigcap_{n=1}^N K_n \right)^{\sim} = O \cup \widetilde{K_N}$. Then, $K_N \subseteq O$. This completes the proof of the proposition. $\square$

Proposition 5.12 Let $\mathcal{X}$ be a Hausdorff topological space and $K_{\alpha} \subseteq \mathcal{X}$ be a compact subset, $\forall \alpha \in \Lambda$, where Λ is an index set. Assume that $\bigcap_{\alpha \in \Lambda} K_{\alpha} = \emptyset$. Then, there exists a finite set $\Lambda_N \subseteq \Lambda$ such that $\bigcap_{\alpha \in \Lambda_N} K_{\alpha} = \emptyset$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: Λ is finite; Case 2: Λ is infinite. Case 1: Λ is finite. Choose $\Lambda_N := \Lambda$ and the result follows.

Case 2: Λ is infinite. $\forall \alpha \in \Lambda$, K_{α} is closed in $\mathcal{X}$ by Proposition 5.5. Note that $\bigcup_{\alpha \in \Lambda} \widetilde{K_{\alpha}} = \left(\bigcap_{\alpha \in \Lambda} K_{\alpha} \right)^{\sim} = \mathcal{X} \supseteq K_{\alpha_0}$, for some $\alpha_0 \in \Lambda$. By the compactness of K_{α_0} , there exists a finite set $\bar{\Lambda}_N \subseteq \Lambda$ such that $K_{\alpha_0} \subseteq \bigcup_{\alpha \in \bar{\Lambda}_N} \widetilde{K_{\alpha}}$. Let $\Lambda_N := \bar{\Lambda}_N \cup \{\alpha_0\}$, which is a finite set. Then, $\bigcap_{\alpha \in \Lambda_N} K_{\alpha} = K_{\alpha_0} \cap \left(\bigcap_{\alpha \in \bar{\Lambda}_N} K_{\alpha} \right) = \emptyset$. This completes the proof of the proposition. $\square$

Proposition 5.13 *Let $\mathcal{X} := (X, \mathcal{O})$ be a Hausdorff topological space, $K_1, K_2 \subseteq \mathcal{X}$ be compact subsets, and $K_1 \cap K_2 = \emptyset$. Then, $\exists O_1, O_2 \in \mathcal{O}$ such that $K_1 \subseteq O_1$, $K_2 \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$.*

Proof Since $(X, \mathcal{O})$ is Hausdorff, then, $\forall x_1 \in K_1$ and $\forall x_2 \in K_2$, we must have $x_1 \neq x_2$ and $\exists O_{x_1, x_2}^{(1)}, O_{x_1, x_2}^{(2)} \in \mathcal{O}$ such that $x_j \in O_{x_1, x_2}^{(j)}$, $j = 1, 2$, and $O_{x_1, x_2}^{(1)} \cap O_{x_1, x_2}^{(2)} = \emptyset$. Then, $\forall x_1 \in K_1$, $K_2 \subseteq \bigcup_{x_2 \in K_2} O_{x_1, x_2}^{(2)}$. By the compactness of K_2 , there exists a finite set $K_{2, x_1} \subseteq K_2$ such that $K_2 \subseteq \bigcup_{x_2 \in K_{2, x_1}} O_{x_1, x_2}^{(2)} =: O_{x_1}^{(2)} \in \mathcal{O}$. It is clear that $x_1 \in \bigcap_{x_2 \in K_{2, x_1}} O_{x_1, x_2}^{(1)} =: O_{x_1}^{(1)} \in \mathcal{O}$ and $O_{x_1}^{(1)} \cap O_{x_1}^{(2)} = \emptyset$. Then, $K_1 \subseteq \bigcup_{x_1 \in K_1} O_{x_1}^{(1)}$. By the compactness of K_1 , there exists a finite set $K_{1N} \subseteq K_1$ such that $K_1 \subseteq \bigcup_{x_1 \in K_{1N}} O_{x_1}^{(1)} =: O_1 \in \mathcal{O}$. Note that $K_2 \subseteq \bigcap_{x_1 \in K_{1N}} O_{x_1}^{(2)} =: O_2 \in \mathcal{O}$. Clearly, $O_1 \cap O_2 = \emptyset$. This completes the proof of the proposition. $\square$

Proposition 5.14 *A compact Hausdorff topological space is normal.*

Proof Let $(X, \mathcal{O})$ be a compact Hausdorff topological space. Fix any closed sets $F_1, F_2 \subseteq X$ with $F_1 \cap F_2 = \emptyset$. By Proposition 5.5, F_1 and F_2 are compact. By Proposition 5.13, $\exists O_1, O_2 \in \mathcal{O}$ such that $F_1 \subseteq O_1$, $F_2 \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$. Hence, $(X, \mathcal{O})$ is normal. This completes the proof of the proposition. $\square$

Proposition 5.15 *Let X be a set and $\mathcal{O}$ and $\mathcal{O}_1$ be topologies on X . Assume that $\mathcal{O}_1$ is weaker than $\mathcal{O}$, that is $\mathcal{O}_1 \subset \mathcal{O}$, and $(X, \mathcal{O})$ is compact. Then, $(X, \mathcal{O}_1)$ is compact.*

Proof Let $\mathcal{U}_1 \subseteq \mathcal{O}_1$ be any open covering of $(X, \mathcal{O}_1)$. Then, it is an open covering of $(X, \mathcal{O})$. By the compactness of $(X, \mathcal{O})$, there exists finite subcovering $\mathcal{U}_N \subseteq \mathcal{U}_1$ of $(X, \mathcal{O})$. Clearly $\mathcal{U}_N \subseteq \mathcal{O}_1$ is a finite subcovering of $(X, \mathcal{O}_1)$. Therefore, $(X, \mathcal{O}_1)$ is compact. This completes the proof of the proposition. $\square$

Proposition 5.16 *Let X be a set and $\mathcal{O}$ and $\mathcal{O}_1$ be topologies on X . Assume that $\mathcal{O}_1$ is stronger than $\mathcal{O}$, that is $\mathcal{O} \subset \mathcal{O}_1$, and $(X, \mathcal{O})$ is Hausdorff. Then, $(X, \mathcal{O}_1)$ is Hausdorff.*

Proof $\forall x, y \in X$ with $x \neq y$, $\exists O_1, O_2 \in \mathcal{O}$ such that $x \in O_1$, $y \in O_2$, and $O_1 \cap O_2 = \emptyset$, since $(X, \mathcal{O})$ is Hausdorff. Note that $O_1, O_2 \in \mathcal{O}_1$ since $\mathcal{O} \subseteq \mathcal{O}_1$. Hence, $(X, \mathcal{O}_1)$ is Hausdorff. This completes the proof of the proposition. $\square$

Proposition 5.17 *Let $(X, \mathcal{O})$ be a compact Hausdorff space. Then, any weaker topology $\mathcal{O}_1 \subset \mathcal{O}$ is not Hausdorff, and any stronger topology $\mathcal{O}_2 \supset \mathcal{O}$ is not compact.*

Proof Since $\mathcal{O}_1 \subset \mathcal{O}$, then $\exists O_0 \in \mathcal{O} \setminus \mathcal{O}_1$. Then, $\widetilde{O_0}$ is closed in $(X, \mathcal{O})$. By Proposition 5.5, $\widetilde{O_0}$ is compact in $(X, \mathcal{O})$. By Proposition 5.15, $\widetilde{O_0}$ is compact in $(X, \mathcal{O}_1)$. Suppose $(X, \mathcal{O}_1)$ is Hausdorff. Then, by Proposition 5.5, $\widetilde{O_0}$ is closed in $(X, \mathcal{O}_1)$. This implies that $O_0 \in \mathcal{O}_1$, which contradicts with $O_0 \in \mathcal{O} \setminus \mathcal{O}_1$. Hence, $(X, \mathcal{O}_1)$ is not Hausdorff.

Since $\mathcal{O}_2 \supset \mathcal{O}$, then $\exists O_0 \in \mathcal{O}_2 \setminus \mathcal{O}$. Suppose $(X, \mathcal{O}_2)$ is compact. Then, by Proposition 5.5, $\widetilde{O_0}$ is compact in $(X, \mathcal{O}_2)$. By Proposition 5.15, $\widetilde{O_0}$ is compact in $(X, \mathcal{O})$. By Proposition 5.5, $\widetilde{O_0}$ is closed in $(X, \mathcal{O})$. Then, $O_0 \in \mathcal{O}$, which contradicts with $O_0 \in \mathcal{O}_2 \setminus \mathcal{O}$. Hence, $(X, \mathcal{O}_2)$ is not compact.

This completes the proof of the proposition. $\square$

Proposition 5.18 *Let $\mathcal{X}$ be a compact topological space, $\mathcal{Y}$ be a Hausdorff topological space, $\mathcal{Z}$ be a topological space, $f : \mathcal{X} \rightarrow \mathcal{Y}$ be surjective and continuous, and $g : \mathcal{Y} \rightarrow \mathcal{Z}$ be such that $T := g \circ f : \mathcal{X} \rightarrow \mathcal{Z}$ is continuous. Then, g is continuous.*

Proof By Proposition 5.7 and the compactness of $\mathcal{X}$, $\mathcal{Y} = f(\mathcal{X})$ is compact. Then, $\mathcal{Y}$ is compact and Hausdorff.

Claim 5.18.1 $\forall O_Z \subseteq \mathcal{Z}, g_{\text{inv}}(O_Z) = f(T_{\text{inv}}(O_Z))$.

Proof of claim: $\forall y_0 \in g_{\text{inv}}(O_Z) \subseteq \mathcal{Y}, \exists x_0 \in \mathcal{X}$ such that $y_0 = f(x_0)$ since f is surjective. Then, $g(y_0) = g(f(x_0)) = T(x_0) \in O_Z$. This implies that $x_0 \in T_{\text{inv}}(O_Z)$. Then, $y_0 \in f(T_{\text{inv}}(O_Z))$. Hence, $g_{\text{inv}}(O_Z) \subseteq f(T_{\text{inv}}(O_Z))$.

On the other hand, $\forall y_0 \in f(T_{\text{inv}}(O_Z)), \exists x_0 \in T_{\text{inv}}(O_Z)$ such that $y_0 = f(x_0)$. Then, $O_Z \ni T(x_0) = g(f(x_0)) = g(y_0)$. This implies that $y_0 \in g_{\text{inv}}(O_Z)$. Hence, $f(T_{\text{inv}}(O_Z)) \subseteq g_{\text{inv}}(O_Z)$.

Therefore, we have $g_{\text{inv}}(O_Z) = f(T_{\text{inv}}(O_Z))$. This completes the proof of the claim. $\square$

$\forall$ closed set $F \subseteq \mathcal{Z}$. $T_{\text{inv}}(F)$ is closed in $\mathcal{X}$ by Proposition 3.10 and the continuity of T . Then, $T_{\text{inv}}(F)$ is compact by Proposition 5.5. This implies that, by Proposition 5.7, $f(T_{\text{inv}}(F))$ is compact. By Proposition 5.5, $g_{\text{inv}}(F) = f(T_{\text{inv}}(F))$ is closed. Then, by Proposition 3.10 and the arbitrariness of F , g is continuous.

This completes the proof of the proposition. $\square$

Definition 5.19 *A compact and connected Hausdorff topological space containing more than one point is called a continuum.*

Proposition 5.20 *Let $\mathcal{X} := (X, \mathcal{O})$ be a Hausdorff topological space, $(K_n)_{n=1}^{\infty}$ be a sequence of compact and connected subsets of $\mathcal{X}$, and $K_{n+1} \subseteq K_n, \forall n \in \mathbb{N}$. Then, $\bigcap_{n=1}^{\infty} K_n =: K$ is compact and connected.*

Proof By Proposition 5.5, K_n is closed, $\forall n \in \mathbb{N}$. Then, K is closed. Then, K is closed in the subset topology of K_1 by Proposition 3.5. By Proposition 5.5, K is compact.

Suppose K is not connected. Then, $\exists O_1, O_2 \in \mathcal{O}$ such that $K = O_{K1} \cup O_{K2} := (O_1 \cap K) \cup (O_2 \cap K)$, $O_{K1} \neq \emptyset$, $O_{K2} \neq \emptyset$, and $O_{K1} \cap O_{K2} = \emptyset$. Define $\bar{K}_i := K \setminus O_{Ki}$, $i = 1, 2$. Then, $i = 1, 2$, $\bar{K}_i = K \cap \widetilde{O_i}$, which is closed and therefore compact by Proposition 5.5. Note that $O_{K1} = \bar{K}_2$ and $O_{K2} = \bar{K}_1$. Then, O_{K1} and O_{K2} are disjoint compact sets. By Proposition 5.13, $\exists \bar{O}_1, \bar{O}_2 \in \mathcal{O}$ such that $O_{Ki} \subseteq \bar{O}_i$, $i = 1, 2$, and $\bar{O}_1 \cap \bar{O}_2 = \emptyset$. Then, $\bar{O}_1 \cup \bar{O}_2 \supseteq K$, by Proposition 5.11, $\exists n_0 \in \mathbb{N}$ such that $\bar{O}_1 \cup \bar{O}_2 \supseteq K_{n_0}$. Note that $\bar{O}_i \cap K_{n_0} \supseteq O_{Ki} \neq \emptyset$ and is open in the subset topology of K_{n_0} , $i = 1, 2$, then, $\bar{O}_1 \cap K_{n_0}$ and $\bar{O}_2 \cap K_{n_0}$ form a separation of K_{n_0} . This contradicts with the assumption that K_{n_0} is connected. Hence, K is connected.

This completes the proof of the proposition. $\square$

We note that the union of finitely many compact subsets of a topological space is compact.

5.2 Countable and Sequential Compactness

Definition 5.21 Let $(X, \mathcal{O})$ be a topological space. A subset $K \subseteq X$ is said to be countably compact if any countable open covering of K admits a finite subcovering.

Proposition 5.22 A topological space is compact if, and only if, it is Lindelöf and countably compact.

Proof This is straightforward and therefore omitted. $\square$

Proposition 5.23 A second countable topological space is compact if, and only if, it is countably compact.

Proof A second countable topological space is Lindelöf, by Proposition 3.24. Then, the result follows from Proposition 5.22. $\square$

Proposition 5.24 Let $\mathcal{X}$ be a countably compact topological space, $\mathcal{Y}$ be a topological space, and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous. Then, $f(\mathcal{X})$ is countably compact.

Proof Fix any countable open covering $(O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}_{\mathcal{Y}}$ of $f(\mathcal{X})$, where Λ is a countable index set. By Proposition 2.5 and the continuity of f , $(f_{\text{inv}}(O_\alpha))_{\alpha \in \Lambda}$ is a countable open covering of $\mathcal{X}$. By the countable compactness of $\mathcal{X}$, there exists a finite set $\Lambda_N \subseteq \Lambda$ such that $\bigcup_{\alpha \in \Lambda_N} f_{\text{inv}}(O_\alpha) = \mathcal{X}$. Then, we have $\bigcup_{\alpha \in \Lambda_N} O_\alpha \supseteq \bigcup_{\alpha \in \Lambda_N} f(f_{\text{inv}}(O_\alpha)) = f(\bigcup_{\alpha \in \Lambda_N} f_{\text{inv}}(O_\alpha)) = f(\mathcal{X})$, by Proposition 2.5. Hence, $f(\mathcal{X})$ is countably compact. This completes the proof of the proposition. $\square$

Definition 5.25 Let $\mathcal{X}$ be a topological space. It is said to have the Bolzano-Weierstrass property if $\forall (x_n)_{n=1}^\infty \subseteq \mathcal{X}$, there exists a cluster point $x \in \mathcal{X}$ of the sequence.

Proposition 5.26 A topological space $\mathcal{X}$ is countably compact if, and only if, it has the Bolzano-Weierstrass property.

Proof “If” Suppose $\mathcal{X}$ is not countably compact. Then, there exists an open covering $(O_n)_{n=1}^\infty \subseteq \mathcal{O}$ of $\mathcal{X}$ which does not have any finite sub-covering. $\forall n \in \mathbb{N}$, $\exists x_n \in \bigcup_{i=1}^n O_i = \bigcap_{i=1}^n \widetilde{O_i}$. Then, $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$. By the Bolzano-Weierstrass property, $\exists x_0 \in \mathcal{X}$ such that x_0 is a cluster point of $(x_n)_{n=1}^\infty$. By Proposition 3.3, $x_0 \in (x_n)_{n=1}^\infty$. $\forall n \in \mathbb{N}$, since $x_0 \in (x_i)_{i=n}^\infty$, $(x_i)_{i=n}^\infty \subseteq \bigcap_{i=1}^n \widetilde{O_i}$, and $\bigcap_{i=1}^n \widetilde{O_i}$ is closed, then $x_0 \in \bigcap_{i=1}^n \widetilde{O_i}$. Then, $x_0 \in \bigcap_{n=1}^\infty \widetilde{O_i}$, which contradicts with the fact that $(O_n)_{n=1}^\infty$ is an open covering of $\mathcal{X}$. Therefore, $\mathcal{X}$ is countably compact.

“Only if” Fix any sequence $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$. Let $B_n := (x_i)_{i=n}^\infty$, $\forall n \in \mathbb{N}$. Suppose $\bigcap_{n=1}^\infty \overline{B_n} = \emptyset$. Then, $(\overline{B_n})_{n \in \mathbb{N}}$ form a countable open covering of $\mathcal{X}$. By the countable compactness of $\mathcal{X}$, $\exists N \in \mathbb{N}$ such that $\bigcup_{n=1}^N \overline{B_n} = \mathcal{X}$. Hence, we have $\bigcap_{n=1}^N \overline{B_n} = \emptyset$. This contradicts with the fact that $x_{N+1} \in \bigcap_{n=1}^N \overline{B_n}$. Therefore, $\bigcap_{n=1}^\infty \overline{B_n} \neq \emptyset$. Let $x_0 \in \bigcap_{n=1}^\infty \overline{B_n} \subseteq \mathcal{X}$. $\forall O \in \mathcal{O}$ with $x_0 \in O$, $\forall n \in \mathbb{N}$, $x_0 \in \overline{B_n}$ implies that $O \cap B_n \neq \emptyset$ by Proposition 3.3. Then, $\exists m \in \mathbb{N}$ with $m \geq n$ such that $x_m \in O$. This shows that x_0 is a cluster point of $(x_n)_{n=1}^\infty$. Hence, $\mathcal{X}$ has the Bolzano-Weierstrass property.

This completes the proof of the proposition. $\square$

Definition 5.27 A topological space $\mathcal{X}$ is said to be sequentially compact if $\forall (x_n)_{n=1}^\infty \subseteq \mathcal{X}$, there exists a subsequence $(x_{n_i})_{i=1}^\infty$ that converges to some $x_0 \in \mathcal{X}$.

Proposition 5.28 Let $\mathcal{X}$ be a topological space. Then, the following statements hold.

- (i) If it is sequentially compact, then $\mathcal{X}$ is countably compact.
- (ii) If it is first countable and countably compact, then $\mathcal{X}$ is sequentially compact.

Proof (i) $\mathcal{X}$ is sequentially compact implies that $\mathcal{X}$ has the Bolzano-Weierstrass property since the limit of a convergent subsequence is a cluster point for the subsequence and hence a cluster point for the original sequence. Then, $\mathcal{X}$ is countably compact by Proposition 5.26.

(ii) Let $\mathcal{X}$ be countably compact and first countable. Fix any $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$. By Proposition 5.26, $\mathcal{X}$ has the Bolzano-Weierstrass property. Then, $\exists x_0 \in \mathcal{X}$ such that x_0 is a cluster point of $(x_n)_{n=1}^\infty$. Let $\mathcal{B}_{x_0} := (B_n)_{n=1}^\infty \subseteq \mathcal{O}$ be a countable basis at x_0 . (In case $\mathcal{B}_{x_0}$ is finite, we will pad $\mathcal{X}$ as

additional basis sets to make it countably infinite.) Let $n_0 = 0$. $\forall i \in \mathbb{N}$, $\exists n_i \in \mathbb{N}$ with $n_i > n_{i-1}$ such that $x_{n_i} \in \bigcap_{j=1}^i B_j \in \mathcal{O}$. Clearly, $(x_{n_i})_{i=1}^\infty$ is a subsequence of $(x_n)_{n=1}^\infty$ and admits a limit point x_0 . Hence, $\mathcal{X}$ is sequentially compact.

This completes the proof of the proposition. $\square$

The relationships between different compactness concepts are: compactness implies countable compactness, which is equivalent to the Bolzano-Weierstrass property; sequential compactness implies countable compactness.

5.3 Real-Valued Functions and Compactness

Proposition 5.29 *Let $\mathcal{X}$ be a countably compact topological space and $f : \mathcal{X} \rightarrow \mathbb{R}$ be continuous. Then, f is bounded, that is, $\exists a, b \in \mathbb{R}$ such that $a \leq f(x) \leq b$, $\forall x \in \mathcal{X}$. Furthermore, if $\mathcal{X} \neq \emptyset$, $\exists x_m, x_M \in \mathcal{X}$ such that $f(x_m) = \min_{\bar{x} \in \mathcal{X}} f(\bar{x}) \leq f(x) \leq \max_{\bar{x} \in \mathcal{X}} f(\bar{x}) = f(x_M)$, $\forall x \in \mathcal{X}$, i. e., f achieves its minimum and maximum on $\mathcal{X}$.*

Proof Note that $\mathbb{R} = \bigcup_{n \in \mathbb{N}} (-n, n)$. Then, $\mathcal{X} = \bigcup_{n \in \mathbb{N}} f_{\text{inv}}((-n, n))$, by Proposition 2.5. By the continuity of f , $f_{\text{inv}}((-n, n))$ is open, $\forall n \in \mathbb{N}$. By the countable compactness of $\mathcal{X}$, $\exists N \in \mathbb{N}$ such that $\mathcal{X} = \bigcup_{n=1}^N f_{\text{inv}}((-n, n)) = f_{\text{inv}}((-N, N))$. By Proposition 2.5, $f(\mathcal{X}) \subseteq (-N, N)$. Hence, f is bounded.

Let $\mathcal{X} \neq \emptyset$. Define $M := \sup_{x \in \mathcal{X}} f(x)$ and $m := \inf_{x \in \mathcal{X}} f(x)$. Then, $-N \leq m \leq M \leq N$. $\forall n \in \mathbb{N}$, $\exists x_n \in \mathcal{X}$ such that $M - 1/n < f(x_n) \leq M$. We thus obtain a sequence $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$. By Proposition 5.26 and the countable compactness of $\mathcal{X}$, $(x_n)_{n=1}^\infty$ admits a cluster point $x_M \in \mathcal{X}$. By Proposition 3.66, the sequence $(f(x_n))_{n=1}^\infty$ admits a cluster point $f(x_M)$. By construction, $\lim_{n \in \mathbb{N}} f(x_n) = M$, which is the only cluster point of $(f(x_n))_{n=1}^\infty$. Then, we have $f(x_M) = M$. By an argument that is similar to the above, $\exists x_m \in \mathcal{X}$ such that $f(x_m) = m$. This completes the proof of the proposition. $\square$

Proposition 5.30 *Let $\mathcal{X}$ be a countably compact topological space and $f : \mathcal{X} \rightarrow \mathbb{R}$ be upper semicontinuous. Then, f is bounded from above, that is, $\exists b \in \mathbb{R}$ such that $f(x) \leq b$, $\forall x \in \mathcal{X}$. Furthermore, if $\mathcal{X} \neq \emptyset$, then $\exists x_M \in \mathcal{X}$ such that $f(x) \leq \max_{\bar{x} \in \mathcal{X}} f(\bar{x}) = f(x_M)$, $\forall x \in \mathcal{X}$, i. e., f achieves its maximum on $\mathcal{X}$.*

Proof Note that $\mathbb{R} = \bigcup_{n \in \mathbb{N}} (-\infty, n)$. Then, $\mathcal{X} = \bigcup_{n \in \mathbb{N}} f_{\text{inv}}((-\infty, n))$, by Proposition 2.5. By the upper semicontinuity of f , $f_{\text{inv}}((-\infty, n))$ is open, $\forall n \in \mathbb{N}$. By the countable compactness of $\mathcal{X}$, $\exists N \in \mathbb{N}$ such that $\mathcal{X} = \bigcup_{n=1}^N f_{\text{inv}}((-\infty, n)) = f_{\text{inv}}((-\infty, N))$. By Proposition 2.5, $f(\mathcal{X}) \subseteq (-\infty, N)$.

Let $\mathcal{X} \neq \emptyset$. Define $M := \sup_{x \in \mathcal{X}} f(x)$. Then, $-\infty < M \leq N$. $\forall n \in \mathbb{N}$, $\exists x_n \in \mathcal{X}$ such that $M - 1/n < f(x_n) \leq M$. We thus obtain a sequence $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$. By Proposition 5.26 and the countable compactness of $\mathcal{X}$, $(x_n)_{n=1}^\infty$ admits a cluster point $x_M \in \mathcal{X}$. By Proposition 3.66, the sequence $(f(x_n))_{n=1}^\infty$ admits a cluster point $f(x_M)$. By construction, $\lim_{n \in \mathbb{N}} f(x_n) = M$, which is the only cluster point of $(f(x_n))_{n=1}^\infty$. Then, we have $f(x_M) = M$. This completes the proof of the proposition. $\square$

Lemma 5.31 (Dini's Lemma) *Let $\mathcal{X} := (X, \mathcal{O})$ be a countably compact topological space and $(f_n)_{n=1}^\infty$ be a sequence of upper semicontinuous real-valued functions on $\mathcal{X}$. Assume that $\forall x \in \mathcal{X}$, $(f_n(x))_{n=1}^\infty \subseteq \mathbb{R}$ is nonincreasing and converges to 0. Then, $(f_n)_{n=1}^\infty$ converges to the zero function uniformly on $\mathcal{X}$.*

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $O_n := \{x \in \mathcal{X} \mid f_n(x) < \epsilon\}$, $\forall n \in \mathbb{N}$. By the upper semicontinuity of f_n , $O_n \in \mathcal{O}$, $\forall n \in \mathbb{N}$. Since $\lim_{n \in \mathbb{N}} f_n(x) = 0$, $\forall x \in \mathcal{X}$, then $\mathcal{X} = \bigcup_{n \in \mathbb{N}} O_n$. By the countable compactness of $\mathcal{X}$, $\exists N \in \mathbb{N}$ such that $\mathcal{X} = \bigcup_{n=1}^N O_n$. $\forall n \in \mathbb{N}$ with $n \geq N$, $\forall x \in \mathcal{X}$, $\exists i \in \{1, \dots, N\}$ such that $x \in O_i$, then $0 \leq f_n(x) \leq f_i(x) < \epsilon$. Hence, the result is established. This completes the proof of the lemma. $\square$

Proposition 5.32 *Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $(f_n)_{n=1}^\infty$ be a sequence of upper semicontinuous real-valued functions on $\mathcal{X}$. Assume that $\forall x \in \mathcal{X}$, $(f_n(x))_{n=1}^\infty \subseteq \mathbb{R}$ is nonincreasing and converges to $f_0(x)$. Then,*

- (i) f_0 is upper semicontinuous;
- (ii) if, in addition, $\mathcal{X}$ is countably compact and f_0 is lower semicontinuous, then, $(f_n)_{n=1}^\infty$ converges to f_0 uniformly on $\mathcal{X}$.

Proof (i) $\forall a \in \mathbb{R}$, since, $\forall x \in \mathcal{X}$, $f_0(x) < a \Leftrightarrow f_n(x) < a$ for all $n \geq m$, for some $m \in \mathbb{N}$; and $f_{i \text{ inv}}((-\infty, a)) \subseteq f_{j \text{ inv}}((-\infty, a))$, $\forall i \leq j$, then $f_{0 \text{ inv}}((-\infty, a)) = \bigcup_{n \in \mathbb{N}} f_{n \text{ inv}}((-\infty, a))$. By the upper semicontinuity of f_n , $f_{n \text{ inv}}((-\infty, a)) \in \mathcal{O}$, $\forall n \in \mathbb{N}$. Hence, $f_{0 \text{ inv}}((-\infty, a)) \in \mathcal{O}$. Therefore, f_0 is upper semicontinuous.

(ii) f_0 is continuous by Proposition 3.16. Then, by Proposition 3.16, $(f_n - f_0)_{n=1}^\infty$ is a sequence of upper semicontinuous functions, which is nonincreasing and converges to 0 pointwise. By Dini's Lemma, $(f_n - f_0)_{n=1}^\infty$ converges to the zero function uniformly. Then, $(f_n)_{n=1}^\infty$ converges to f_0 uniformly.

This completes the proof of the proposition. $\square$

Proposition 5.33 *Let $\mathcal{X}$ be a topological space and $(f_n)_{n=1}^\infty$ be a sequence of upper semicontinuous real-valued functions on $\mathcal{X}$ that converges uniformly to $f : \mathcal{X} \rightarrow \mathbb{R}$. Then, f is upper semicontinuous.*

Proof $\forall x_0 \in \mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$, we have $|f_N(x) - f(x)| < \epsilon/3$, $\forall x \in \mathcal{X}$. Since f_N is upper semicontinuous, then, by Proposition 3.15, $\exists U \in \mathcal{O}$ with $x_0 \in U$, $\forall x \in U$, we have $f_N(x) < f_N(x_0) + \epsilon/3$. Then, $\forall x \in U$, we have

$$f(x) - f(x_0) = f(x) - f_N(x) + f_N(x) - f_N(x_0) + f_N(x_0) - f(x_0) < \epsilon$$

Therefore, f is upper semicontinuous at x_0 . By Proposition 3.15, f is upper semicontinuous.

This completes the proof of the proposition. $\square$

5.4 Compactness in Metric Spaces

Lemma 5.34 *A sequentially compact metric space is totally bounded.*

Proof Let $\mathcal{X}$ be a sequentially compact metric space. Suppose $\mathcal{X}$ is not totally bounded. Then, $\exists \epsilon_0 \in (0, \infty) \subset \mathbb{R}$, $\mathcal{X}$ can not be covered by finitely many open balls with radius ϵ_0 . Clearly, $\mathcal{X} \neq \emptyset$. Fix an $x_1 \in \mathcal{X}$. $\forall n \in \mathbb{N}$ with $n \geq 2$, $\exists x_n \in \mathcal{X} \setminus \left(\bigcup_{i=1}^{n-1} \mathcal{B}(x_i, \epsilon_0) \right)$. The sequence $(x_n)_{n=1}^\infty$ is such that $\rho(x_n, x_m) \geq \epsilon_0$, $\forall n, m \in \mathbb{N}$ with $n \neq m$. Clearly, this sequence does not have any convergent subsequence. This contradicts with the assumption that $\mathcal{X}$ is sequentially compact. Therefore, $\mathcal{X}$ must be totally bounded. This completes the proof of the lemma. $\square$

Definition 5.35 *Let $\mathcal{X}$ be a sequentially compact metric space and $(O_\alpha)_{\alpha \in \Lambda}$ be an open covering of $\mathcal{X}$, where Λ is an index set. The Lebesgue number of $(O_\alpha)_{\alpha \in \Lambda}$ is defined by*

$$\epsilon := \sup \{ \delta \in (0, \infty) \subset \mathbb{R} \mid \forall x \in \mathcal{X}, \exists \bar{x} \in \mathcal{X} \ni \rho(x, \bar{x}) \geq \delta \text{ and } \mathcal{B}(x, \delta) \subseteq O_\alpha \text{ for some } \alpha \in \Lambda \}$$

Lemma 5.36 *Let $\mathcal{X}$ be a sequentially compact metric space containing at least two distinct points and $(O_\alpha)_{\alpha \in \Lambda}$ be an open covering of $\mathcal{X}$, where Λ is an index set. Then, the Lebesgue number ϵ of $(O_\alpha)_{\alpha \in \Lambda}$ is positive and belongs to $\mathbb{R}$.*

Proof Define a function $\phi : \mathcal{X} \rightarrow \mathbb{R}$ by, $\forall x \in \mathcal{X}$,

$$\begin{aligned} \phi(x) &= \sup \{ c \in (0, \infty) \subset \mathbb{R} \mid \mathcal{B}(x, c) \subseteq O_\alpha \text{ for some } \alpha \in \Lambda \text{ and} \\ &\quad \exists \bar{x} \in \mathcal{X} \ni \rho(\bar{x}, x) \geq c \} \end{aligned}$$

By Proposition 5.34, $\mathcal{X}$ is totally bounded. Then, $\forall x \in \mathcal{X}$, $\exists c_x \in (0, \infty) \subset \mathbb{R}$, such that $\nexists \bar{x} \in \mathcal{X}$ with $\rho(\bar{x}, x) \geq c_x$. Then, $\phi(x) \leq c_x$. Since $(O_\alpha)_{\alpha \in \Lambda}$ is an open cover of $\mathcal{X}$, $\exists c_1 \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(x, c_1) \subseteq O_\alpha$ for some $\alpha \in \Lambda$. Since $\mathcal{X}$ contains at least two distinct points, then $\exists \bar{x} \in \mathcal{X}$

with $\bar{x} \neq x$ such that $\rho(\bar{x}, x) =: c_2 > 0$. Then, $c := \min\{c_1, c_2\} > 0$ and $\phi(x) \geq c$. Hence, $0 < \phi(x) < +\infty, \forall x \in \mathcal{X}$.

Next, we will show that ϕ is continuous. $\forall x, y \in \mathcal{X}$, we will distinguish two exhaustive and mutually exclusive cases: Case 1: $\phi(x) - \rho(x, y) \leq 0$; Case 2: $\phi(x) - \rho(x, y) =: l > 0$. Case 1: $\phi(x) - \rho(x, y) \leq 0$. Then, we have $\phi(y) > 0 \geq \phi(x) - \rho(x, y)$. Case 2: $\phi(x) - \rho(x, y) =: l > 0$. Then, $\forall \delta \in (0, l)$, $\exists \alpha \in \Lambda$ and $\exists \bar{x} \in \mathcal{X}$ such that $\mathcal{B}(x, \phi(x) - \delta) \subseteq O_\alpha$ and $\rho(x, \bar{x}) \geq \phi(x) - \delta$. Then, $\rho(\bar{x}, y) \geq \rho(x, \bar{x}) - \rho(x, y) \geq \phi(x) - \delta - \rho(x, y) = l - \delta$. $\forall \bar{y} \in \mathcal{B}(y, l - \delta)$, we have $\rho(x, \bar{y}) \leq \rho(x, y) + \rho(y, \bar{y}) < l - \delta + \rho(x, y) = \phi(x) - \delta$. Then, $\bar{y} \in \mathcal{B}(x, \phi(x) - \delta) \subseteq O_\alpha$. This implies that $\mathcal{B}(y, l - \delta) \subseteq O_\alpha$. Hence, we have $\phi(y) \geq l - \delta$. By the arbitrariness of δ , we have $\phi(y) \geq l = \phi(x) - \rho(x, y)$. Hence, in both cases, we have arrived at $\phi(y) \geq \phi(x) - \rho(x, y)$. Then, $\phi(x) - \phi(y) \leq \rho(x, y)$. This further implies that $\phi(y) - \phi(x) \leq \rho(y, x)$. Hence, we have $|\phi(x) - \phi(y)| \leq \rho(x, y)$. Hence, ϕ is continuous.

By the sequential compactness of $\mathcal{X}$, $\mathcal{X} \neq \emptyset$, and Propositions 5.28 and 5.29, $\exists x_m \in \mathcal{X}$ such that $\phi(x_m) = \min_{x \in \mathcal{X}} \phi(x) \in (0, \infty) \subset \mathbb{R}$. Then, the Lebesgue number $\epsilon = \phi(x_m)$ is positive and belongs to $\mathbb{R}$. This completes the proof of the lemma. $\square$

Theorem 5.37 (Borel-Lebesgue Theorem) *Let $\mathcal{X}$ be a metric space. Then, the following are equivalent.*

- (i) $\mathcal{X}$ is compact.
- (ii) $\mathcal{X}$ is countably compact.
- (iii) $\mathcal{X}$ is sequentially compact.
- (iv) $\mathcal{X}$ has the Bolzano-Weierstrass property.

Proof (i) $\Rightarrow$ (ii). This follows directly from Definitions 5.1 and 5.21.

(ii) $\Rightarrow$ (iii). Clearly, $\mathcal{X}$ is first countable. By Proposition 5.28, $\mathcal{X}$ is sequentially compact.

(iii) $\Rightarrow$ (i). We will distinguish three exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X}$ is a singleton set; Case 3: $\mathcal{X}$ contains at least two distinct points. Case 1: $\mathcal{X} = \emptyset$. Clearly, $\mathcal{X}$ is compact. Case 2: $\mathcal{X}$ is a singleton set. Clearly, $\mathcal{X}$ is compact. Case 3: $\mathcal{X}$ contains at least two distinct points. Fix any open covering $(O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}$ of $\mathcal{X}$. Let ϵ be the Lebesgue number of the covering, which is a positive real number by Lemma 5.36. By Proposition 5.34, $\mathcal{X}$ is totally bounded. Then, there exists a finite set $X_N \subseteq \mathcal{X}$ such that $\mathcal{X} = \bigcup_{x \in X_N} \mathcal{B}(x, \epsilon/2)$. $\forall x \in X_N$, $\mathcal{B}(x, \epsilon/2) \subseteq O_{\alpha_x}$ for some $\alpha_x \in \Lambda$. Hence, $\mathcal{X} = \bigcup_{x \in X_N} O_{\alpha_x}$, which is a finite subcovering. Therefore, $\mathcal{X}$ is compact.

(ii) $\Leftrightarrow$ (iv). This is proved in Proposition 5.26.

This completes the proof of the theorem. $\square$

Proposition 5.38 *A metric space $\mathcal{X}$ is compact if, and only if, it is complete and totally bounded.*

Proof “Only if” Fix any Cauchy sequence $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$. By the compactness of $\mathcal{X}$ and Borel-Lebesgue Theorem, there exists a subsequence $(x_{n_i})_{i=1}^\infty$ that converges to $x_0 \in \mathcal{X}$. Then, it is straightforward to show that $\lim_{n \in \mathbb{N}} x_n = x_0$. (Here, the notation $\lim_{n \in \mathbb{N}} x_n$ makes sense since metric spaces are Hausdorff.) Hence, $\mathcal{X}$ is complete. By Borel-Lebesgue Theorem and Proposition 5.34, $\mathcal{X}$ is totally bounded.

“If” Fix any sequence $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$. We will construct subsequences as following. Take $(x_n^{(1)})_{n=1}^\infty = (x_n)_{n=1}^\infty$. $\forall m \in \mathbb{N}$ with $m \geq 2$, since $\mathcal{X}$ is covered by finite many balls of radius $1/m$, we may find a subsequence $(x_n^{(m)})_{n=1}^\infty$ of $(x_n^{(m-1)})_{n=1}^\infty$ such that $(x_n^{(m)})_{n=1}^\infty \subseteq \mathcal{B}(x^{(m)}, 1/m)$ for some $x^{(m)} \in \mathcal{X}$. Now, consider the sequence of diagonal elements $(x_n^{(n)})_{n=1}^\infty \subseteq \mathcal{X}$, which is a subsequence of the original sequence $(x_n)_{n=1}^\infty$. Clearly, this sequence is Cauchy by construction. By the completeness of $\mathcal{X}$, $\lim_{n \in \mathbb{N}} x_n^{(n)} = x_0 \in \mathcal{X}$. Hence, we have shown that $\mathcal{X}$ is sequentially compact. By Borel-Lebesgue Theorem, $\mathcal{X}$ is compact.

This completes the proof of the proposition. $\square$

Proposition 5.39 Let $\mathcal{X} := (X, \rho)$ be a compact metric space, $\mathcal{Y} := (Y, \sigma)$ be a metric space, and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous. Then, f is uniformly continuous.

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{X}$, by the continuity of f , $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$ such that $\sigma(f(x), f(\bar{x})) < \epsilon/2$, $\forall \bar{x} \in \mathcal{B}(x, \delta_x)$. Then, $\mathcal{X} = \bigcup_{x \in \mathcal{X}} \mathcal{B}(x, \delta_x/2)$. By the compactness of $\mathcal{X}$, there exists a finite set $X_N \subseteq \mathcal{X}$ such that $\mathcal{X} = \bigcup_{x \in X_N} \mathcal{B}(x, \delta_x/2)$. Let $\delta = \min\{\inf_{x \in X_N} \delta_x/2, 1\} \in (0, \infty) \subset \mathbb{R}$. $\forall x_1, x_2 \in \mathcal{X}$ with $\rho(x_1, x_2) < \delta$, $\exists x_0 \in X_N$ such that $x_1 \in \mathcal{B}(x_0, \delta_{x_0}/2)$. Then, $x_2 \in \mathcal{B}(x_0, \delta_{x_0})$. This implies that $x_1, x_2 \in \mathcal{B}(x_0, \delta_{x_0})$. Then, $\sigma(f(x_1), f(x_2)) \leq \sigma(f(x_1), f(x_0)) + \sigma(f(x_0), f(x_2)) < \epsilon$. Hence, f is uniformly continuous. This completes the proof of the proposition. $\square$

Proposition 5.40 Let $n \in \mathbb{Z}_+$ and $K \subseteq \mathbb{R}^n$. K is compact if, and only if, K is closed and bounded.

Proof When $n \in \mathbb{N}$, note that $\mathbb{R}^n$ is a complete metric space. Then, the result follows from Propositions 5.38 and 4.39. When $n = 0$, note that $\mathbb{R}^n$ is a singleton set and is compact. Then, the result follows immediately. $\square$

5.5 The Ascoli-Arzelá Theorem

Lemma 5.41 Let $(f_n)_{n=1}^\infty$ be a sequence of functions of a countable set D to a metric space $\mathcal{Y}$ such that $\forall x \in D$, $\overline{(f_n(x))_{n \in \mathbb{N}}}$ is compact. Then, there

exists a subsequence $(f_{n_k})_{k=1}^\infty$ such that $\forall x \in D$, the sequence $(f_{n_k}(x))_{k=1}^\infty$ converges to some point $f_0(x) \in \mathcal{Y}$.

Proof We will distinguish three exhaustive and mutually exclusive cases: Case 1: $D = \emptyset$; Case 2: $D \neq \emptyset$ and is finite; Case 3: D is countably infinite. Case 1: $D = \emptyset$. Then, $(f_n)_{n=1}^\infty$ is the subsequence we seek. Case 2: $D \neq \emptyset$ and is finite. Take $D = \{x_1, \dots, x_m\}$ for some $m \in \mathbb{N}$. Let $\left(f_n^{(0)}\right)_{n=1}^\infty = (f_n)_{n=1}^\infty$. $\forall k = 1, \dots, m$, $\overline{\left(f_n^{(k-1)}(x_k)\right)_{n \in \mathbb{N}}} \subseteq \overline{(f_n(x_k))_{n \in \mathbb{N}}}$ is a closed set and therefore a compact set by Proposition 5.5. By Borel-Lebesgue Theorem, there exists a subsequence $\left(f_n^{(k)}\right)_{n=1}^\infty$ of $\left(f_n^{(k-1)}\right)_{n=1}^\infty$ such that $\left(f_n^{(k)}(x_k)\right)_{n=1}^\infty$ converges. Hence, the sequence $\left(f_n^{(m)}\right)_{n=1}^\infty$ is the subsequence we seek. Case 3: D is countably infinite. Take $D = (x_m)_{m=1}^\infty$. Let $\left(f_n^{(0)}\right)_{n=1}^\infty = (f_n)_{n=1}^\infty$. $\forall k \in \mathbb{N}$, $\overline{\left(f_n^{(k-1)}(x_k)\right)_{n \in \mathbb{N}}} \subseteq \overline{(f_n(x_k))_{n \in \mathbb{N}}}$ is a closed set and therefore a compact set by Proposition 5.5. By Borel-Lebesgue Theorem, there exists a subsequence $\left(f_n^{(k)}\right)_{n=1}^\infty$ of $\left(f_n^{(k-1)}\right)_{n=1}^\infty$ such that $\left(f_n^{(k)}(x_k)\right)_{n=1}^\infty$ converges. Now, the diagonal sequence $\left(f_k^{(k)}\right)_{k=1}^\infty$ is the subsequence we seek. This completes the proof of the lemma. $\square$

Lemma 5.42 Let $\mathcal{X} := (X, \mathcal{O})$ be a compact topological space, $\mathcal{Y} := (Y, \rho)$ be a metric space, $\mathcal{F}$ be an equicontinuous family of functions of $\mathcal{X}$ to $\mathcal{Y}$, and $(f_n)_{n=1}^\infty \subseteq \mathcal{F}$ be such that $\lim_{n \in \mathbb{N}} f_n(x) = f_0(x)$, $\forall x \in \mathcal{X}$, for some $f_0 : \mathcal{X} \rightarrow \mathcal{Y}$. Then, f_0 is continuous and $(f_n)_{n=1}^\infty$ converges to f_0 uniformly.

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x_0 \in \mathcal{X}$, by the equicontinuity of $\mathcal{F}$, $\exists U_{x_0, \epsilon} \in \mathcal{O}$ with $x_0 \in U_{x_0, \epsilon}$ such that $\forall x \in U_{x_0, \epsilon}$, $\forall f \in \mathcal{F}$, $\rho(f(x), f(x_0)) < \epsilon$.

By $\lim_{n \in \mathbb{N}} f_n(x_0) = f_0(x_0)$, $\exists N_{x_0, \epsilon} \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq N_{x_0, \epsilon}$, we have $\rho(f_n(x_0), f_0(x_0)) < \epsilon$.

$\forall x_0 \in \mathcal{X}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x \in U_{x_0, \epsilon}$, we have, by Propositions 4.30, 3.66, and 3.67,

$$\rho(f_0(x), f_0(x_0)) = \lim_{n \in \mathbb{N}} \rho(f_n(x), f_n(x_0)) \leq \epsilon$$

Hence, f_0 is continuous at x_0 . By the arbitrariness of x_0 and Proposition 3.9, f_0 is continuous.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\mathcal{X} \subseteq \bigcup_{x \in X} U_{x, \epsilon}$. By the compactness of $\mathcal{X}$, there exists a finite set $X_N \subseteq X$ such that $\mathcal{X} \subseteq \bigcup_{x \in X_N} U_{x, \epsilon}$. Let $N := \max\{\sup_{x \in X_N} N_{x, \epsilon}, 1\} \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq N$, $\forall x \in \mathcal{X}$, $\exists x_0 \in X_N$ such that $x \in U_{x_0, \epsilon}$. Then, we have

$$\rho(f_n(x), f_0(x)) \leq \rho(f_n(x), f_n(x_0)) + \rho(f_n(x_0), f_0(x_0)) + \rho(f_0(x_0), f_0(x)) < 3\epsilon$$

Hence, we have $(f_n)_{n=1}^\infty$ converges to f_0 uniformly.

This completes the proof of the lemma. $\square$

Lemma 5.43 *Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $\mathcal{Y} := (Y, \rho)$ be a metric space, $(f_n)_{n=1}^\infty$ be an equicontinuous sequence of functions of $\mathcal{X}$ to $\mathcal{Y}$, and $(f_n(x))_{n=1}^\infty$ converge at every $x \in D$, where $D \subseteq \mathcal{X}$ is dense. Assume that $\forall x \in \mathcal{X}$, $\overline{(f_n(x))_{n=1}^\infty} \subseteq \mathcal{Y}$ is complete. Then, $\exists$ continuous function $f_0 : \mathcal{X} \rightarrow \mathcal{Y}$ such that $\lim_{n \in \mathbb{N}} f_n(x) = f_0(x)$, $\forall x \in \mathcal{X}$.*

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{X}$, since $(f_n)_{n=1}^\infty$ is equicontinuous, then $\exists O \in \mathcal{O}$ with $x \in O$ such that, $\forall \bar{x} \in O$, $\rho(f_n(x), f_n(\bar{x})) < \epsilon$, $\forall n \in \mathbb{N}$.

Since D is dense, then, $\exists x_0 \in D \cap O$. $(f_n(x_0))_{n=1}^\infty \subseteq \mathcal{Y}$ is Cauchy since it is convergent. Then, $\exists N \in \mathbb{N}$, $\forall n, m \geq N$, we have $\rho(f_n(x_0), f_m(x_0)) < \epsilon$. Note that

$$\begin{aligned} \rho(f_n(x), f_m(x)) &\leq \rho(f_n(x), f_n(x_0)) + \rho(f_n(x_0), f_m(x_0)) \\ &\quad + \rho(f_m(x_0), f_m(x)) < 3\epsilon \end{aligned}$$

Therefore, $\overline{(f_n(x))_{n=1}^\infty} \subseteq \overline{(f_n(x_0))_{n=1}^\infty} \subseteq \mathcal{Y}$ is a Cauchy sequence, which converges since $\overline{(f_n(x))_{n=1}^\infty}$ is complete. Then, we may define $f_0 : \mathcal{X} \rightarrow \mathcal{Y}$ by $f_0(x) = \lim_{n \in \mathbb{N}} f_n(x)$.

By Proposition 4.30, 3.66, and 3.67, we have, $\forall \bar{x} \in O$,

$$\rho(f_0(x), f_0(\bar{x})) = \lim_{n \in \mathbb{N}} \rho(f_n(x), f_n(\bar{x})) \leq \epsilon$$

Hence, f_0 is continuous at x . By the arbitrariness of x and Proposition 3.9, f_0 is continuous. This completes the proof of the lemma. $\square$

Theorem 5.44 (Ascoli-Arzelá Theorem) *Let $\mathcal{X} := (X, \mathcal{O})$ be a separable topological space, $\mathcal{Y} := (Y, \rho)$ be a metric space, $\mathcal{F}$ be an equicontinuous family of functions of $\mathcal{X}$ to $\mathcal{Y}$, and $(f_n)_{n=1}^\infty \subseteq \mathcal{F}$ be such that $\overline{(f_n(x))_{n \in \mathbb{N}}} \subseteq \mathcal{Y}$ is compact, $\forall x \in \mathcal{X}$. Then, there exists a subsequence $(f_{n_k})_{k=1}^\infty$ of $(f_n)_{n=1}^\infty$ that converges pointwise to a continuous function $f_0 : \mathcal{X} \rightarrow \mathcal{Y}$ and the convergence is uniform on any compact subset of $\mathcal{X}$.*

Proof Let $D \subseteq \mathcal{X}$ be a countable dense set. By Proposition 5.41, there exists a subsequence $(f_{n_k})_{k=1}^\infty$ of $(f_n)_{n=1}^\infty$ that converges pointwise on D to a function $\bar{f} : D \rightarrow \mathcal{Y}$. Note that $\forall x \in \mathcal{X}$, $\overline{(f_n(x))_{n \in \mathbb{N}}} \subseteq \mathcal{Y}$ is compact implies that $\overline{(f_n(x))_{n \in \mathbb{N}}}$ is complete by Proposition 5.38. Then, by Proposition 5.43, there exists a continuous function $f_0 : \mathcal{X} \rightarrow \mathcal{Y}$ to which $(f_{n_k})_{k=1}^\infty$ converges pointwise. By Proposition 5.42, $(f_{n_k})_{k=1}^\infty$ converges to f_0 uniformly on compact subsets of $\mathcal{X}$. This completes the proof of the theorem. $\square$

5.6 Product Spaces

Lemma 5.45 *Let $\mathcal{A}$ be a collection of subsets of a set X with the finite intersection property. Then, there is a collection $\mathcal{M}$ of subsets in X such*

that $\mathcal{A} \subseteq \mathcal{M}$, $\mathcal{M}$ has the finite intersection property, and $\mathcal{M}$ is maximal with respect to this property, that is $\forall \mathcal{C} \subseteq {}^X 2$ with the finite intersection property and $\mathcal{M} \subseteq \mathcal{C}$, we have $\mathcal{C} = \mathcal{M}$.

Proof Clearly, $X \neq \emptyset$ by the fact that $\mathcal{A}$ has the finite intersection property. Define $M := \{\mathcal{C} \subseteq {}^X 2 \mid \mathcal{A} \subseteq \mathcal{C} \text{ and } \mathcal{C} \text{ has the finite intersection property}\}$. Set containment $\subseteq$ forms an antisymmetric partial ordering on M and $\mathcal{A} \in M \neq \emptyset$. Let $N \subseteq M$ be any nonempty totally ordered subset. Let $\mathcal{B} := \bigcup_{\mathcal{C} \in N} \mathcal{C} \subseteq {}^X 2$. Fix any finite subcollection $\mathcal{B}_N \subseteq \mathcal{B}$. We will distinguish two exhaustive and mutually exclusive cases: $\mathcal{B}_N = \emptyset$; Case 2: $\mathcal{B}_N \neq \emptyset$. Case 1: $\mathcal{B}_N = \emptyset$. Then, $\bar{B} := \bigcap_{B \in \mathcal{B}_N} B = X \neq \emptyset$. Case 2: $\mathcal{B}_N \neq \emptyset$. Take $\mathcal{B}_N = \{B_1, \dots, B_n\}$ for some $n \in \mathbb{N}$. $\forall B_i \in \mathcal{B}_N$, $\exists \mathcal{C}_i \in N$ such that $B_i \in \mathcal{C}_i$. Since N is totally ordered by $\subseteq$, without loss of generality, assume that $\mathcal{C}_1 \subseteq \mathcal{C}_2 \subseteq \dots \subseteq \mathcal{C}_n$. Then, $B_i \in \mathcal{C}_n$, $\forall B_i \in \mathcal{B}_N$. By the definition of the set M , we have $\bigcap_{i=1}^n B_i \neq \emptyset$. In both cases, we have arrived at $\bigcap_{B \in \mathcal{B}_N} B \neq \emptyset$. Hence, $\mathcal{B}$ has finite intersection property. Since $N \neq \emptyset$, then $\mathcal{A} \subseteq \mathcal{B}$. Hence, $\mathcal{B} \in M$ and is an upper bound of N . By Zorn's Lemma, there exists a maximal element $\mathcal{M}$ of M . This completes the proof of the lemma. $\square$

Lemma 5.46 *Let $\mathcal{B}$ be a collection of subsets of a set X that is maximal with respect to the finite intersection property. Then, each intersection of finite number of sets in $\mathcal{B}$ is again in $\mathcal{B}$, and each set that meets every set in $\mathcal{B}$ is itself in $\mathcal{B}$.*

Proof Fix any finite subcollection $\mathcal{B}_N \subseteq \mathcal{B}$. Let $\bar{B} := \bigcap_{B \in \mathcal{B}_N} B$. Then, $\bar{B} \neq \emptyset$. Let $\bar{\mathcal{B}} := \mathcal{B} \cup \{\bar{B}\} \supseteq \mathcal{B}$. It is easy to see that $\bar{\mathcal{B}}$ has the finite intersection property. Then, by maximality of $\mathcal{B}$, we have $\bar{\mathcal{B}} = \mathcal{B}$. Hence, $\bar{B} \in \mathcal{B}$.

Let $C \subseteq X$ be such that $C \cap B \neq \emptyset$, $\forall B \in \mathcal{B}$. Let $\bar{\mathcal{B}} := \mathcal{B} \cup \{C\} \supseteq \mathcal{B}$. Fix any finite subcollection $\mathcal{M} \subseteq \bar{\mathcal{B}}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $C \notin \mathcal{M}$; Case 2: $C \in \mathcal{M}$. Case 1: $C \notin \mathcal{M}$. Then, $\mathcal{M} \subseteq \mathcal{B}$. Then, $\bigcap_{B \in \mathcal{M}} B \neq \emptyset$ since $\mathcal{B}$ has the finite intersection property. Case 2: $C \in \mathcal{M}$. Then, $\bar{\mathcal{M}} := \mathcal{M} \setminus \{C\} \subseteq \mathcal{B}$ and is a finite set. Then, $\bigcap_{B \in \bar{\mathcal{M}}} B \in \mathcal{B}$. Then, $\bigcap_{B \in \mathcal{M}} B = C \cap (\bigcap_{B \in \bar{\mathcal{M}}} B) \neq \emptyset$. Hence, in both cases, we have $\bigcap_{B \in \mathcal{M}} B \neq \emptyset$. This shows that $\bar{\mathcal{B}}$ has the finite intersection property. By the maximality of $\mathcal{B}$, we have $\bar{\mathcal{B}} = \mathcal{B}$. Hence, $C \in \mathcal{B}$.

This completes the proof of the lemma. $\square$

Theorem 5.47 (Tychonoff Theorem) *Let $\mathcal{X}_\alpha := (X_\alpha, \mathcal{O}_\alpha)$ be a compact topological space, $\forall \alpha \in \Lambda$, where Λ is an index set. Then, the product topological space $\mathcal{X} := (X, \mathcal{O}) := \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha$ is compact.*

Proof (Bourbaki) We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\Lambda = \emptyset$; Case 2: $\Lambda \neq \emptyset$. Case 1: $\Lambda = \emptyset$. Then, X is a singleton set, which is clearly compact.

Case 2: $\Lambda \neq \emptyset$. We will prove this case by Proposition 5.3. Note that the a basis for $\mathcal{X}$ is given by $\mathcal{B} := \{\prod_{\alpha \in \Lambda} O_\alpha \subseteq X \mid O_\alpha \in \mathcal{O}_\alpha, \forall \alpha \in \Lambda, \text{ and } O_\alpha = X_\alpha \text{ for all } \alpha\text{'s except finitely many } \alpha\text{'s}\}$. Fix a collection $\mathcal{A}$ of closed sets in $\mathcal{X}$ with the finite intersection property. By Lemma 5.45, let $\mathcal{C}$ be a collection of subsets of $\mathcal{X}$ such that $\mathcal{A} \subseteq \mathcal{C}$ and $\mathcal{C}$ is maximal with respect to the finite intersection property.

$\forall \alpha \in \Lambda$, let $\mathcal{C}_\alpha := \{C_\alpha \subseteq X_\alpha \mid \exists C \in \mathcal{C} \ni \pi_\alpha(C) = C_\alpha\}$. $\forall$ finite subcollection $\mathcal{C}_{\alpha N} \subseteq \mathcal{C}_\alpha$, $\forall C_\alpha \in \mathcal{C}_{\alpha N}$, $\exists C_{C_\alpha} \in \mathcal{C}$ such that $C_\alpha = \pi_\alpha(C_{C_\alpha})$. Since $\mathcal{C}$ has the finite intersection property, then $\bigcap_{C_\alpha \in \mathcal{C}_{\alpha N}} C_{C_\alpha} \neq \emptyset$. Then, by Proposition 2.5, $\bigcap_{C_\alpha \in \mathcal{C}_{\alpha N}} C_\alpha = \bigcap_{C_\alpha \in \mathcal{C}_{\alpha N}} \pi_\alpha(C_{C_\alpha}) \supseteq \pi_\alpha(\bigcap_{C_\alpha \in \mathcal{C}_{\alpha N}} C_{C_\alpha}) \neq \emptyset$. Hence, $\mathcal{C}_\alpha$ has finite intersection property. Let $\bar{\mathcal{C}}_\alpha := \{\bar{C}_\alpha \mid C_\alpha \in \mathcal{C}_\alpha\}$. Then, $\bar{\mathcal{C}}_\alpha$ has finite intersection property. By the compactness of $\mathcal{X}_\alpha$ and Proposition 5.3, $\exists x_\alpha \in \bigcap_{C_\alpha \in \mathcal{C}_\alpha} \bar{C}_\alpha$.

Let $x \in \mathcal{X}$ be given by $\pi_\alpha(x) = x_\alpha$, $\forall \alpha \in \Lambda$. Consider a set S of the form $S = \pi_{\alpha \text{ inv}}(O_\alpha)$ for some $\alpha \in \Lambda$ and for some $O_\alpha \in \mathcal{O}_\alpha$ with $x_\alpha \in O_\alpha$. Then, $S \cap C \neq \emptyset$, $\forall C \in \mathcal{C}$, since $x_\alpha \in \pi_\alpha(C)$. Hence, $S \in \mathcal{C}$ by Lemma 5.46 and the maximality of $\mathcal{C}$. $\forall B \in \mathcal{B}$ with $x \in B$, $\exists n \in \mathbb{N}$, $\exists \alpha_1, \dots, \alpha_n \in \Lambda$, and $\exists O_{\alpha_i} \in \mathcal{O}_{\alpha_i}$, $i = 1, \dots, n$, such that $x_{\alpha_i} \in O_{\alpha_i}$ and $B = \bigcap_{i=1}^n \pi_{\alpha_i \text{ inv}}(O_{\alpha_i})$. Then, $B \in \mathcal{C}$ by Lemma 5.46.

$\forall F \in \mathcal{A}$, F is closed. Then, $F \cap B \neq \emptyset$, $\forall B \in \mathcal{B}$ with $x \in B$, by the finite intersection property of $\mathcal{C}$. By Proposition 3.3, $x \in \bar{F} = F$. Hence, $x \in \bigcap_{F \in \mathcal{A}} F \neq \emptyset$. By Proposition 5.3, $\mathcal{X}$ is compact.

This completes the proof of the theorem. $\square$

Proposition 5.48 *Let $\mathcal{X}_\alpha := (X_\alpha, \mathcal{O}_\alpha)$ be a sequentially compact topological space, $\forall \alpha \in \Lambda$, where Λ is a countable index set. Then, the product topological space $\mathcal{X} := (X, \mathcal{O}) := \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha$ is sequentially compact.*

Proof We will distinguish three exhaustive and mutually exclusive cases: Case 1: $\Lambda = \emptyset$; Case 2: $\Lambda \neq \emptyset$ and is finite; Case 3: Λ is countably infinite. Case 1: $\Lambda = \emptyset$. Then, $X = \{\emptyset\}$ is a singleton set. Clearly, $\mathcal{X}$ is sequentially compact.

Case 2: $\Lambda \neq \emptyset$ and is finite. Without loss of generality, assume that $\Lambda = \{1, \dots, m\}$ for some $m \in \mathbb{N}$. Fix a sequence $(x_n^{(0)})_{n=1}^\infty \subseteq \mathcal{X}$. $\forall k \in \{1, \dots, m\}$, $(\pi_k(x_n^{(k-1)}))_{n=1}^\infty \subseteq \mathcal{X}_k$. By the sequential compactness of $\mathcal{X}_k$, there exists a subsequence $(x_n^{(k)})_{n=1}^\infty$ of $(x_n^{(k-1)})_{n=1}^\infty$ such that $(\pi_k(x_n^{(k)}))_{n=1}^\infty$ converges to $x_k \in \mathcal{X}_k$. Let $x \in \mathcal{X}$ be given by $\pi_k(x) = x_k$, $\forall k \in \{1, \dots, m\}$. The sequence $(x_n^{(m)})_{n=1}^\infty \subseteq \mathcal{X}$ is a subsequence of $(x_n^{(0)})_{n=1}^\infty$. Clearly, $\forall k \in \{1, \dots, m\}$, $(\pi_k(x_n^{(m)}))_{n=1}^\infty$ converges to x_k . By Proposition 3.67, $(x_n^{(m)})_{n=1}^\infty$ converges to x . Hence, $\mathcal{X}$ is sequentially compact.

Case 3: Λ is countably infinite. Without loss of generality, assume that $\Lambda = \mathbb{N}$. Fix a sequence $\left(x_n^{(0)}\right)_{n=1}^{\infty} \subseteq \mathcal{X}$. $\forall k \in \mathbb{N}$, $\left(\pi_k(x_n^{(k-1)})\right)_{n=1}^{\infty} \subseteq \mathcal{X}_k$. By the sequential compactness of $\mathcal{X}_k$, there exists a subsequence $\left(x_n^{(k)}\right)_{n=1}^{\infty}$ of $\left(x_n^{(k-1)}\right)_{n=1}^{\infty}$ such that $\left(\pi_k(x_n^{(k)})\right)_{n=1}^{\infty}$ converges to $x_k \in \mathcal{X}_k$. Let $x \in \mathcal{X}$ be given by $\pi_k(x) = x_k$, $\forall k \in \mathbb{N}$. Now consider the diagonal sequence $\left(x_n^{(n)}\right)_{n=1}^{\infty} \subseteq \mathcal{X}$, which is a subsequence of $\left(x_n^{(0)}\right)_{n=1}^{\infty}$. Clearly, $\forall k \in \mathbb{N}$, $\left(\pi_k(x_n^{(n)})\right)_{n=1}^{\infty}$ converges to x_k . By Proposition 3.67, $\left(x_n^{(n)}\right)_{n=1}^{\infty}$ converges to x . Hence, $\mathcal{X}$ is sequentially compact.

This completes the proof of the proposition. $\square$

5.7 Locally Compact Spaces

5.7.1 Fundamental notion

Definition 5.49 A topological space $\mathcal{X}$ is locally compact if $\forall x \in \mathcal{X}$, $\exists O \in \mathcal{O}$ with $x \in O$ such that $\overline{O}$ is compact.

Clearly, a compact space is locally compact.

Example 5.50 $\mathbb{R}^n$ is a locally compact space but not a compact space, $\forall n \in \mathbb{N}$. $\diamond$

Proposition 5.51 A topological space $\mathcal{X}$ is locally compact if, and only if, the collection $\mathcal{B}_l := \{O \in \mathcal{O} \mid \overline{O} \text{ is compact}\}$ forms a basis for $\mathcal{O}$.

Proof “Only if” Let $\mathcal{X}$ be locally compact. $\forall x \in \mathcal{X}$, $\exists B \in \mathcal{B}_l$ such that $x \in B$. $\forall O \in \mathcal{O}$, $\forall x \in O$, $\exists B \in \mathcal{B}_l$ such that $x \in B$ and $\overline{B}$ is compact. Let $B_1 := O \cap B$. Then, $x \in B_1 \in \mathcal{O}$ and $B_1 \subseteq O$. Note that $\overline{B_1} \subseteq \overline{B}$ and is closed. By Proposition 5.5, $\overline{B_1}$ is compact. Then, $B_1 \in \mathcal{B}_l$. Therefore, $\mathcal{B}_l$ is a basis for $\mathcal{O}$.

“If” This is straightforward.

This completes the proof of the proposition. $\square$

Proposition 5.52 Let $\mathcal{X}$ be a locally compact topological space and $K \subseteq \mathcal{X}$ be compact. Then, $\exists O \in \mathcal{O}$, such that $K \subseteq O \subseteq \overline{O}$ and $\overline{O}$ is compact.

Proof $\forall x \in K$, by the local compactness of $\mathcal{X}$, $\exists O_x \in \mathcal{O}$ such that $x \in O_x \subseteq \overline{O_x}$ and $\overline{O_x}$ is compact. Then, $K \subseteq \bigcup_{x \in K} O_x$. By the compactness of K , there exists a finite set $K_N \subseteq K$ such that $K \subseteq \bigcup_{x \in K_N} O_x =: O \in \mathcal{O}$. By Proposition 3.3, $\overline{O} = \overline{\bigcup_{x \in K_N} O_x} = \bigcup_{x \in K_N} \overline{O_x}$, which is compact. This completes the proof of the proposition. $\square$

Proposition 5.53 Let $\mathcal{X}$ be a locally compact Hausdorff topological space and $\mathcal{B}_l$ be the basis of $\mathcal{X}$ defined in Proposition 5.51. $\forall O \in \mathcal{O}$, $\forall x \in O$, $\exists B \in \mathcal{B}_l$ such that $x \in B \subseteq \overline{B} \subseteq O$ and $\overline{B}$ is compact.

Proof $\forall O \in \mathcal{O}, \forall x \in O$, we will distinguish two exhaustive and mutually exclusive cases: Case 1: $O = \mathcal{X}$; Case 2: $O \subset \mathcal{X}$. Case 1: $O = \mathcal{X}$. The result holds by Definition 5.49. Case 2: $O \subset \mathcal{X}$. $\forall y \in \tilde{O}$, since $\mathcal{X}$ is Hausdorff, then $\exists O_y^{(1)}, O_y^{(2)} \in \mathcal{O}$ such that $x \in O_y^{(1)}, y \in O_y^{(2)}$, and $O_y^{(1)} \cap O_y^{(2)} = \emptyset$. By Proposition 5.51 and Definition 3.17, $\exists B_y \in \mathcal{B}_l$ such that $x \in B_y \subseteq O_y^{(1)}$. Then, $B_y \subseteq \widetilde{O_y^{(2)}}$ and $\overline{B_y} \subseteq \widetilde{O_y^{(2)}}$, since $O_y^{(2)}$ is closed in $\mathcal{X}$. This implies that $y \notin \overline{B_y}$. Note that $\tilde{O}$ is closed in $\mathcal{X}$, then $\tilde{O} \cap \overline{B_y}$ is closed in $\mathcal{X}$. Since $\overline{B_y}$ is compact, then, by Propositions 3.5 and 5.5, $\tilde{O} \cap \overline{B_y}$ is compact. Note that $\bigcap_{y \in \tilde{O}} (\tilde{O} \cap \overline{B_y}) = \emptyset$. By Proposition 5.12, there exists a finite set $D_N \subseteq \tilde{O}$ such that $\bigcap_{y \in D_N} (\tilde{O} \cap \overline{B_y}) = \emptyset$. Clearly, $D_N \neq \emptyset$ since $O \subset \mathcal{X} \neq \emptyset$. Then, we have $x \in \bigcap_{y \in D_N} B_y \subseteq \bigcap_{y \in D_N} \overline{B_y} \subseteq O$. Since $\bigcap_{y \in D_N} B_y \in \mathcal{O}$, then $\exists B \in \mathcal{B}_l$ such that $x \in B \subseteq \bigcap_{y \in D_N} B_y$ and $\overline{B}$ is compact. Note that $\overline{B} \subseteq \bigcap_{y \in D_N} \overline{B_y} \subseteq O$. This completes the proof of the proposition. $\square$

Proposition 5.54 *Let $\mathcal{X}$ be a locally compact Hausdorff topological space, $U \in \mathcal{O}$, and $K \subseteq U$ be compact. Then, $\exists V \in \mathcal{O}$ such that $K \subseteq V \subseteq \overline{V} \subseteq U$ and $\overline{V}$ is compact.*

Proof $\forall x \in K$, by Proposition 5.53, $\exists V_x \in \mathcal{O}$ such that $x \in V_x \subseteq \overline{V_x} \subseteq U$ and $\overline{V_x}$ is compact. Then, $K \subseteq \bigcup_{x \in K} V_x$. By the compactness of K , there exists a finite set $K_N \subseteq K$ such that $K \subseteq \bigcup_{x \in K_N} V_x =: V \in \mathcal{O}$. By Proposition 3.3, $\overline{V} = \bigcup_{x \in K_N} \overline{V_x} \subseteq U$ and is compact. This completes the proof of the proposition. $\square$

Proposition 5.55 *Let $\mathcal{X}$ be a locally compact space and $F \subseteq \mathcal{X}$. Then, F is closed if, and only if, for any closed and compact set $K \subseteq \mathcal{X}$, we have $F \cap K$ is closed.*

Proof “Only if” Let F be closed. For any closed and compact $K \subseteq \mathcal{X}$, $F \cap K$ is closed.

“If” $\forall x \in \overline{F}$. By local compactness of $\mathcal{X}$, $\exists O \in \mathcal{O}$ such that $x \in O \subseteq \overline{O}$ and $\overline{O}$ is compact. By the assumption, $\overline{O} \cap F$ is closed. $\forall U \in \mathcal{O}$ with $x \in U$, we have $x \in O \cap U \in \mathcal{O}$. By Proposition 3.3, $(O \cap U) \cap F \neq \emptyset$, since $x \in \overline{F}$. Then, we have $U \cap (\overline{O} \cap F) \neq \emptyset$. Hence, we have $x \in \overline{O} \cap F = \overline{O} \cap F \subseteq F$, by Proposition 3.3. Hence, $\overline{F} \subseteq F$ and F is closed.

This completes the proof of the proposition. $\square$

Proposition 5.56 *Let $\mathcal{X}$ be a locally compact Hausdorff topological space and $(O_n)_{n=1}^\infty \subseteq \mathcal{O}$ be a sequence of open dense subsets in $\mathcal{X}$. Then, $\bigcap_{n=1}^\infty O_n$ is dense. Therefore, $\mathcal{X}$ is second category everywhere.*

Proof $\forall U \in \mathcal{O}$ with $U \neq \emptyset$, since O_1 is dense, then $\exists x_1 \in U \cap O_1$. By Proposition 5.53, $\exists V_1 \in \mathcal{O}$ such that $x_1 \in V_1 \subseteq \overline{V_1} \subseteq U \cap O_1$ and $\overline{V_1}$ is

compact. $\forall n \in \mathbb{N}$ with $n \geq 2$, since O_n is dense, then $\exists x_n \in V_{n-1} \cap O_n$. By Proposition 5.53, $\exists V_n \in \mathcal{O}$ such that $x_n \in V_n \subseteq \overline{V_n} \subseteq V_{n-1} \cap O_n$ and $\overline{V_n}$ is compact. Hence, $(\overline{V_n})_{n=1}^\infty$ is a sequence of nonempty closed sets which is nonincreasing (that is $\overline{V_{n+1}} \subseteq \overline{V_n}$, $\forall n \in \mathbb{N}$). Clearly, this sequence have the finite intersection property. By the compactness of $\overline{V_1}$ and Proposition 5.3, $\bigcap_{n=1}^\infty \overline{V_n} \neq \emptyset$. Clearly, $\bigcap_{n=1}^\infty \overline{V_n} \subseteq U \cap (\bigcap_{n=1}^\infty O_n) \neq \emptyset$. Therefore, $\bigcap_{n=1}^\infty O_n$ is dense in $\mathcal{X}$, by the arbitrariness of U .

By Proposition 3.38, $\mathcal{X}$ is second category everywhere. This completes the proof of the proposition. $\square$

Proposition 5.57 *Let $\mathcal{X} := (X, \mathcal{O})$ be a Hausdorff space, $Y \subseteq X$ be a dense subset, and $\mathcal{O}_Y$ be the subset topology on Y . Assume that $\mathcal{Y} := (Y, \mathcal{O}_Y)$ is locally compact. Then, Y is open in $\mathcal{X}$.*

Proof $\forall y \in Y$, by the local compactness of $\mathcal{Y}$, $\exists U_y \in \mathcal{O}_Y$ such that $y \in U_y \subseteq U_{yc} \subseteq Y$, where U_{yc} is the closure of U_y in $\mathcal{Y}$, and U_{yc} is compact. Then, U_{yc} is compact in $\mathcal{X}$ and is closed, by Proposition 5.5. Then, $U_{yc} \supseteq \overline{U_y}$. On the other hand, by Proposition 3.5, $U_{yc} = \overline{U_y} \cap Y \subseteq \overline{U_y}$. Hence, we have $U_{yc} = \overline{U_y}$. Since $U_y \in \mathcal{O}_Y$, $\exists O \in \mathcal{O}$ such that $U_y = O \cap Y$. By Proposition 3.37, we have $U_{yc} = \overline{O \cap Y} = \overline{O}$. Hence, we have $y \in U_y \subseteq O \subseteq \overline{O} = U_{yc} \subseteq Y$. Hence, Y is open in $\mathcal{X}$. This completes the proof of the proposition. $\square$

Lemma 5.58 *Let $\mathcal{X} := (X, \mathcal{O})$ be a locally compact Hausdorff topological space, $Y \in \mathcal{O}$, and $\mathcal{O}_Y$ be the subset topology on Y . Then, $\mathcal{Y} := (Y, \mathcal{O}_Y)$ is locally compact.*

Proof $\forall x \in Y$, by Proposition 5.53, $\exists U \in \mathcal{O}$ such that $x \in U \subseteq \overline{U} \subseteq Y$ and $\overline{U}$ is compact. Then, $U \in \mathcal{O}_Y$ and the closure of U in $\mathcal{Y}$ is $\overline{U}$, by Proposition 3.5. Then, $\overline{U}$ is compact in $\mathcal{Y}$. Hence, $\mathcal{Y}$ is locally compact. This completes the proof of the lemma. $\square$

Lemma 5.59 *Let $\mathcal{X} := (X, \mathcal{O})$ be a locally compact Hausdorff topological space, $Y \subseteq \mathcal{X}$ be closed, and $\mathcal{O}_Y$ be the subset topology on Y . Then, $\mathcal{Y} := (Y, \mathcal{O}_Y)$ is locally compact.*

Proof $\forall x \in Y$, by the local compactness of $\mathcal{X}$, $\exists U \in \mathcal{O}$ such that $x \in U \subseteq \overline{U}$ and $\overline{U}$ is compact. Then, $\bar{U} := U \cap Y \in \mathcal{O}_Y$. Let $\bar{U}_c$ be the closure of $\bar{U}$ in $\mathcal{Y}$. Then, $\bar{U}_c \subseteq \overline{U} \cap Y \subseteq \overline{U}$. Note that $\bar{U}_c$ is a closed set in $\mathcal{Y}$ and therefore a closed set in $\mathcal{X}$ since Y is closed in $\mathcal{X}$ and Proposition 3.5. Then, $\bar{U}_c$ is compact by Proposition 5.5. Thus, we have $x \in \bar{U} \subseteq \bar{U}_c \subseteq Y$ and $\bar{U}_c$ is compact. Hence, $\mathcal{Y}$ is locally compact. This completes the proof of the lemma. $\square$

Proposition 5.60 *Let $\mathcal{X} := (X, \mathcal{O})$ be a locally compact Hausdorff topological space, $Y \subseteq X$, and $\mathcal{O}_Y$ be the subset topology on Y . $\mathcal{Y} := (Y, \mathcal{O}_Y)$ is locally compact if, and only if, Y is relatively open in $\overline{Y}$.*

Proof “Only if” Let $\mathcal{O}_{\overline{Y}}$ be the subset topology on $\overline{Y}$. Then, $(\overline{Y}, \mathcal{O}_{\overline{Y}})$ is Hausdorff. By Proposition 5.57, Y is relative open in $\mathcal{O}_{\overline{Y}}$.

“If” Since $\overline{Y}$ is closed in $\mathcal{X}$, then, by Lemma 5.59, $(\overline{Y}, \mathcal{O}_{\overline{Y}})$ is locally compact. Clearly, $(\overline{Y}, \mathcal{O}_{\overline{Y}})$ is Hausdorff. In $(\overline{Y}, \mathcal{O}_{\overline{Y}})$, by Lemma 5.58, $\mathcal{Y}$ is locally compact.

This completes the proof of the proposition. $\square$

5.7.2 Partition of unity

Definition 5.61 Let $\mathcal{X}$ be a topological space. For a function $f : \mathcal{X} \rightarrow \mathbb{R}$, the support of f is the set $\text{supp } f := \overline{\{x \in \mathcal{X} \mid f(x) \neq 0\}}$. A collection of real-valued functions $(\phi_\alpha)_{\alpha \in \Gamma}$ on $\mathcal{X}$ is said to be subordinate to a collection of open sets $(O_\lambda)_{\lambda \in \Lambda}$, if $\forall \alpha \in \Gamma, \exists \lambda \in \Lambda$ such that $\text{supp } \phi_\alpha \subseteq O_\lambda$.

Proposition 5.62 Let $\mathcal{X}$ be a locally compact Hausdorff topological space, $U \in \mathcal{O}$, and $K \subseteq U$ be compact. Then, there exists a continuous function $f : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R}$ such that $f|_K = 1$ and $f|_{\tilde{U}} = 0$. Furthermore, $\text{supp } f \subseteq U$ is compact.

Proof Since the set $Q := \mathbb{Q} \cap [0, 1]$ is countable, then, by recursively applying Proposition 5.54, we may find $(O_r)_{r \in Q} \subseteq \mathcal{O}$ such that the following two properties are satisfied:

1. $\forall r \in Q, K \subseteq O_r \subseteq \overline{O_r} \subseteq U$ and $\overline{O_r}$ is compact;
2. $\forall r, s \in Q$ with $r < s$, $\overline{O_r} \subseteq O_s$.

Define the real-valued function $\bar{f} : X \rightarrow \mathbb{R}$ by

$$\bar{f}(x) = \inf(\{r \in Q \mid x \in O_r\} \cup \{1\})$$

Clearly, $\bar{f} : X \rightarrow [0, 1]$, $\bar{f}(x) = 0, \forall x \in O_0$, and $\bar{f}(x) = 1, \forall x \in \widetilde{O_1}$. By 1, we have $K \subseteq O_0$ and $O_1 \subseteq U$. Next, we will show that $\bar{f}$ is continuous. $\forall x_0 \in X$, we will show that $\bar{f}$ is continuous at x_0 . Let $a_0 = \bar{f}(x_0) \in [0, 1]$. $\forall B \subseteq \mathbb{R}$ with B being open and $a_0 \in B$, $\exists a_1, a_2, a_3, a_4 \in \mathbb{Q}$ such that $a_1 < a_2 < a_0 < a_3 < a_4$ and $(a_1, a_4) \subseteq B$. Let $\bar{a}_2 = \max\{a_2, 0\}$ and $\bar{a}_3 = \min\{a_3, 1\}$. Then, we must have $a_1 < \bar{a}_2 \leq a_0 \leq \bar{a}_3 < a_4$ and $\bar{a}_2, \bar{a}_3 \in Q$. We will distinguish three exhaustive and mutually exclusive cases: Case 1: $a_0 \in (0, 1)$; Case 2: $a_0 = 0$; Case 3: $a_0 = 1$.

Case 1: $a_0 \in (0, 1)$. Then, we must have $a_1 < \bar{a}_2 < a_0 < \bar{a}_3 < a_4$. Let $V = \widetilde{\overline{O_{\bar{a}_2}}} \cap O_{\bar{a}_3} \in \mathcal{O}$. $\forall x \in V$, we have $x \in O_{\bar{a}_3}$ and $\bar{f}(x) \leq \bar{a}_3$. Also, $x \in \widetilde{\overline{O_{\bar{a}_2}}}$ implies that $\bar{f}(x) \geq \bar{a}_2$. Hence, $\bar{f}(V) \subseteq [\bar{a}_2, \bar{a}_3] \subset (a_1, a_4) \subseteq B$. $\bar{f}(x_0) = a_0 < \bar{a}_3$ implies that $x_0 \in O_{\bar{a}_3}$. $\bar{f}(x_0) = a_0 > \bar{a}_2$ implies that $\exists \tilde{a}_2 \in (\bar{a}_2, a_0) \cap Q$ such that $x_0 \in \widetilde{\overline{O_{\tilde{a}_2}}} \subseteq \widetilde{\overline{O_{\bar{a}_2}}}$. Therefore $x_0 \in V$. This shows that $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $\bar{f}(V) \subseteq B$.

Case 2: $a_0 = 0$. Then, we must have $a_1 < 0 = a_0 < \bar{a}_3 < a_4$. Take $V = O_{\bar{a}_3} \in \mathcal{O}$. We must have $x_0 \in V$. $\forall x \in V, 0 \leq \bar{f}(x) \leq \bar{a}_3$. Hence,

$\bar{f}(V) \subseteq [0, \bar{a}_3] \subset (a_1, a_4) \subseteq B$. Hence, $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $\bar{f}(V) \subseteq B$.

Case 3: $a_0 = 1$. Then, we must have $a_1 < \bar{a}_2 < a_0 = 1 < a_4$. Take $V = \overline{O_{\bar{a}_2}} \in \mathcal{O}$. Since $\bar{f}(x_0) = a_0 = 1$, then $x_0 \in \overline{O_{\frac{1+\bar{a}_2}{2}}} \subseteq \overline{O_{\bar{a}_2}} = V$. $\forall x \in V$, $\bar{f}(x) \geq \bar{a}_2$. Hence, $\bar{f}(V) \subseteq [\bar{a}_2, 1] \subset (a_1, a_4) \subseteq B$. Hence, $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $\bar{f}(V) \subseteq B$.

Therefore, in all cases, $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $\bar{f}(V) \subseteq B$. Then, $\bar{f}$ is continuous at x_0 . By the arbitrariness of x_0 and Proposition 3.9, $\bar{f}$ is continuous. Define $f : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R}$ by $f(x) = 1 - \bar{f}(x)$, $\forall x \in \mathcal{X}$. Clearly, f is continuous by Proposition 3.12, $f|_K = 1$, and $f|_{\bar{U}} = 0$. Note that $\text{supp } f = \overline{\{x \in \mathcal{X} \mid \bar{f}(x) < 1\}} \subseteq \overline{O_1} \subseteq U$ and is compact by Proposition 5.5. This completes the proof of the proposition. $\square$

Theorem 5.63 (Partition of Unity) *Let $\mathcal{X}$ be a locally compact Hausdorff topological space, $K \subseteq \mathcal{X}$ be compact, and $(O_\lambda)_{\lambda \in \Lambda} \subseteq \mathcal{O}$ be an open covering of K , where Λ is an index set. Let $\mathcal{F}$ be a collection of continuous real-valued functions on $\mathcal{X}$ such that*

- (i) *the constantly zero function is in $\mathcal{F}$;*
- (ii) *$\forall f, g \in \mathcal{F}$, $f + g \in \mathcal{F}$;*
- (iii) *$\forall f, g \in \mathcal{F}$ with $\text{supp } f \subseteq M := \{x \in \mathcal{X} \mid g(x) \neq 0\}$, define $f/g : \mathcal{X} \rightarrow \mathbb{R}$ by*

$$(f/g)(x) = \begin{cases} f(x)/g(x) & \forall x \in \text{supp } f \\ 0 & \forall x \in \text{supp } f \end{cases}$$

then $f/g \in \mathcal{F}$;

- (iv) *$\forall O \in \mathcal{O}$, $\forall x_0 \in O$, $\exists f \in \mathcal{F}$ such that $f(x_0) = 1$, $f|_{\bar{O}} = 0$, and $f : \mathcal{X} \rightarrow [0, 1]$.*

Then, there exists a finite collection $\Phi \subseteq \mathcal{F}$ of continuous nonnegative real-valued functions on $\mathcal{X}$, which is subordinate to $(O_\lambda)_{\lambda \in \Lambda}$, such that $\left(\sum_{\phi \in \Phi} \phi\right)|_K = 1$, $\sum_{\phi \in \Phi} \phi(x) \in [0, 1] \subset \mathbb{R}$, $\forall x \in \mathcal{X}$, and $\text{supp } \phi$ is compact, $\forall \phi \in \Phi$.

Remark 5.64 Note that in (iii) above f/g is always continuous since $(f/g)|_{\widetilde{\text{supp } f}} = 0$ is continuous, $(f/g)|_M$ is continuous, $\text{supp } f$ and M are open, $\text{supp } f \cup M = \mathcal{X}$, and Theorem 3.11 implies the result. $\diamond$

Proof Let $O := \bigcup_{\lambda \in \Lambda} O_\lambda \supseteq K$. By Proposition 5.54, there exists $U \in \mathcal{O}$ such that $K \subseteq U \subseteq \bar{U} \subseteq O$ and $\bar{U}$ is compact. $\forall \lambda \in \Lambda$, let $U_\lambda := U \cap O_\lambda \in \mathcal{O}$. Then, $K \subseteq \bigcup_{\lambda \in \Lambda} U_\lambda = U$.

$\forall x_0 \in K, \exists \lambda_0 \in \Lambda$ such that $x_0 \in U_{\lambda_0}$. By Proposition 5.53, $\exists \bar{U}_{\lambda_0} \in \mathcal{O}$ such that $x_0 \in \bar{U}_{\lambda_0} \subseteq \overline{U_{\lambda_0}} \subseteq U_{\lambda_0}$ and $\bar{U}_{\lambda_0}$ is compact. By (iv), there exists a continuous function $f_{x_0} : \mathcal{X} \rightarrow [0, 1]$ in $\mathcal{F}$ such that $f_{x_0}(x_0) = 1$ and $f_{x_0}|_{\bar{U}_{\lambda_0}} = 0$. Let $V_{x_0} := \{x \in \mathcal{X} \mid f_{x_0}(x) > 0\}$. Then, $x_0 \in V_{x_0} \in \mathcal{O}$ since f_{x_0} is continuous. Note that $V_{x_0} \subseteq \bar{U}_{\lambda_0}$, which implies that $\text{supp } f_{x_0} = \overline{V_{x_0}} \subseteq \bar{U}_{\lambda_0} \subseteq U_{\lambda_0} \subseteq O_{\lambda_0}$, which is compact by Proposition 5.5. Hence, f_{x_0} has compact support.

$\forall x_0 \in \bar{U} \setminus K, \tilde{K} \in \mathcal{O}$, by Proposition 5.5. By (iv), there exists a continuous function $g_{x_0} : \mathcal{X} \rightarrow [0, 1]$ in $\mathcal{F}$ such that $g_{x_0}(x_0) = 1$ and $g_{x_0}|_K = 0$. Let $W_{x_0} := \{x \in \mathcal{X} \mid g_{x_0}(x) > 0\}$. Then, $x_0 \in W_{x_0} \in \mathcal{O}$ since g_{x_0} is continuous. Clearly, $W_{x_0} \subseteq \tilde{K}$. Note that $\bar{U} \subseteq (\bigcup_{x \in K} V_x) \cup (\bigcup_{x \in \bar{U} \setminus K} W_x)$. By the compactness of $\bar{U}$, there exist finite sets $K_N \subseteq K$ and $U_N \subseteq \bar{U} \setminus K$ such that $\bar{U} \subseteq (\bigcup_{x \in K_N} V_x) \cup (\bigcup_{x \in U_N} W_x)$. By the construction of W_x 's, $K \subseteq \bigcup_{x \in K_N} V_x$. Let $f := \sum_{x \in K_N} f_x \in \mathcal{F}$ and $g := \sum_{x \in U_N} g_x \in \mathcal{F}$ (here $f, g \in \mathcal{F}$ by (i) and (ii)). Then, f is nonnegative, $f(x) > 0, \forall x \in K, g$ is nonnegative, $f(x) + g(x) > 0, \forall x \in \bar{U}$.

Define $\Phi := \{\phi : \mathcal{X} \rightarrow [0, 1] \mid \phi = f_x/(f + g), x \in K_N\}$, which is a finite set. Note that $\forall x \in K_N, \text{supp } f_x \subseteq U_{\lambda_0} \subseteq U \subseteq \bar{U} \subseteq \{x \in \mathcal{X} \mid (f + g)(x) \neq 0\}$, for some $\lambda_0 \in \Lambda$. Then, $f_x/(f + g) \in \mathcal{F}$. Hence, $\Phi \subseteq \mathcal{F}$. $\forall \phi \in \Phi, \exists x \in K_N$ such that $\phi = f_x/(f + g)$. Then, ϕ is continuous by the fact that $\phi \in \mathcal{F}$. $\text{supp } \phi = \text{supp } f_x$ is compact and $\text{supp } \phi \subseteq O_\lambda$, for some $\lambda \in \Lambda$. Hence, Φ is subordinate to $(O_\lambda)_{\lambda \in \Lambda}$. Clearly, ϕ is nonnegative, $\forall \phi \in \Phi, \forall x \in K$, we have

$$\sum_{\phi \in \Phi} \phi(x) = \frac{f(x)}{f(x) + g(x)} = \frac{f(x)}{f(x)} = 1$$

$\forall x \in \mathcal{X}$, we have

$$0 \leq \sum_{\phi \in \Phi} \phi(x) = (f/(f + g))(x) \leq 1$$

This completes the proof of the theorem. $\square$

Corollary 5.65 *Let $\mathcal{X}$ be a locally compact Hausdorff topological space, $K \subseteq \mathcal{X}$ be compact, and $(O_\lambda)_{\lambda \in \Lambda} \subseteq \mathcal{O}$ be an open covering of K , where Λ is an index set. Then, there exists a finite collection Φ of continuous nonnegative real-valued functions on $\mathcal{X}$, which is subordinate to $(O_\lambda)_{\lambda \in \Lambda}$, such that $\left(\sum_{\phi \in \Phi} \phi\right)|_K = 1, \sum_{\phi \in \Phi} \phi(x) \in [0, 1] \subset \mathbb{R}, \forall x \in \mathcal{X}$, and $\text{supp } \phi$ is compact, $\forall \phi \in \Phi$.*

Proof Let $\mathcal{F}$ be the collection of continuous real-valued functions on $\mathcal{X}$. Clearly, $\mathcal{F}$ satisfies (i) – (iii) in Theorem 5.63. Note that $\{x_0\}$ is compact, by Proposition 5.62, (iv) of Theorem 5.63 is also satisfied by $\mathcal{F}$. Then, the result follows. $\square$

5.7.3 The Alexandroff one-point compactification

Theorem 5.66 (Alexandroff One-Point Compactification) *Let $\mathcal{X} := (X, \mathcal{O})$ be a locally compact Hausdorff topological space. The Alexandroff one-point compactification of $\mathcal{X}$ is the set $X_c := X \cup \{\omega\}$ with the topology $\mathcal{O}_c := \{O_c \subseteq X_c \mid O_c \in \mathcal{O} \text{ or } X_c \setminus O_c \text{ is compact in } \mathcal{X}\}$. Then, $\mathcal{X}_c := (X_c, \mathcal{O}_c)$ is a compact Hausdorff space and the identity map $\text{id} : X \rightarrow X_c \setminus \{\omega\}$ is a homeomorphism. The element ω is called the point at infinity in $\mathcal{X}_c$.*

Proof We first show that $\mathcal{O}_c$ is a topology on X_c . (i) $\emptyset \in \mathcal{O}$, then $\emptyset \in \mathcal{O}_c$; $X_c \setminus X_c = \emptyset$ is compact in $\mathcal{X}$, then $X_c \in \mathcal{O}_c$. (ii) $\forall O_{c1}, O_{c2} \in \mathcal{O}_c$, we will distinguish four exhaustive and mutually exclusive cases: Case 1: $O_{c1}, O_{c2} \in \mathcal{O}$; Case 2: $X_c \setminus O_{c1}$ and $X_c \setminus O_{c2}$ are compact in $\mathcal{X}$; Case 3: $X_c \setminus O_{c1}$ is compact in $\mathcal{X}$ and $O_{c2} \in \mathcal{O}$; Case 4: $O_{c1} \in \mathcal{O}$ and $X_c \setminus O_{c2}$ is compact in $\mathcal{X}$. Case 1: $O_{c1}, O_{c2} \in \mathcal{O}$. Then, $O_{c1} \cap O_{c2} \in \mathcal{O}$ and hence $O_{c1} \cap O_{c2} \in \mathcal{O}_c$. Case 2: $X_c \setminus O_{c1}$ and $X_c \setminus O_{c2}$ are compact in $\mathcal{X}$. Then, $X_c \setminus (O_{c1} \cap O_{c2}) = (X_c \setminus O_{c1}) \cup (X_c \setminus O_{c2})$, which is compact in $\mathcal{X}$. This implies that $O_{c1} \cap O_{c2} \in \mathcal{O}_c$. Case 3: $X_c \setminus O_{c1}$ is compact in $\mathcal{X}$ and $O_{c2} \in \mathcal{O}$. Let $\bar{O}_{c1} := O_{c1} \setminus \{\omega\}$, then $\bar{O}_{c1} = X \setminus (X_c \setminus O_{c1})$. Since $X_c \setminus O_{c1}$ is compact in $\mathcal{X}$ and $\mathcal{X}$ is Hausdorff, then, by Proposition 5.5, $X_c \setminus O_{c1}$ is closed in $\mathcal{X}$. Then, $\bar{O}_{c1} \in \mathcal{O}$. Note that $O_{c1} \cap O_{c2} = \bar{O}_{c1} \cap O_{c2} \in \mathcal{O}$. Then, $O_{c1} \cap O_{c2} \in \mathcal{O}_c$. Case 4: $O_{c1} \in \mathcal{O}$ and $X_c \setminus O_{c2}$ is compact in $\mathcal{X}$. By an argument that is similar to Case 3, we have $O_{c1} \cap O_{c2} \in \mathcal{O}_c$. Hence, in all four cases, we have $O_{c1} \cap O_{c2} \in \mathcal{O}_c$.

(iii) $\forall (O_{c\lambda})_{\lambda \in \Lambda} \subseteq \mathcal{O}_c$, where Λ is an index set, we will distinguish two exhaustive and mutually exclusive cases: Case A: $(O_{c\lambda})_{\lambda \in \Lambda} \subseteq \mathcal{O}$; Case B: $\exists \lambda_0 \in \Lambda$ such that $\omega \in O_{c\lambda_0}$. Case A: $(O_{c\lambda})_{\lambda \in \Lambda} \subseteq \mathcal{O}$. Then, $\bigcup_{\lambda \in \Lambda} O_{c\lambda} \in \mathcal{O}$ and hence $\bigcup_{\lambda \in \Lambda} O_{c\lambda} \in \mathcal{O}_c$. Case B: $\exists \lambda_0 \in \Lambda$ such that $\omega \in O_{c\lambda_0}$. Then, $X_c \setminus O_{c\lambda_0}$ is compact in $\mathcal{X}$. We may partition Λ into two disjoint set Λ_1 and Λ_2 such that $\Lambda = \Lambda_1 \cup \Lambda_2$, $\Lambda_1 \cap \Lambda_2 = \emptyset$, $\forall \lambda \in \Lambda_1$, $O_{c\lambda} \in \mathcal{O}$, $\forall \lambda \in \Lambda_2$, $X_c \setminus O_{c\lambda}$ is compact in $\mathcal{X}$. Note that

$$\begin{aligned} X_c \setminus \left(\bigcup_{\lambda \in \Lambda} O_{c\lambda} \right) &= \left(\bigcap_{\lambda \in \Lambda_1} (X_c \setminus O_{c\lambda}) \right) \cap \left(\bigcap_{\lambda \in \Lambda_2} (X_c \setminus O_{c\lambda}) \right) \\ &= \left(\bigcap_{\lambda \in \Lambda_1} (X \setminus O_{c\lambda}) \right) \cap \left(\bigcap_{\lambda \in \Lambda_2} (X_c \setminus O_{c\lambda}) \right) \cap (X_c \setminus O_{c\lambda_0}) \end{aligned}$$

$\forall \lambda \in \Lambda_1$, $X \setminus O_{c\lambda}$ is a closed set in $\mathcal{X}$. $\forall \lambda \in \Lambda_2$, $X_c \setminus O_{c\lambda}$ is compact in $\mathcal{X}$ and therefore closed in $\mathcal{X}$ by Proposition 5.5. Hence, $X_c \setminus \left(\bigcup_{\lambda \in \Lambda} O_{c\lambda} \right)$ is a closed subset of $X_c \setminus O_{c\lambda_0}$ and hence compact in $\mathcal{X}$ by Proposition 5.5. Then, $\bigcup_{\lambda \in \Lambda} O_{c\lambda} \in \mathcal{O}_c$. Hence, in both cases, we have $\bigcup_{\lambda \in \Lambda} O_{c\lambda} \in \mathcal{O}_c$. Summarizing the above, $\mathcal{O}_c$ is a topology on X_c .

Next, we show that $\mathcal{X}_c$ is compact. Fix an open covering $(O_{c\lambda})_{\lambda \in \Lambda} \subseteq \mathcal{O}_c$ of $\mathcal{X}_c$. We may partition Λ into two disjoint set Λ_1 and Λ_2 such that $\Lambda = \Lambda_1 \cup \Lambda_2$, $\Lambda_1 \cap \Lambda_2 = \emptyset$, $\forall \lambda \in \Lambda_1$, $O_{c\lambda} \in \mathcal{O}$, $\forall \lambda \in \Lambda_2$, $X_c \setminus O_{c\lambda}$ is compact

in $\mathcal{X}$. Since $\omega \in X_c$, then $\exists \lambda_0 \in \Lambda_2$ such that $\omega \in O_{c\lambda_0}$. Then, $X_c \setminus O_{c\lambda_0}$ is compact in $\mathcal{X}$. $\forall \lambda \in \Lambda_2$, let $\bar{O}_{c\lambda} := O_{c\lambda} \setminus \{\omega\}$, then $\bar{O}_{c\lambda} = X \setminus (X_c \setminus O_{c\lambda})$. By Proposition 5.5 and the compactness of $X_c \setminus O_{c\lambda}$, $\bar{O}_{c\lambda} \in \mathcal{O}$. Note that

$$X_c = \left(\bigcup_{\lambda \in \Lambda_1} O_{c\lambda} \right) \cup \left(\bigcup_{\lambda \in \Lambda_2} O_{c\lambda} \right) \cup O_{c\lambda_0} = O_{c\lambda_0} \cup (X_c \setminus O_{c\lambda_0})$$

Then, $X_c \setminus O_{c\lambda_0} \subseteq \left(\bigcup_{\lambda \in \Lambda_1} O_{c\lambda} \right) \cup \left(\bigcup_{\lambda \in \Lambda_2} \bar{O}_{c\lambda} \right)$. By the compactness of $X_c \setminus O_{c\lambda_0}$, there exist finite sets $\Lambda_{N1} \subseteq \Lambda_1$ and $\Lambda_{N2} \subseteq \Lambda_2$ such that $X_c \setminus O_{c\lambda_0} \subseteq \left(\bigcup_{\lambda \in \Lambda_{N1}} O_{c\lambda} \right) \cup \left(\bigcup_{\lambda \in \Lambda_{N2}} \bar{O}_{c\lambda} \right)$. Then, $X_c = \left(\bigcup_{\lambda \in \Lambda_{N1}} O_{c\lambda} \right) \cup \left(\bigcup_{\lambda \in \Lambda_{N2}} O_{c\lambda} \right) \cup O_{c\lambda_0}$. Hence, $\mathcal{X}_c$ is compact.

Next, we show that $\mathcal{X}_c$ is Hausdorff. $\forall x_1, x_2 \in \mathcal{X}_c$ with $x_1 \neq x_2$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x_1, x_2 \in \mathcal{X}$; Case 2: $x_1 = \omega$ or $x_2 = \omega$. Case 1: $x_1, x_2 \in \mathcal{X}$. Since $\mathcal{X}$ is Hausdorff, then $\exists O_1, O_2 \in \mathcal{O}$ such that $x_1 \in O_1$, $x_2 \in O_2$, and $O_1 \cap O_2 = \emptyset$. Clearly, $O_1, O_2 \in \mathcal{O}_c$. Case 2: $x_1 = \omega$ or $x_2 = \omega$. Without loss of generality, assume $x_2 = \omega$. Then, $x_1 \in \mathcal{X}$. By local compactness of $\mathcal{X}$, $\exists O_1 \in \mathcal{O}$ such that $x_1 \in O_1 \subseteq \bar{O}_1 \subseteq \mathcal{X}$ and $\bar{O}_1$ is compact. Then, $O_1 \in \mathcal{O}_c$ and $O_2 := X_c \setminus \bar{O}_1 \in \mathcal{O}_c$. Clearly, $x_2 \in O_2$ and $O_1 \cap O_2 = \emptyset$. Hence, in both cases, we have obtained $O_1, O_2 \in \mathcal{O}_c$ such that $x_1 \in O_1$, $x_2 \in O_2$, and $O_1 \cap O_2 = \emptyset$. Hence, $\mathcal{X}_c$ is Hausdorff.

Finally, we show that $\text{id} : \mathcal{X} \rightarrow \mathcal{X}_c \setminus \{\omega\}$ is a homeomorphism. Clearly, id is bijective. $\forall O_c \in \mathcal{O}_c$, we either have $O_c \in \mathcal{O}$, which implies that $O_c \cap (X_c \setminus \{\omega\}) = O_c \in \mathcal{O}$; or we have $X_c \setminus O_c$ is compact, which implies that $\bar{O}_c := O_c \setminus \{\omega\} = X \setminus (X_c \setminus O_c) \in \mathcal{O}$, by Proposition 5.5, and hence $O_c \cap (X_c \setminus \{\omega\}) = \bar{O}_c \in \mathcal{O}$. Hence, the subset topology $\mathcal{O}_{\bar{c}}$ on $\mathcal{X}_c \setminus \{\omega\}$ with respect to $\mathcal{X}_c$ is contained in $\mathcal{O}$. It is easy to see that $\mathcal{O} \subseteq \mathcal{O}_{\bar{c}}$. Then, $\mathcal{O} = \mathcal{O}_{\bar{c}}$. Hence, $\text{id} : \mathcal{X} \rightarrow \mathcal{X}_c \setminus \{\omega\}$ is a homeomorphism.

This completes the proof of the theorem. $\square$

5.7.4 Proper functions

Definition 5.67 Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous. f is said to be proper if $\forall$ compact set $K \subseteq \mathcal{Y}$, we have $f_{\text{inv}}(K) \subseteq \mathcal{X}$ is compact. f is said to be countably proper if $\forall$ compact set $K \subseteq \mathcal{Y}$, we have $f_{\text{inv}}(K) \subseteq \mathcal{X}$ is countably compact.

Proposition 5.68 Let $\mathcal{X} := (X, \mathcal{O}_X)$ and $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be locally compact Hausdorff topological spaces and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous. Let $\mathcal{X}_c := (X_c, \mathcal{O}_{X_c})$ and $\mathcal{Y}_c := (Y_c, \mathcal{O}_{Y_c})$ be the Alexandroff one-point compactifications of $\mathcal{X}$ and $\mathcal{Y}$, respectively, where $X_c = X \cup \{\omega_x\}$ and $Y_c = Y \cup \{\omega_y\}$. Define a function $f_c : \mathcal{X}_c \rightarrow \mathcal{Y}_c$ by $f_c(x) = f(x)$, $\forall x \in X$, and $f_c(\omega_x) = \omega_y$. Then, f is proper if, and only if, f_c is continuous.

Proof “Only if” Let f be proper. $\forall O_{Y_c} \in \mathcal{O}_{Y_c}$, we will distinguish two exhaustive and mutually exclusive cases: Case 1: $O_{Y_c} \in \mathcal{O}_Y$; Case

2: $Y_c \setminus O_{Y_c}$ is compact in $\mathcal{Y}$. Case 1: $O_{Y_c} \in \mathcal{O}_Y$. Then, $f_{\text{cinv}}(O_{Y_c}) = f_{\text{inv}}(O_{Y_c}) \in \mathcal{O}_X$. This implies that $f_{\text{cinv}}(O_{Y_c}) \in \mathcal{O}_{X_c}$. Case 2: $Y_c \setminus O_{Y_c}$ is compact in $\mathcal{Y}$. Then, $f_{\text{cinv}}(Y_c \setminus O_{Y_c}) = f_{\text{inv}}(Y_c \setminus O_{Y_c})$ is compact in $\mathcal{X}$ by the properness of f . Then, by Proposition 2.5, $f_{\text{cinv}}(O_{Y_c}) = X_c \setminus f_{\text{cinv}}(Y_c \setminus O_{Y_c}) \in \mathcal{O}_{X_c}$. Hence, in both cases, we have $f_{\text{cinv}}(O_{Y_c}) \in \mathcal{O}_{X_c}$. Hence, f_c is continuous.

“If” Let f_c be continuous. Fix a compact set $K \subseteq \mathcal{Y}$. Then, $\omega_y \in Y_c \setminus K \in \mathcal{O}_{Y_c}$. By the continuity of f_c , we have $\omega_x \in f_{\text{cinv}}(Y_c \setminus K) \in \mathcal{O}_{X_c}$. This implies that $X_c \setminus f_{\text{cinv}}(Y_c \setminus K)$ is compact in $\mathcal{X}$. By Proposition 2.5, $f_{\text{inv}}(K) = f_{\text{cinv}}(K) = X_c \setminus f_{\text{cinv}}(Y_c \setminus K)$. Hence, f is proper.

This completes the proof of the proposition. $\square$

Proposition 5.69 *Let $\mathcal{X}$ be a topological space, $\mathcal{Y}$ be a metric space, $F \subseteq \mathcal{X}$ be closed, and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous and proper. Then, $f(F) \subseteq \mathcal{Y}$ is closed.*

Proof $\forall y \in \overline{f(F)}$, by Proposition 4.13, $\exists (y_n)_{n=1}^\infty \subseteq f(F)$ such that $\lim_{n \in \mathbb{N}} y_n = y$. Then, $\forall n \in \mathbb{N}$, $\exists x_n \in F$ such that $f(x_n) = y_n$. Let $K = \{y\} \cup (y_n)_{n=1}^\infty$. Then, K is compact. By the properness of f , $f_{\text{inv}}(K)$ is compact. By Propositions 3.5 and 5.5, $F \cap f_{\text{inv}}(K)$ is compact. Note that $(x_n)_{n=1}^\infty \subseteq F \cap f_{\text{inv}}(K)$. Then, by Proposition 5.4, $(x_n)_{n=1}^\infty$ admits a cluster point $x \in F \cap f_{\text{inv}}(K)$. By the continuity of f and Propositions 3.66, we have the sequence $(f(x_n))_{n=1}^\infty = (y_n)_{n=1}^\infty$ admits a cluster point $f(x)$. Since $\mathcal{Y}$ is Hausdorff and $\lim_{n \in \mathbb{N}} y_n = y$, then $y = f(x) \in f(F)$. Therefore, $\overline{f(F)} \subseteq f(F)$ and $f(F)$ is closed. This completes the proof of the proposition. $\square$

In the above proposition, the assumption on $\mathcal{Y}$ may be relaxed to first countable Hausdorff space.

Proposition 5.70 *Let $\mathcal{X}$ be a topological space, $\mathcal{Y}$ be a locally compact Hausdorff topological space, $F \subseteq \mathcal{X}$ be closed, and $f : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous and proper. Then, $f(F) \subseteq \mathcal{Y}$ is closed.*

Proof $\forall y \in \overline{f(F)}$, by the local compactness of $\mathcal{Y}$, $\exists O \in \mathcal{O}_Y$ such that $y \in O$ and $\overline{O}$ is compact in $\mathcal{Y}$. Then, $y \in \overline{f(F)} \cap O$. $\forall U \in \mathcal{O}_Y$ with $y \in U$, by Proposition 3.3, we have $\emptyset \neq (U \cap O) \cap f(F) = U \cap (O \cap f(F))$. Hence, $y \in \overline{f(F)} \cap \overline{O}$ by Proposition 3.3. By Proposition 3.68, there exists a net $(y_\alpha)_{\alpha \in \mathcal{A}} \subseteq f(F) \cap O$ such that $\lim_{\alpha \in \mathcal{A}} y_\alpha = y$. $\forall \alpha \in \mathcal{A}$, $\exists x_\alpha \in F$ such that $y_\alpha = f(x_\alpha)$. Then, $x_\alpha \in F \cap f_{\text{inv}}(O)$. Then, the net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq F \cap f_{\text{inv}}(\overline{O})$. By the properness of f , we have $f_{\text{inv}}(\overline{O}) \subseteq \mathcal{X}$ is compact. By Proposition 3.5, $F \cap f_{\text{inv}}(\overline{O})$ is closed relative to $f_{\text{inv}}(\overline{O})$, which further implies that $F \cap f_{\text{inv}}(\overline{O})$ is compact by Proposition 5.5. By Proposition 5.4, the net $(x_\alpha)_{\alpha \in \mathcal{A}}$ admits a cluster point $x \in F \cap f_{\text{inv}}(\overline{O})$. By the continuity of f and Proposition 3.66, $f(x)$ is a cluster point of the net $(f(x_\alpha))_{\alpha \in \mathcal{A}} = (y_\alpha)_{\alpha \in \mathcal{A}}$. Since $\mathcal{Y}$ is Hausdorff and $\lim_{\alpha \in \mathcal{A}} y_\alpha = y$, then $y = f(x) \in f(F)$. Hence, $\overline{f(F)} \subseteq f(F)$ and $f(F)$ is closed. This completes the proof of the proposition. $\square$

5.8 σ -Compact Spaces

Definition 5.71 A topological space is said to be σ -compact if it is the union of countably infinitely many compact sets.

Proposition 5.72 Let $\mathcal{X}$ be a locally compact topological space. Then, the following statements are equivalent.

- (i) $\mathcal{X}$ is Lindelöf.
- (ii) $\mathcal{X}$ is σ -compact.
- (iii) $\exists (O_n)_{n=1}^\infty \subseteq \mathcal{O}$ such that $\forall n \in \mathbb{N}, \overline{O_n} \subseteq O_{n+1}$ is compact and $\mathcal{X} = \bigcup_{n=1}^\infty O_n$. This sequence is called an exhaustion of $\mathcal{X}$.

Furthermore, if $\mathcal{X}$ is Hausdorff, then the above is equivalent to

- (iv) $\exists \phi : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$ that is proper and continuous.

Proof (i) $\Rightarrow$ (ii). $\forall x \in \mathcal{X}$, by local compactness of $\mathcal{X}$, $\exists O_x \in \mathcal{O}$ such that $x \in O_x \subseteq \overline{O_x}$ and $\overline{O_x}$ is compact. $\mathcal{X} = \bigcup_{x \in \mathcal{X}} O_x$. Since $\mathcal{X}$ is Lindelöf, then $\exists$ a countable set $X_C \subseteq \mathcal{X}$ such that $\mathcal{X} = \bigcup_{x \in X_C} O_x$. We will distinguish three exhaustive and mutually exclusive cases: Case 1: $X_C = \emptyset$; Case 2: $X_C \neq \emptyset$ is finite; Case 3: X_C is countably infinite. Case 1: $X_C = \emptyset$. Let $K_n = \emptyset, \forall n \in \mathbb{N}$, which are clearly compact. Then, $\mathcal{X} = \emptyset = \bigcup_{n=1}^\infty K_n$. Hence, $\mathcal{X}$ is σ -compact. Case 2: $X_C \neq \emptyset$ is finite. Without loss of generality, assume that $X_C = \{x_1, \dots, x_n\}$ for some $n \in \mathbb{N}$. Let $K_i = \overline{O_{x_i}}, i = 1, \dots, n$, and $K_i = \emptyset, i = n+1, n+2, \dots$. Clearly, K_i 's are compact and $\mathcal{X} = \bigcup_{i=1}^\infty K_i$. Hence, $\mathcal{X}$ is σ -compact. Case 3: X_C is countably infinite. Without loss of generality, assume that $X_C = \{x_1, x_2, \dots\}$. Let $K_i = \overline{O_{x_i}}, \forall i \in \mathbb{N}$. Clearly, K_i 's are compact and $\mathcal{X} = \bigcup_{i=1}^\infty K_i$. Hence, $\mathcal{X}$ is σ -compact. In all cases, $\mathcal{X}$ is σ -compact.

(ii) $\Rightarrow$ (iii). Let $\mathcal{X} = \bigcup_{n=1}^\infty K_n$, where K_n is compact, $\forall n \in \mathbb{N}$. Without loss of generality, we assume that $K_n \subseteq K_{n+1}, \forall n \in \mathbb{N}$, since, otherwise, we may let $\bar{K}_n = \bigcup_{i=1}^n K_i$ and consider $\mathcal{X} = \bigcup_{n=1}^\infty \bar{K}_n$ instead. Let $O_0 := \emptyset \in \mathcal{O}$. Then, $\overline{O_0} = \emptyset$ is compact. $\forall n \in \mathbb{N}, K_n \cup \overline{O_{n-1}}$ is compact, by Proposition 5.52, $\exists O_n \in \mathcal{O}$ such that $K_n \cup \overline{O_{n-1}} \subseteq O_n \subseteq \overline{O_n}$ and $\overline{O_n}$ is compact. Then, $(O_n)_{n=1}^\infty$ is an exhaustion of $\mathcal{X}$ that we seek.

(iii) $\Rightarrow$ (i). Let $(U_n)_{n=1}^\infty$ be an exhaustion of $\mathcal{X}$. Fix any open covering $(O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}$ of $\mathcal{X}$, where Λ is an index set. $\forall n \in \mathbb{N}, \overline{U_n} \subseteq \bigcup_{\alpha \in \Lambda} O_\alpha$. By the compactness of $\overline{U_n}$, there exists a finite set $\Lambda_{Nn} \subseteq \Lambda$ such that $\overline{U_n} \subseteq \bigcup_{\alpha \in \Lambda_{Nn}} O_\alpha$. Then, $\mathcal{X} = \bigcup_{n=1}^\infty \overline{U_n} \subseteq \bigcup_{n=1}^\infty \bigcup_{\alpha \in \Lambda_{Nn}} O_\alpha$, which is a countable subcovering. Hence, $\mathcal{X}$ is Lindelöf.

Now, assume that $\mathcal{X}$ is Hausdorff.

(iii) $\Rightarrow$ (iv). Let $(O_n)_{n=1}^\infty$ be an exhaustion of $\mathcal{X}$. $\forall n \in \mathbb{N}, \overline{O_n} \subseteq O_{n+1}$ is compact. Then, by Proposition 5.62, there exists a continuous function $\phi_n : \mathcal{X} \rightarrow [0, 1]$ such that $\phi_n|_{\overline{O_n}} = 1$ and $\phi_n|_{\widetilde{\overline{O_{n+1}}}} = 0$. Let $\phi := \sum_{n=1}^\infty (1 - \phi_n)$.

Clearly, $\phi : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$. $\forall x_0 \in \mathcal{X}$, $\exists n_0 \in \mathbb{N}$ such that $x_0 \in O_{n_0}$. Then, $\phi|_{O_{n_0}} = \sum_{i=1}^{n_0-1} (1 - \phi_i)|_{O_{n_0}}$, which is clearly continuous. $\forall U \in \mathcal{O}_{\mathbb{R}}$ with $\phi(x_0) \in U$, $\exists V \in \mathcal{O}$ with $x_0 \in V \subseteq O_{n_0}$ such that $\phi|_{O_{n_0}}(V) \subseteq U$. Then, $\phi(V) \subseteq U$. Hence, ϕ is continuous at x_0 . By the arbitrariness of x_0 and Proposition 3.9, ϕ is continuous. Fix any compact subset $K \subseteq \mathbb{R}$. Then, K is closed and bounded, by Heine-Borel Theorem. $\exists N \in \mathbb{N}$ such that $K \subseteq [-N, N] \subset \mathbb{R}$. $\forall x \in \widetilde{O_{N+2}}$, we have $x \in \widetilde{O_n}$, $n = 1, \dots, N+2$, and hence $\phi_n(x) = 0$, $n = 1, \dots, N+1$, and $\phi(x) \geq N+1 > N$. Then, $\phi_{\text{inv}}(K) \subseteq O_{N+2} \subseteq \widetilde{O_{N+2}}$. By Proposition 3.10, $\phi_{\text{inv}}(K)$ is closed in $\mathcal{X}$. By the compactness of $\widetilde{O_{N+2}}$ and Proposition 5.5, $\phi_{\text{inv}}(K)$ is compact. Hence, ϕ is proper.

(iv) $\Rightarrow$ (ii). Let $K_n = [-n, n] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$. Then, K_n 's are compact in $\mathbb{R}$ and $\mathbb{R} \subseteq \bigcup_{n=1}^{\infty} K_n$. By the properness of ϕ , $\phi_{\text{inv}}(K_n)$ is compact in $\mathcal{X}$, $\forall n \in \mathbb{N}$. By Proposition 2.5, $\mathcal{X} = \bigcup_{n=1}^{\infty} \phi_{\text{inv}}(K_n)$. Hence, $\mathcal{X}$ is σ -compact.

This completes the proof of the proposition. $\square$

5.9 Paracompact Spaces

Definition 5.73 Let $\mathcal{X}$ be a topological space and $\mathcal{A}$ be a collection of subsets in $\mathcal{X}$. $\mathcal{A}$ is said to be locally finite if $\forall x \in \mathcal{X}$, $\exists U \in \mathcal{O}$ with $x \in U$ such that U meets only finitely many members of $\mathcal{A}$, that is $U \cap A \neq \emptyset$ for finitely many $A \in \mathcal{A}$.

Proposition 5.74 Let $\mathcal{X}$ be a topological space and $(E_\lambda)_{\lambda \in \Lambda}$ be a locally finite collection of subsets of $\mathcal{X}$, where Λ is an index set.

(i) Let $E = \bigcup_{\lambda \in \Lambda} E_\lambda$. Then, $\overline{E} = \bigcup_{\lambda \in \Lambda} \overline{E_\lambda}$.

(ii) Let $K \subseteq \mathcal{X}$ be compact. Then, K meets only finitely many members of $(E_\lambda)_{\lambda \in \Lambda}$.

Proof (i). $\forall x \in \overline{E}$, by local finiteness of $(E_\lambda)_{\lambda \in \Lambda}$, $\exists U \in \mathcal{O}$ with $x \in U$, then U meets only finitely many members of $(E_\lambda)_{\lambda \in \Lambda}$. Let $\Lambda_N \subseteq \Lambda$ be the finite set such that $U \cap E_\lambda \neq \emptyset$, $\forall \lambda \in \Lambda_N$. Suppose $x \notin \bigcup_{\lambda \in \Lambda} \overline{E_\lambda}$. Then, $\forall \lambda \in \Lambda_N$, $x \notin \overline{E_\lambda}$. $\exists U_\lambda \in \mathcal{O}$ with $x \in U_\lambda$ such that $U_\lambda \cap E_\lambda = \emptyset$, by Proposition 3.3. Let $O := (\bigcap_{\lambda \in \Lambda_N} U_\lambda) \cap U \in \mathcal{O}$. Then, $x \in O$ and $O \cap E_\lambda = \emptyset$, $\forall \lambda \in \Lambda$. Then, $O \cap E = \emptyset$. This contradicts with $x \in \overline{E}$, by Proposition 3.3. Hence, $x \in \bigcup_{\lambda \in \Lambda} \overline{E_\lambda}$. Then, $\overline{E} \subseteq \bigcup_{\lambda \in \Lambda} \overline{E_\lambda}$.

On the other hand, $\forall x \in \bigcup_{\lambda \in \Lambda} \overline{E_\lambda}$, $\exists \lambda_0 \in \Lambda$ such that $x \in \overline{E_{\lambda_0}}$. $\forall O \in \mathcal{O}$ with $x \in O$, $O \cap E_{\lambda_0} \neq \emptyset$. Then, $O \cap E \neq \emptyset$ and hence $x \in \overline{E}$. Therefore, we have $\bigcup_{\lambda \in \Lambda} \overline{E_\lambda} \subseteq \overline{E}$.

Hence, $\overline{E} = \bigcup_{\lambda \in \Lambda} \overline{E_\lambda}$.

(ii). Let $K \subseteq \mathcal{X}$ be compact. $\forall x \in K$, by the local finiteness of $(E_\lambda)_{\lambda \in \Lambda}$, $\exists U_x \in \mathcal{O}$ with $x \in U_x$ such that U_x meets only finitely many members of

$(E_\lambda)_{\lambda \in \Lambda}$. Then, $K \subseteq \bigcup_{x \in K} U_x$. By the compactness of K , there exists a finite set $K_N \subseteq K$ such that $K \subseteq \bigcup_{x \in K_N} U_x$. Clearly, $\bigcup_{x \in K_N} U_x$ meets only finitely many members in $(E_\lambda)_{\lambda \in \Lambda}$. Hence, K meets only finitely many members in $(E_\lambda)_{\lambda \in \Lambda}$.

This completes the proof of the proposition. $\square$

Definition 5.75 A topological space $\mathcal{X}$ is said to be paracompact if every open covering of $\mathcal{X}$ has a locally finite open refinement.

Proposition 5.76 A closed subset of a paracompact space is paracompact in the subset topology.

Proof Let $\mathcal{X}$ be a paracompact space, $F \subseteq \mathcal{X}$ be closed, and $\mathcal{O}_F$ be the subset topology on F . Let $(O_{F\alpha})_{\alpha \in \Lambda} \subseteq \mathcal{O}_F$ be any open covering of F , where Λ is an index set (open in the subset topology $\mathcal{O}_F$). By Proposition 3.4, $\forall \alpha \in \Lambda$, $\exists O_\alpha \in \mathcal{O}$ such that $O_{F\alpha} = O_\alpha \cap F$. Then, $F \subseteq \bigcup_{\alpha \in \Lambda} O_\alpha$ and $\mathcal{X} \subseteq (\bigcup_{\alpha \in \Lambda} O_\alpha) \cup \tilde{F}$. By the paracompactness of $\mathcal{X}$, there exists a locally finite open refinement $\mathcal{V}$ of $(O_\alpha)_{\alpha \in \Lambda} \cup \{\tilde{F}\}$. Let $\bar{\mathcal{V}} := \{V \in \mathcal{V} \mid V \subseteq O_\alpha, \text{ for some } \alpha \in \Lambda\}$. Then, $F \subseteq \bigcup_{V \in \bar{\mathcal{V}}} V$, since other V 's in $\mathcal{V}$ are subset of $\tilde{F}$. Clearly, $\bar{\mathcal{V}}$ is locally finite. Then, $\{V \cap F \mid V \in \bar{\mathcal{V}}\} \subseteq \mathcal{O}_F$ is a locally finite open refinement of $(O_{F\alpha})_{\alpha \in \Lambda}$ (open in the subset topology $\mathcal{O}_F$). Hence, $(F, \mathcal{O}_F)$ is paracompact. This completes the proof of the proposition. $\square$

Theorem 5.77 Metric spaces are paracompact.

Proof Let $\mathcal{X}$ be a metric space. Let $(O_\alpha)_{\alpha \in \Lambda}$ be any open covering of $\mathcal{X}$, where Λ is an index set. By Well-Ordering Principle, Λ may be well-ordered by $\preceq$. We will construct an open refinement of $(O_\alpha)_{\alpha \in \Lambda}$ in the following steps. Let $X_0 = \mathcal{X}$.

Step n ($n \in \mathbb{N}$): $\forall \alpha \in \Lambda$, let $P_{\alpha n} := \{x \in O_\alpha \cap X_{n-1} \mid \alpha \text{ is the least element of } \{\beta \in \Lambda \mid x \in O_\beta\}\}$. Let $Q_{\alpha n} := \{x \in P_{\alpha n} \mid \mathcal{B}(x, 3/2^n) \subseteq O_\alpha\}$ and $D_{\alpha n} := \bigcup_{x \in Q_{\alpha n}} \mathcal{B}(x, 1/2^n) \in \mathcal{O}$. Let $X_n := X_{n-1} \setminus (\bigcup_{\alpha \in \Lambda} D_{\alpha n})$.

We will now show that $\mathcal{V} := \{D_{\alpha n} \mid \alpha \in \Lambda, n \in \mathbb{N}\}$ is a locally finite open refinement of $(O_\alpha)_{\alpha \in \Lambda}$. $\forall x_0 \in \mathcal{X}$. Then, $\{\beta \in \Lambda \mid x_0 \in O_\beta\} \neq \emptyset$. Let α_0 be the least element of the set, which exists since Λ is well-ordered. Then, $x_0 \in O_{\alpha_0}$. Then, $\exists n_0 \in \mathbb{N}$ such that $\mathcal{B}(x_0, 3/2^{n_0}) \subseteq O_{\alpha_0}$, since O_{α_0} is open. Then, $x_0 \in D_{\alpha_0 n_0}$ or $x_0 \in D_{\beta n}$ for some $\beta \in \Lambda$ and for some $n \in \mathbb{N}$ with $n < n_0$. Hence, $\mathcal{V}$ covers $\mathcal{X}$. Clearly, $D_{\alpha n} \subseteq O_\alpha$ and is open by construction, $\forall \alpha \in \Lambda$ and $\forall n \in \mathbb{N}$. Then, $\mathcal{V}$ is an open refinement of $(O_\alpha)_{\alpha \in \Lambda}$.

Fix any $x_0 \in \mathcal{X}$. Consider the set $\{\beta \in \Lambda \mid \exists n \in \mathbb{N}, x_0 \in D_{\beta n}\} \neq \emptyset$. Let α_1 be the least element of this set. Then, $x_0 \in D_{\alpha_1 n_0}$ for some $n_0 \in \mathbb{N}$. Furthermore, $\exists j_0 \in \mathbb{N}$ such that $\mathcal{B}(x_0, 2^{-j_0}) \subseteq D_{\alpha_1 n_0}$. Consider the open set $\mathcal{B}(x_0, 2^{-j_0-n_0}) \ni x_0$.

$\forall n \geq n_0 + j_0, \forall \alpha \in \Lambda$, since $\mathcal{B}(x_0, 2^{-j_0}) \subseteq D_{\alpha_1 n_0} \subseteq O_{\alpha_1}$, then $\forall x \in \mathcal{B}(x_0, 2^{-j_0-n_0})$, we have $\mathcal{B}(x, 2^{-j_0} - 2^{-j_0-n_0}) \subseteq \mathcal{B}(x_0, 2^{-j_0}) \subseteq D_{\alpha_1 n_0}$. $\forall y \in D_{\alpha n}$, since $n > n_0$, then $\exists \bar{x} \in Q_{\alpha n}$ with $\bar{x} \notin D_{\alpha_1 n_0}$ such that $y \in \mathcal{B}(\bar{x}, 2^{-n})$. Then, $\exists \bar{x} \in \mathcal{B}(y, 2^{-n})$ such that $\bar{x} \in D_{\alpha n} \setminus D_{\alpha_1 n_0}$. Note that $2^{-n} \leq 2^{-j_0-n_0} \leq 2^{-j_0} - 2^{-j_0-n_0}$. Therefore, $x \notin D_{\alpha n}$. Hence, $\mathcal{B}(x_0, 2^{-j_0-n_0}) \cap D_{\alpha n} = \emptyset$. Thus, $\mathcal{B}(x_0, 2^{-j_0-n_0})$ does not intersect any $D_{\alpha n}, \forall \alpha \in \Lambda, \forall n \geq n_0 + j_0$.

Claim 5.77.1 $\forall n < n_0 + j_0, \mathcal{B}(x_0, 2^{-j_0-n_0})$ intersects at most one of the set $D_{\alpha n}$'s, $\alpha \in \Lambda$.

Proof of claim: Suppose the result is not true. Then, $\exists \alpha, \beta \in \Lambda$ such that $\mathcal{B}(x_0, 2^{-j_0-n_0}) \cap D_{\alpha n} \neq \emptyset, \mathcal{B}(x_0, 2^{-j_0-n_0}) \cap D_{\beta n} \neq \emptyset$, and $\alpha \neq \beta$. Without loss of generality, assume that $\alpha \leq \beta$. Let $p \in \mathcal{B}(x_0, 2^{-j_0-n_0}) \cap D_{\alpha n}$ and $q \in \mathcal{B}(x_0, 2^{-j_0-n_0}) \cap D_{\beta n}$. By the definition of $D_{\alpha n}$, $\exists \bar{p} \in Q_{\alpha n} \subseteq D_{\alpha n}$ such that $p \in \mathcal{B}(\bar{p}, 2^{-n}) \subseteq D_{\alpha n}$. Similarly, $\exists \bar{q} \in Q_{\beta n} \subseteq D_{\beta n}$ such that $q \in \mathcal{B}(\bar{q}, 2^{-n}) \subseteq D_{\beta n}$. Then, by the definition of $Q_{\alpha n}$, we have $\mathcal{B}(\bar{p}, 3/2^n) \subseteq O_\alpha$. Similarly, $\mathcal{B}(\bar{q}, 3/2^n) \subseteq O_\beta$. Note that

$$\begin{aligned} \rho(\bar{p}, \bar{q}) &\leq \rho(\bar{p}, p) + \rho(p, x_0) + \rho(x_0, q) + \rho(q, \bar{q}) \\ &< 2^{-n} + 2^{-j_0-n_0} + 2^{-j_0-n_0} + 2^{-n} \leq 3/2^n \end{aligned}$$

Then, $\bar{q} \in O_\alpha$. But, $\bar{q} \in Q_{\beta n} \subseteq P_{\beta n}, \alpha \leq \beta$, and $\alpha \neq \beta$ implies that $\bar{q} \notin O_\alpha$. This is a contradiction. Therefore, the result must hold. This completes the proof of the claim. $\square$

Hence, $\mathcal{B}(x_0, 2^{-j_0-n_0})$ can meet at most $n_0 + j_0 - 1$ sets in $\mathcal{V}$. Hence, $\mathcal{V}$ is locally finite.

Therefore, we have obtained a locally finite open refinement of $(O_\alpha)_{\alpha \in \Lambda}$. Hence, $\mathcal{X}$ is paracompact. This completes the proof of the theorem. $\square$

Definition 5.78 Let $\mathcal{X}$ be a topological space and $(E_\alpha)_{\alpha \in \Lambda}$ be a collection of subsets of $\mathcal{X}$, where Λ is an index set. $(E_\alpha)_{\alpha \in \Lambda}$ is said to be star-finite if $\forall \alpha_0 \in \Lambda, E_{\alpha_0}$ meets only finitely many members in the collection.

It is easy to see that a star-finite open covering of $\mathcal{X}$ is locally finite. But the converse need not hold.

Proposition 5.79 A σ -compact locally compact space is paracompact.

Proof Let $\mathcal{X}$ be a σ -compact locally compact space. Fix any open covering $(O_\alpha)_{\alpha \in \Lambda}$ of $\mathcal{X}$, where Λ is an index set. By Proposition 5.72, there exists an exhaustion $(U_n)_{n=1}^\infty$ of $\mathcal{X}$. Let $U_{-1} = \emptyset$ and $U_0 = \emptyset$. Let $V_{\alpha n} := O_\alpha \cap (U_{n+1} \setminus \overline{U_{n-2}}) \in \mathcal{O}, \forall \alpha \in \Lambda, \forall n \in \mathbb{N}$. It is easy to see that $\mathcal{V} := \{V_{\alpha n} \mid \alpha \in \Lambda, n \in \mathbb{N}\}$ is an open refinement of $(O_\alpha)_{\alpha \in \Lambda}$ that covers $\mathcal{X}$. $\forall n \in \mathbb{N}$, by the compactness of $\overline{U_n}$ and Proposition 5.5, $\overline{U_n} \setminus U_{n-1} \subseteq \bigcup_{\alpha \in \Lambda} V_{\alpha n}$ is compact. Then, there exists a finite set $\mathcal{V}_n \subseteq$

$\{V_{\alpha n} \mid \alpha \in \Lambda\}$ such that $\overline{U_n} \setminus U_{n-1} \subseteq \bigcup_{V \in \mathcal{V}_n} V$. Then, $\mathcal{V}_L := \bigcup_{n=1}^{\infty} \mathcal{V}_n$ is an open refinement of $(O_{\alpha})_{\alpha \in \Lambda}$ that covers $\mathcal{X}$ since $\bigcup_{n=1}^{\infty} (\overline{U_n} \setminus U_{n-1}) = \mathcal{X}$. $\forall V \in \mathcal{V}_L$, $\exists n_0 \in \mathbb{N}$ such that $V \in \mathcal{V}_{n_0}$. Then, V does not intersect any member of $\mathcal{V}_n$ with $n \in \mathbb{N}$ and $|n - n_0| > 2$. Hence, $\mathcal{V}_L$ is star-finite. Then, $\mathcal{V}_L$ is a locally finite open refinement of $(O_{\alpha})_{\alpha \in \Lambda}$. Then, $\mathcal{X}$ is paracompact. This completes the poof of the proposition. $\square$

Proposition 5.80 *Let $\mathcal{X}$ be a locally compact Hausdorff topological space. $\mathcal{X}$ is paracompact if, and only if, any open covering of $\mathcal{X}$ has a star-finite open refinement (that covers $\mathcal{X}$).*

Proof “Only if” Let $\mathcal{U}$ be any open covering of $\mathcal{X}$. $\forall x \in \mathcal{X}$, $\exists U_x \in \mathcal{U}$ such that $x \in U_x \in \mathcal{O}$. By Proposition 5.53, $\exists O_x \in \mathcal{O}$ such that $x \in O_x \subseteq \overline{O_x} \subseteq U_x$ and $\overline{O_x}$ is compact. Hence, $(O_x)_{x \in \mathcal{X}}$ is an open refinement of $\mathcal{U}$. By the paracompactness of $\mathcal{X}$, there exists a locally finite open refinement $\mathcal{V}$ of $(O_x)_{x \in \mathcal{X}}$ that covers $\mathcal{X}$. $\forall V \in \mathcal{V}$, $\exists x \in \mathcal{X}$ such that $V \subseteq O_x$. Then, $\overline{V} \subseteq \overline{O_x}$. By Propositions 5.5 and 3.5, we have $\overline{V}$ is compact. By Proposition 5.74, $\overline{V}$ meets only finitely many members of $\mathcal{V}$. Hence, $\mathcal{V}$ is star-finite. Therefore, $\mathcal{V}$ is a star-finite open refinement of $\mathcal{U}$.

“If” Let $\mathcal{U}$ be any open covering of $\mathcal{X}$. Then, there is a star-finite open refinement $\mathcal{V}$ of $\mathcal{U}$. Then, $\mathcal{V}$ is a locally finite open refinement of $\mathcal{U}$. Hence, $\mathcal{X}$ is paracompact.

This completes the proof of the proposition. $\square$

Proposition 5.81 *A paracompact Hausdorff topological space is normal.*

Proof Let $\mathcal{X}$ be a paracompact Hausdorff topological space. Clearly, $\mathcal{X}$ is Tychonoff. $\forall x_0 \in \mathcal{X}$ and $\forall$ closed set $F \subseteq \mathcal{X}$ with $x_0 \notin F$, we will show that $\exists O_1, O_2 \in \mathcal{O}$ such that $x_0 \in O_1$, $F \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$. $\forall x \in F$, then $x \neq x_0$. Since $\mathcal{X}$ is Hausdorff, $\exists O_x^{(1)}, O_x^{(2)} \in \mathcal{O}$ such that $x_0 \in O_x^{(1)}$, $x \in O_x^{(2)}$, and $O_x^{(1)} \cap O_x^{(2)} = \emptyset$. Then, $\mathcal{X} \subseteq \widetilde{F} \cup \left(\bigcup_{x \in F} O_x^{(2)} \right)$. By the paracompactness of $\mathcal{X}$, there exists a locally finite open refinement $\mathcal{V} \subseteq \mathcal{O}$ of $\{\widetilde{F}\} \cup \left\{ O_x^{(2)} \mid x \in F \right\}$. Let $\widetilde{\mathcal{V}} := \left\{ V \in \mathcal{V} \mid V \subseteq O_x^{(2)}, \text{ for some } x \in F \right\}$. Then, $\widetilde{\mathcal{V}}$ is an open covering of F , since $\forall V \in \mathcal{V} \setminus \widetilde{\mathcal{V}}$ we have $V \subseteq \widetilde{F}$. Furthermore, $\widetilde{\mathcal{V}}$ is locally finite. $\forall V \in \widetilde{\mathcal{V}}$, $\exists x \in F$ such that $V \subseteq O_x^{(2)}$, $x_0 \in O_x^{(1)}$, and $O_x^{(2)} \cap O_x^{(1)} = \emptyset$. Then, $V \subseteq O_x^{(2)} \subseteq \widetilde{O_x^{(1)}}$. Hence, we have $\overline{V} \subseteq \widetilde{O_x^{(1)}}$ and $x_0 \notin \overline{V}$. By Proposition 5.74, $\overline{\bigcup_{V \in \widetilde{\mathcal{V}}} V} = \bigcup_{V \in \widetilde{\mathcal{V}}} \overline{V} =: \widetilde{O_1}$. Then, $O_1 \in \mathcal{O}$, $F \subseteq \bigcup_{V \in \widetilde{\mathcal{V}}} V =: O_2 \in \mathcal{O}$, $O_2 \subseteq \widetilde{O_1}$, and $x_0 \in O_1$. Hence, $\exists O_1, O_2 \in \mathcal{O}$ such that $x_0 \in O_1$, $F \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$. Then, $\mathcal{X}$ is regular.

Next, we show that $\mathcal{X}$ is normal. Fix any closed sets $F_1, F_2 \subseteq \mathcal{X}$ with $F_1 \cap F_2 = \emptyset$. $\forall x \in F_2$, then $x \notin F_1$. Since $\mathcal{X}$ is regular, $\exists O_x^{(1)}, O_x^{(2)} \in \mathcal{O}$ such that $F_1 \subseteq O_x^{(1)}$, $x \in O_x^{(2)}$, and $O_x^{(1)} \cap O_x^{(2)} = \emptyset$. Then, $\mathcal{X} \subseteq \widetilde{F_2} \cup$

$(\bigcup_{x \in F_2} O_x^{(2)})$. By the paracompactness of $\mathcal{X}$, there exists a locally finite open refinement $\mathcal{V} \subseteq \mathcal{O}$ of $\{\widetilde{F_2}\} \cup \{O_x^{(2)} \mid x \in F_2\}$. Let $\bar{\mathcal{V}} := \{V \in \mathcal{V} \mid V \subseteq O_x^{(2)}, \text{ for some } x \in F_2\}$. Then, $\bar{\mathcal{V}}$ is an open covering of F_2 , since $\forall V \in \mathcal{V} \setminus \bar{\mathcal{V}}$ we have $V \subseteq \widetilde{F_2}$. Furthermore, $\bar{\mathcal{V}}$ is locally finite. $\forall V \in \bar{\mathcal{V}}$, $\exists x \in F_2$ such that $V \subseteq O_x^{(2)}$, $F_1 \subseteq \widetilde{O_x^{(1)}}$, and $O_x^{(2)} \cap O_x^{(1)} = \emptyset$. Then, $V \subseteq \widetilde{O_x^{(2)}} \subseteq \widetilde{O_x^{(1)}}$. Hence, we have $\bar{V} \subseteq \widetilde{O_x^{(1)}}$ and $F_1 \cap \bar{V} = \emptyset$. By Proposition 5.74, $\bigcup_{V \in \bar{\mathcal{V}}} \bar{V} = \bigcup_{V \in \bar{\mathcal{V}}} \widetilde{V} =: \widetilde{O_1}$. Then, $O_1 \in \mathcal{O}$, $F_2 \subseteq \bigcup_{V \in \bar{\mathcal{V}}} V =: O_2 \in \mathcal{O}$, $O_2 \subseteq \widetilde{O_1}$, and $F_1 \cap \widetilde{O_1} = \emptyset$. Hence, $\exists O_1, O_2 \in \mathcal{O}$ such that $F_1 \subseteq O_1$, $F_2 \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$. Then, $\mathcal{X}$ is normal.

This completes the proof of the proposition. $\square$

5.10 The Stone-Čech Compactification

Definition 5.82 Let $\mathcal{X} := (X, \mathcal{O})$ be a completely regular topological space, $I = [0, 1] \subseteq \mathbb{R}$, and $\mathcal{F}$ be the family of continuous functions of $\mathcal{X}$ to I . By Corollary 3.62, $\mathcal{X}$ is homeomorphic to $E(\mathcal{X}) \subseteq I^{\mathcal{F}}$, where E is the equivalence map. Let $F = \overline{E(\mathcal{X})}$ and $\mathcal{O}_F$ be the subset topology of F . Then, $\beta(\mathcal{X}) := (F, \mathcal{O}_F)$ is a compact Hausdorff topological space, by Tychonoff Theorem and Proposition 5.5. $\beta(\mathcal{X})$ is said to be the Stone-Čech compactification of $\mathcal{X}$.

Proposition 5.83 Let $\mathcal{X} := (X, \mathcal{O})$ be a completely regular topological space, $I = [0, 1] \subseteq \mathbb{R}$, $\mathcal{F}$ be the family of continuous functions of $\mathcal{X}$ to I , and $E : \mathcal{X} \rightarrow E(\mathcal{X}) \subseteq I^{\mathcal{F}}$ be the equivalence map. Then, there exists a unique compact Hausdorff topological space $\beta(\mathcal{X})$ with the following properties:

- (i) the space $E(\mathcal{X})$ is dense in $\beta(\mathcal{X})$;
- (ii) each bounded continuous real-valued function on $E(\mathcal{X})$ extends to a bounded continuous real-valued function on $\beta(\mathcal{X})$;
- (iii) if $\mathcal{X}$ is a dense subset of a compact Hausdorff topological space $\mathcal{Y}$, then, there exists a unique continuous mapping $\phi : \beta(\mathcal{X}) \rightarrow \mathcal{Y}$ such that ϕ is surjective and $\phi|_{E(\mathcal{X})} = E_{\text{inv}}$.

Furthermore, if $\mathcal{X}$ is locally compact, then $E(\mathcal{X})$ is an open subset of $\beta(\mathcal{X})$.

Proof We will first show that $\beta(\mathcal{X})$ defined in Definition 5.82 satisfies properties (i), (ii), and (iii). (i). By Proposition 3.5, $E(\mathcal{X})$ is dense in $F = \overline{E(\mathcal{X})} = \beta(\mathcal{X})$. (ii). Fix any bounded continuous real-valued function $g : E(\mathcal{X}) \rightarrow \mathbb{R}$. Then, $\exists N \in \mathbb{N}$ such that $g : E(\mathcal{X}) \rightarrow [-N, N] \subset \mathbb{R}$. Then, $\bar{g} := ((g + N)/(2N)) \circ E \in \mathcal{F}$ by Proposition 3.12. Now, consider the function $\bar{h} := \pi_{\bar{g}}|_{\beta(\mathcal{X})} : \beta(\mathcal{X}) \rightarrow [0, 1] \subset \mathbb{R}$, which is continuous by

Proposition 3.27. Note that $\bar{h}|_{E(\mathcal{X})} = \bar{g} \circ E_{\text{inv}} = (g + N)/(2N)$. Then, by Proposition 3.12, $h := 2N\bar{h} - N$ is a continuous function of $\beta(\mathcal{X})$ to $[-N, N]$. Furthermore, $h|_{E(\mathcal{X})} = g$. Hence, h is the desired extension that we seek. Clearly, $h : \beta(\mathcal{X}) \rightarrow [-N, N] \subset \mathbb{R}$.

(iii). Let $\mathcal{Y} := (Y, \mathcal{O}_{\mathcal{Y}})$ be a compact Hausdorff space such that $\mathcal{X}$ is dense in $\mathcal{Y}$. Let $\mathcal{G}$ be the family of continuous real-valued functions of $\mathcal{Y}$ to I . Let $E^{(\mathcal{Y})} : \mathcal{Y} \rightarrow I^{\mathcal{G}}$ be the equivalence map and $\pi_g^{(\mathcal{G})} : I^{\mathcal{G}} \rightarrow I$ be the projection function, $\forall g \in \mathcal{G}$. By Propositions 5.14 and 3.61, $\mathcal{Y}$ is completely regular. By Corollary 3.62, $E^{(\mathcal{Y})} : \mathcal{Y} \rightarrow E^{(\mathcal{Y})}(\mathcal{Y}) \subseteq I^{\mathcal{G}}$ is a homeomorphism. Define a mapping $\psi : \mathcal{G} \rightarrow \mathcal{F}$ by $\psi(g) = g|_{\mathcal{X}}$, $\forall g \in \mathcal{G}$. By Proposition 3.56, ψ is injective. By Tychonoff Theorem, $I^{\mathcal{G}}$ is a compact Hausdorff space. Define a mapping $\Psi : I^{\mathcal{F}} \rightarrow I^{\mathcal{G}}$ by $\pi_g^{(\mathcal{G})}(\Psi(i_f)) = \pi_{\psi(g)}(i_f)$, $\forall i_f \in I^{\mathcal{F}}$, $\forall g \in \mathcal{G}$. $\forall x \in \mathcal{X}$, $E^{(\mathcal{Y})}(x) \in I^{\mathcal{G}}$ satisfies $\pi_g^{(\mathcal{G})}(E^{(\mathcal{Y})}(x)) = g(x) = g|_{\mathcal{X}}(x) = \psi(g)(x) = \pi_{\psi(g)}(E(x))$, $\forall g \in \mathcal{G}$. Then, we have $E^{(\mathcal{Y})}(x) = \Psi(E(x))$, $\forall x \in \mathcal{X}$. Hence, $\Psi \circ E = E^{(\mathcal{Y})}|_{\mathcal{X}}$.

Claim 5.83.1 Ψ is continuous.

Proof of claim: Fix any basis open set $U \subseteq I^{\mathcal{G}}$. Then, $U = \prod_{g \in \mathcal{G}} U_g$, where $U_g \subseteq I$ is open, $\forall g \in \mathcal{G}$, and $U_g = I$ for all g 's except finitely many g 's, say $g \in \mathcal{G}_N \subseteq \mathcal{G}$. Let $V := \prod_{f \in \mathcal{F}} V_f \subseteq I^{\mathcal{F}}$ be given by $V_f = I$, $\forall f \notin \text{range}(\psi)$, and $V_f = U_g$, $\forall f \in \text{range}(\psi)$ and $f = \psi(g)$. The set V is well defined since ψ is injective. $\forall i_f \in V$, $\forall g \in \mathcal{G}$, $\pi_g^{(\mathcal{G})}(\Psi(i_f)) = \pi_{\psi(g)}(i_f) \in V_{\psi(g)} = U_g$. Hence, we have $\Psi(i_f) \in U$. Then, $V \subseteq \Psi_{\text{inv}}(U)$. $\forall i_f \in I^{\mathcal{F}} \setminus V$, $\exists f_0 \in \text{range}(\psi)$ such that $\pi_{f_0}(i_f) \notin V_{f_0}$. Then, $f_0 = \psi(g_0)$, where $g_0 \in \mathcal{G}_N$. Note that $\pi_{g_0}^{(\mathcal{G})}(\Psi(i_f)) = \pi_{\psi(g_0)}(i_f) \notin V_{\psi(g_0)} = U_{g_0}$. Then, $\Psi(i_f) \notin U$. This shows that $I^{\mathcal{F}} \setminus V \subseteq \Psi_{\text{inv}}(I^{\mathcal{G}} \setminus U)$. Hence, we have $V = \Psi_{\text{inv}}(U)$, by Proposition 2.5. Clearly, V is a basis open set in $I^{\mathcal{F}}$. Hence, Ψ is continuous. This completes the proof of the claim. $\square$

Since $\mathcal{X}$ is dense in $\mathcal{Y}$ and $E^{(\mathcal{Y})}$ is a homeomorphism between $\mathcal{Y}$ and $E^{(\mathcal{Y})}(\mathcal{Y})$, then $E^{(\mathcal{Y})}(\mathcal{X})$ is dense in $E^{(\mathcal{Y})}(\mathcal{Y})$. Since $\mathcal{Y}$ is compact and $E^{(\mathcal{Y})}$ is continuous, then $E^{(\mathcal{Y})}(\mathcal{Y})$ is compact in $I^{\mathcal{G}}$, by Proposition 5.7. Furthermore, by Proposition 5.5, $E^{(\mathcal{Y})}(\mathcal{Y})$ is closed in $I^{\mathcal{G}}$. Then, $\overline{E^{(\mathcal{Y})}(\mathcal{X})} = E^{(\mathcal{Y})}(\mathcal{Y})$, where $\overline{E^{(\mathcal{Y})}(\mathcal{X})}$ is the closure of $E^{(\mathcal{Y})}(\mathcal{X})$ in $I^{\mathcal{G}}$. By the compactness of $\beta(\mathcal{X})$, the continuity of Ψ , and Proposition 5.7, $\Psi(\beta(\mathcal{X})) \subseteq I^{\mathcal{G}}$ is compact. Furthermore, by Proposition 5.5, $\Psi(\beta(\mathcal{X}))$ is closed in $I^{\mathcal{G}}$. Note that $E^{(\mathcal{Y})}(\mathcal{X}) = \Psi(E(\mathcal{X})) \subseteq \Psi(\beta(\mathcal{X}))$. Then, $\Psi(\beta(\mathcal{X})) \supseteq E^{(\mathcal{Y})}(\mathcal{Y})$.

Claim 5.83.2 $\Psi(\beta(\mathcal{X})) = E^{(\mathcal{Y})}(\mathcal{Y})$.

Proof of claim: Suppose $E^{(\mathcal{Y})}(\mathcal{Y}) \subset \Psi(\beta(\mathcal{X}))$, then $\Psi_{\text{inv}}(E^{(\mathcal{Y})}(\mathcal{Y})) \subset \beta(\mathcal{X})$ and $\beta(\mathcal{X}) \cap (I^{\mathcal{F}} \setminus \Psi_{\text{inv}}(E^{(\mathcal{Y})}(\mathcal{Y}))) \neq \emptyset$. Note that $\Psi_{\text{inv}}(E^{(\mathcal{Y})}(\mathcal{Y}))$ is closed in $I^{\mathcal{F}}$ by the closedness of $E^{(\mathcal{Y})}(\mathcal{Y})$, the continuity of Ψ , and Proposition 3.10, which further implies that $I^{\mathcal{F}} \setminus \Psi_{\text{inv}}(E^{(\mathcal{Y})}(\mathcal{Y})) \in \mathcal{O}_{I^{\mathcal{F}}}$. By the denseness of $E(\mathcal{X})$ in $\beta(\mathcal{X})$, we have $E(\mathcal{X}) \cap (I^{\mathcal{F}} \setminus \Psi_{\text{inv}}(E^{(\mathcal{Y})}(\mathcal{Y}))) \neq \emptyset$ and

hence $\exists x \in \mathcal{X}$ such that $E(x) \in I^{\mathcal{F}} \setminus \Psi_{\text{inv}}(E^{(\mathcal{Y})}(\mathcal{Y}))$. By Proposition 2.5, $I^{\mathcal{F}} \setminus \Psi_{\text{inv}}(E^{(\mathcal{Y})}(\mathcal{Y})) = \Psi_{\text{inv}}(I^{\mathcal{G}} \setminus E^{(\mathcal{Y})}(\mathcal{Y}))$. Then, $\Psi(E(x)) \in I^{\mathcal{G}} \setminus E^{(\mathcal{Y})}(\mathcal{Y})$. This contradicts with the fact that $\Psi(E(x)) = E^{(\mathcal{Y})}(x) \in E^{(\mathcal{Y})}(\mathcal{Y})$. Therefore, we must have $\Psi(\beta(\mathcal{X})) = E^{(\mathcal{Y})}(\mathcal{Y})$. This completes the proof of the claim. $\square$

Define $\phi : \beta(\mathcal{X}) \rightarrow \mathcal{Y}$ by $\phi = E^{(\mathcal{Y})}_{\text{inv}} \circ \Psi|_{\beta(\mathcal{X})}$. Clearly, ϕ is continuous and surjective by Proposition 3.12. Clearly, $\phi \circ E = E^{(\mathcal{Y})}_{\text{inv}} \circ E^{(\mathcal{Y})}|_{\mathcal{X}} = \text{id}_{\mathcal{X}}$. Hence $\phi|_{E(\mathcal{X})} = E_{\text{inv}}$.

Let $\bar{\phi} : \beta(\mathcal{X}) \rightarrow \mathcal{Y}$ be any continuous and surjective mapping such that $\bar{\phi}|_{E(\mathcal{X})} = E_{\text{inv}} = \phi|_{E(\mathcal{X})}$. By Proposition 3.56 and the denseness of $E(\mathcal{X})$ in $\beta(\mathcal{X})$, we have $\phi = \bar{\phi}$. Hence, ϕ is unique.

Thus, we have shown that the Stone-Ćech compactification $\beta(\mathcal{X})$ satisfies (i), (ii), and (iii). Next, we will show that $\beta(\mathcal{X})$ is unique. Let $\mathcal{Y}$ be any compact Hausdorff space that satisfies (i), (ii), and (iii). Since $\mathcal{X}$ and $E(\mathcal{X})$ are homeomorphic, we may identify $\mathcal{X}$ as $E(\mathcal{X})$. Then, by (iii), $\exists! \phi : \beta(\mathcal{X}) \rightarrow \mathcal{Y}$, which is continuous and surjective, such that $\phi|_{\mathcal{X}} = \text{id}_{\mathcal{X}}$. Since $\mathcal{Y}$ satisfies (iii), then $\exists! \lambda : \mathcal{Y} \rightarrow \beta(\mathcal{X})$, which is continuous and surjective, such that $\lambda|_{\mathcal{X}} = \text{id}_{\mathcal{X}}$. Then, $\phi \circ \lambda : \mathcal{Y} \rightarrow \mathcal{Y}$ is continuous and satisfies $(\phi \circ \lambda)|_{\mathcal{X}} = \text{id}_{\mathcal{X}}$. By Proposition 3.56, we have $\phi \circ \lambda = \text{id}_{\mathcal{Y}}$. On the other hand, $\lambda \circ \phi : \beta(\mathcal{X}) \rightarrow \beta(\mathcal{X})$ is continuous and satisfies $(\lambda \circ \phi)|_{\mathcal{X}} = \text{id}_{\mathcal{X}}$. Then, by Proposition 3.56, we have $\lambda \circ \phi = \text{id}_{\beta(\mathcal{X})}$. By Proposition 2.4, we have $\lambda = \phi_{\text{inv}}$. Hence, $\beta(\mathcal{X})$ and $\mathcal{Y}$ are homeomorphic. Hence, $\beta(\mathcal{X})$ is unique.

If, in addition, $\mathcal{X}$ is locally compact, then, by Proposition 5.57, $\mathcal{X}$ is open in $\beta(\mathcal{X})$. This completes the proof of the proposition. $\square$

Example 5.84 $\mathbb{R}_e$ with the topology $\mathcal{O}_{\mathbb{R}_e}$ introduced in Example 3.80 is Hausdorff and second countable. It is easy to show that $\mathbb{R}_e$ is compact. By Proposition 5.14, $\mathbb{R}_e$ is a normal topological space. By Urysohn Metrization Theorem 4.53, $\mathbb{R}_e$ is metrizable. Hence, $\mathbb{R}_e$ is a second countable metrizable compact Hausdorff topological space. $\diamond$

Chapter 6

Vector Spaces

6.1 Group

Definition 6.1 A group is the triple $(G, +, e)$, which consists a nonempty set G , an operation $+: G \times G \rightarrow G$, and a unit element $e \in G$, satisfying, $\forall g_1, g_2, g_3 \in G$,

- (i) $(g_1 + g_2) + g_3 = g_1 + (g_2 + g_3)$; (associativeness)
- (ii) $e + g_1 = g_1 + e = g_1$; (unit element)
- (iii) $\exists(-g_1) \in G$ such that $g_1 + (-g_1) = e = (-g_1) + g_1$. (existence of inverse, which is clearly unique)

We have the following result.

Proposition 6.2 Let $(G, +, e)$ be a group and $H \subseteq G$ be nonempty. Then, $(H, +, e)$ is a group (which will be called a subgroup of $(G, +, e)$) if, and only if, $\forall g_1, g_2 \in H$, we have $g_1 + (-g_2) \in H$.

Proof “Necessity” This is obvious.

“Sufficiency” Since $H \neq \emptyset$, then $\exists g \in H$. Then, $e = g + (-g) \in H$. $\forall g_1, g_2 \in H$, $e + (-g_2) = (-g_2) \in H$. Then, $g_1 + g_2 = g_1 + (-(-g_2)) \in H$. Hence, H is closed under $+$. Then, it is straightforward to check that $(H, +, e)$ satisfies all the properties of Definition 6.1. Hence, it is a group.

This completes the proof of the proposition. $\square$

Proposition 6.3 Let $(G, +, e)$ be a group and $g_1, g_2, g_3 \in G$. If $g_1 + g_2 = g_1 + g_3$ then $g_2 = g_3$. On the other hand, if $g_1 + g_3 = g_2 + g_3$ then $g_1 = g_2$.

Proof If $g_1 + g_2 = g_1 + g_3$, then we have

$$\begin{aligned} g_2 &= e + g_2 = ((-g_1) + g_1) + g_2 = (-g_1) + (g_1 + g_2) \\ &= (-g_1) + (g_1 + g_3) = ((-g_1) + g_1) + g_3 = e + g_3 = g_3 \end{aligned}$$

If $g_1 + g_3 = g_2 + g_3$, then we have

$$\begin{aligned} g_1 &= g_1 + e = g_1 + (g_3 + (-g_3)) = (g_1 + g_3) + (-g_3) \\ &= (g_2 + g_3) + (-g_3) = g_2 + (g_3 + (-g_3)) = g_2 + e = g_2 \end{aligned}$$

This completes the proof of the proposition. $\square$

Definition 6.4 Let $(G, +, e)$ be a group, the order of the group is the number of elements in G , if G is finite. $\forall g \in G$, the order of g is the integer $n > 0$ such that $\underbrace{g + \cdots + g}_n = e$.

Definition 6.5 Let $(G, +_G, e_G)$ and $(H, +_H, e_H)$ be two groups, and $T : G \rightarrow H$. T is said to be a homomorphism if, $\forall g_1, g_2 \in G$, we have $T(g_1) +_H T(g_2) = T(g_1 +_G g_2)$. T is said to be an isomorphism if it is bijective and a homomorphism, in this case, the two groups are said to be isomorphic.

Let $T : G \rightarrow H$ be a homomorphism, then $T(e_G) = e_H$.

Definition 6.6 Let $(G, +, e)$ be a group. $g_1, g_2 \in G$ are said to be conjugate if $\exists g_3 \in G$ such that $g_2 = (-g_3) + g_1 + g_3$. Let $(H, +, e)$ be a subgroup of $(G, +, e)$. It is said to be normal (self-conjugate) if, $\forall h \in H, \forall g \in G$, we have $(-g) + h + g \in H$.

Definition 6.7 Let $(G, +, \mathbf{0})$ be a group. It is said to be abelian if, $\forall g_1, g_2 \in G$, we have $g_1 + g_2 = g_2 + g_1$ (commutativity). Then, the unit element $\mathbf{0}$ is also called the zero-element.

Sometimes, we are interested in an algebraic structure that is weaker than a group. For example, the structure of functions $f : X \rightarrow X$ with respect to the function composition operation. This leads us to the following definition.

Definition 6.8 A semigroup is the triple $(G, \circ, e)$, which consists a nonempty set G , an operation $\circ : G \times G \rightarrow G$, and a unit element $e \in G$, satisfying, $\forall g_1, g_2, g_3 \in G$,

$$(i) \quad (g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3); \quad (\text{associativity})$$

$$(ii) \quad e \circ g_1 = g_1 = g_1 \circ e. \quad (\text{unit element})$$

Furthermore, a semigroup that satisfies $g_1 \circ g_2 = g_2 \circ g_1, \forall g_1, g_2 \in G$, is called an abelian semigroup.

6.2 Ring

Definition 6.9 A ring $(R, +, \times, \mathbf{0})$ is an abelian group $(R, +, \mathbf{0})$ with an operation $\times : R \times R \rightarrow R$ such that, $\forall r_1, r_2, r_3 \in R$,

- (i) $(r_1 \times r_2) \times r_3 = r_1 \times (r_2 \times r_3)$; (associativeness)
- (ii) $r_1 \times (r_2 + r_3) = r_1 \times r_2 + r_1 \times r_3$; (right distributiveness)
- (iii) $(r_1 + r_2) \times r_3 = r_1 \times r_3 + r_2 \times r_3$. (left distributiveness)

A ring is commutative if, $\forall r_1, r_2 \in R$, we have $r_1 \times r_2 = r_2 \times r_1$. A ring is with identity element if $\exists \mathbf{1} \in R$ such that, $\forall r \in R$, $\mathbf{1} \times r = r \times \mathbf{1} = r$. (The identity element is unique if it exists.)

Proposition 6.10 Let $(R, +, \times, \mathbf{0})$ be a ring. $S \subseteq R$. Then, $(S, +, \times, \mathbf{0})$ is a ring (which will be called a subring) if $(S, +, \mathbf{0})$ is a subgroup of $(R, +, \mathbf{0})$ and $s_1 \times s_2 \in S$, $\forall s_1, s_2 \in S$.

Proof Clearly, $(S, +, \mathbf{0})$ is an abelian group. Then, it is straightforward to show that $(S, +, \times, \mathbf{0})$ is a ring. $\square$

Proposition 6.11 Let $(R, +, \times, \mathbf{0})$ be a ring, and $r_1, r_2 \in R$, then, we have

- (i) $r_1 \times \mathbf{0} = \mathbf{0} \times r_1 = \mathbf{0}$;
- (ii) $r_1 \times (-r_2) = (-r_1) \times r_2 = -(r_1 \times r_2)$;
- (iii) $(-r_1) \times (-r_2) = r_1 \times r_2$.

Proof Note that $\mathbf{0} + \mathbf{0} \times r_1 = \mathbf{0} \times r_1 = (\mathbf{0} + \mathbf{0}) \times r_1 = \mathbf{0} \times r_1 + \mathbf{0} \times r_1$. By Proposition 6.3, we have $\mathbf{0} \times r_1 = \mathbf{0}$. Similarly, we can show that $r_1 \times \mathbf{0} = \mathbf{0}$.

Note that $r_1 \times r_2 + r_1 \times (-r_2) = r_1 \times (r_2 + (-r_2)) = r_1 \times \mathbf{0} = \mathbf{0}$. Then, we have $r_1 \times (-r_2) = -(r_1 \times r_2)$. Similarly, we can show that $(-r_1) \times r_2 = -(r_1 \times r_2)$.

Note that $(-r_1) \times (-r_2) + (-(r_1 \times r_2)) = (-r_1) \times (-r_2) + r_1 \times (-r_2) = ((-r_1) + r_1) \times (-r_2) = \mathbf{0} \times (-r_2) = \mathbf{0}$. Then, we have $(-r_1) \times (-r_2) = r_1 \times r_2$.

This completes the proof of the proposition. $\square$

Definition 6.12 Let $(R, +, \times, \mathbf{0})$ be a ring and $S \subseteq R$. Then, S is said to be an ideal if $(S, +, \mathbf{0})$ is a subgroup of $(R, +, \mathbf{0})$ and $r \times s, s \times r \in S$, $\forall r \in R$ and $\forall s \in S$.

Definition 6.13 Let $(R, +_R, \times_R, \mathbf{0}_R)$ and $(S, +_S, \times_S, \mathbf{0}_S)$ be two rings and $T : R \rightarrow S$. T is said to be a ring-homomorphism if it is a homomorphism and $T(r_1) \times_S T(r_2) = T(r_1 \times_R r_2)$, $\forall r_1, r_2 \in R$. T is said to be a ring-isomorphism if it is a bijective ring-homomorphism, in this case, the two rings are said to be isomorphic.

6.3 Field

Definition 6.14 Let $(F, +, \times, \mathbf{0})$ be a commutative ring with identity element $\mathbf{1}$. Then, the quintuple $(F, +, \times, \mathbf{0}, \mathbf{1})$ is said to be a field if $(F \setminus \{\mathbf{0}\}, \times, \mathbf{1})$ form an abelian group.

Proposition 6.15 Let $(F, +, \times, \mathbf{0}, \mathbf{1})$ be a field, $A \subseteq F$, and $\mathbf{0}, \mathbf{1} \in A$. Then, $(A, +, \times, \mathbf{0}, \mathbf{1})$ is a field (which will be called a subfield) if, and only if, $\forall a_1, a_2 \in A$, $a_1 + (-a_2) \in A$, and $a_1 \times a_2^{-1} \in A$ when $a_2 \neq \mathbf{0}$, where a_2^{-1} denotes the multiplicative inverse of a_2 .

Proof “Necessity” This is straightforward.

“Sufficiency” By Proposition 6.2, $(A, +, \mathbf{0})$ is a group. Since $(F, +, \mathbf{0})$ is abelian, then $(A, +, \mathbf{0})$ is also an abelian group. $\forall a_1 \in A$ with $a_1 \neq \mathbf{0}$. By the assumption of the proposition, we have $\mathbf{1} \in A$, by the property of field, we have $\mathbf{1} \neq \mathbf{0}$. Then, $a_1^{-1} = \mathbf{1} \times a_1^{-1} \in A$. $\forall a_1, a_2 \in A$. If $a_2 = \mathbf{0}$, then, by Proposition 6.11, $a_1 \times a_2 = \mathbf{0} \in A$. On the other hand, if $a_2 \neq \mathbf{0}$, we have $a_2^{-1} \in A$ and $a_1 \times a_2 = a_1 \times (a_2^{-1})^{-1} \in A$. Hence, by Proposition 6.10, $(A, +, \times, \mathbf{0})$ is a ring. Since $(F, +, \times, \mathbf{0})$ is a commutative ring with identity element $\mathbf{1}$ and $\mathbf{1} \in A$, then $(A, +, \times, \mathbf{0})$ is also a commutative ring with identity element $\mathbf{1}$. Note that $\mathbf{1} \in A \setminus \{\mathbf{0}\}$, then, $A \setminus \{\mathbf{0}\} \neq \emptyset$ and $A \setminus \{\mathbf{0}\} \subseteq F \setminus \{\mathbf{0}\}$. $\forall a_1, a_2 \in A \setminus \{\mathbf{0}\}$, $a_1 \times a_2^{-1} \in A$. We claim that $a_1 \times a_2^{-1} \neq \mathbf{0}$ since, otherwise, $a_1 = a_1 \times \mathbf{1} = a_1 \times (a_2^{-1} \times a_2) = (a_1 \times a_2^{-1}) \times a_2 = \mathbf{0}$ by Proposition 6.11, which is a contradiction. Hence, $a_1 \times a_2^{-1} \in A \setminus \{\mathbf{0}\}$. Since $(F \setminus \{\mathbf{0}\}, \times, \mathbf{1})$ is an abelian group, then, by Proposition 6.2, $(A \setminus \{\mathbf{0}\}, \times, \mathbf{1})$ is also a group and is further abelian. Therefore, $(A, +, \times, \mathbf{0}, \mathbf{1})$ is a field.

This completes the proof of the proposition. $\square$

6.4 Vector Spaces

Associated with every vector space is a set of scalars. This set of scalars can be any algebraic field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$. Examples of fields are the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$. Here, we will abuse the notation to say $x \in \mathcal{F}$ when $x \in F$.

Definition 6.16 A vector space $\mathcal{X}$ over a field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$ is a set X of elements called vectors together with two operations $\oplus$ and $\otimes$. $\oplus : X \times X \rightarrow X$ is called vector addition. It associates any two vectors $x, y \in X$ with a vector $x \oplus y \in X$, the sum of x and y . $\otimes : F \times X \rightarrow X$ is called scalar multiplication. It associates a scalar $\alpha \in F$ and a vector $x \in X$ with a vector $\alpha \otimes x \in X$, the scalar multiple of x by α . Furthermore, the following properties hold for $\forall x, y, z \in X$ and $\forall \alpha, \beta \in F$

$$(i) \quad x \oplus y = y \oplus x; \quad (\text{commutative law})$$

$$(ii) \quad (x \oplus y) \oplus z = x \oplus (y \oplus z); \quad (\text{associative law})$$

- (iii) $\exists$ a null vector $\vartheta \in X$ such that $x \oplus \vartheta = x, \forall x \in X$;
- (iv) $\alpha \otimes (x \oplus y) = (\alpha \otimes x) \oplus (\alpha \otimes y) = \alpha \otimes x \oplus \alpha \otimes y$, where we neglected the parenthesis in the last equality since we assume that $\otimes$ takes precedence over $\oplus$; (distributive law)
- (v) $(\alpha + \beta) \otimes x = \alpha \otimes x \oplus \beta \otimes x$; (distributive law)
- (vi) $(\alpha\beta) \otimes x = \alpha \otimes (\beta \otimes x)$; (associative law)
- (vii) $\mathbf{0} \otimes x = \vartheta$ and $\mathbf{1} \otimes x = x$.

We thus denote the quadruple $(X, \oplus, \otimes, \vartheta)$ by $\mathcal{X}$. The vector space is denoted by $(\mathcal{X}, \mathcal{F})$.

For convenience, $(-\mathbf{1}) \otimes x =: \ominus x$ and called the negative of the vector x . Note that

$$(\ominus x) \oplus x = x \oplus (\ominus x) = \mathbf{1} \otimes x \oplus (-\mathbf{1}) \otimes x = (\mathbf{1} + (-\mathbf{1})) \otimes x = \mathbf{0} \otimes x = \vartheta$$

We will also denote $x \ominus y := x \oplus (\ominus y)$. We will abuse the notation to say $x \in \mathcal{X}$ when $x \in X$. Note that $(\mathcal{X}, \oplus, \vartheta)$ forms an abelian group.

Proposition 6.17 Let $\mathcal{X} := (X, \oplus, \otimes, \vartheta)$ be a vector space over the field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$. $\forall x, y, z \in \mathcal{X}, \forall \alpha, \beta \in \mathcal{F}$, we have

1. $x \oplus y = x \oplus z \Rightarrow y = z$; (cancellation law)
2. $\alpha \otimes x = \alpha \otimes y$ and $\alpha \neq \mathbf{0} \Rightarrow x = y$; (cancellation law)
3. $\alpha \otimes x = \beta \otimes x$ and $x \neq \vartheta \Rightarrow \alpha = \beta$; (cancellation law)
4. $(\alpha - \beta) \otimes x = \alpha \otimes x \ominus \beta \otimes x$; (distributive law)
5. $\alpha \otimes (x \ominus y) = \alpha \otimes x \ominus \alpha \otimes y$; (distributive law)
6. $\alpha \otimes \vartheta = \vartheta$.

We will call ϑ the origin.

Example 6.18 $X = \{\vartheta\}$ with $\vartheta \oplus \vartheta = \vartheta$ and $\alpha \otimes \vartheta = \vartheta, \forall \alpha \in \mathcal{F}$. Then, $(X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathcal{F}$. $\diamond$

Example 6.19 Let $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$ be a field. Then, $(F, +, \cdot, \mathbf{0})$ is a vector space over $\mathcal{F}$. We will abuse the notation and say that $\mathcal{F}$ is a vector space over $\mathcal{F}$. $\diamond$

Example 6.20 Let $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$ be a field, $\mathcal{Y} := (Y, \oplus_Y, \otimes_Y, \vartheta_Y)$ be a vector space over $\mathcal{F}$, and A be a set. $X = \{f : A \rightarrow \mathcal{Y}\}$, that is, X is the set of all $\mathcal{Y}$ -valued functions on A . Define vector addition and scalar multiplication by, $\forall x, y \in X, \forall \alpha \in \mathcal{F}, z_a := x \oplus y \in X$ is given by $z_a(u) = x(u) \oplus_Y y(u), \forall u \in A, z_s := \alpha \otimes x \in X$ is given by $z_s(u) = \alpha \otimes_Y x(u), \forall u \in A$. Let $\vartheta \in X$ be given by $\vartheta(u) = \vartheta_Y, \forall u \in A$.

Now, we will show that $\mathcal{X} := (X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathcal{F}$. $\forall x, y, z \in X, \forall \alpha, \beta \in \mathcal{F}, \forall u \in A$, we have

- (i) $(x \oplus y)(u) = x(u) \oplus_Y y(u) = y(u) \oplus_Y x(u) = (y \oplus x)(u) \Rightarrow x \oplus y = y \oplus x$;
- (ii) $((x \oplus y) \oplus z)(u) = (x \oplus y)(u) \oplus_Y z(u) = (x(u) \oplus_Y y(u)) \oplus_Y z(u) = x(u) \oplus_Y (y(u) \oplus_Y z(u)) = x(u) \oplus_Y (y \oplus z)(u) = (x \oplus (y \oplus z))(u) \Rightarrow (x \oplus y) \oplus z = x \oplus (y \oplus z)$;
- (iii) $(x \oplus \vartheta)(u) = x(u) \oplus_Y \vartheta_Y = x(u) \Rightarrow x \oplus \vartheta = x$;
- (iv) $(\alpha \otimes (x \oplus y))(u) = \alpha \otimes_Y (x \oplus y)(u) = \alpha \otimes_Y (x(u) \oplus_Y y(u)) = \alpha \otimes_Y x(u) \oplus_Y \alpha \otimes_Y y(u) = (\alpha \otimes x)(u) \oplus_Y (\alpha \otimes y)(u) = (\alpha \otimes x \oplus \alpha \otimes y)(u) \Rightarrow \alpha \otimes (x \oplus y) = \alpha \otimes x \oplus \alpha \otimes y$;
- (v) $((\alpha + \beta) \otimes x)(u) = (\alpha + \beta) \otimes_Y x(u) = \alpha \otimes_Y x(u) \oplus_Y \beta \otimes_Y x(u) = (\alpha \otimes x)(u) \oplus_Y (\beta \otimes x)(u) = (\alpha \otimes x \oplus \beta \otimes x)(u) \Rightarrow (\alpha + \beta) \otimes x = \alpha \otimes x \oplus \beta \otimes x$;
- (vi) $((\alpha \beta) \otimes x)(u) = (\alpha \beta) \otimes_Y x(u) = \alpha \otimes_Y (\beta \otimes_Y x(u)) = \alpha \otimes_Y (\beta \otimes x)(u) = (\alpha \otimes (\beta \otimes x))(u) \Rightarrow (\alpha \beta) \otimes x = \alpha \otimes (\beta \otimes x)$;
- (vii) $(\mathbf{0} \otimes x)(u) = \mathbf{0} \otimes_Y x(u) = \vartheta_Y = \vartheta(u) \Rightarrow \mathbf{0} \otimes x = \vartheta$; $(\mathbf{1} \otimes x)(u) = \mathbf{1} \otimes_Y x(u) = x(u) \Rightarrow \mathbf{1} \otimes x = x$.

Therefore, $\mathcal{X}$ is a vector space over $\mathcal{F}$. This vector space will be denoted by $(\mathcal{M}(A, \mathcal{Y}), \mathcal{F})$. $\diamond$

Example 6.21 Let $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$ be a field. $X = F^n$ with $n \in \mathbb{N}$. Define vector addition and scalar multiplication by $x \oplus y := (\xi_1 + \eta_1, \dots, \xi_n + \eta_n) \in X$, $\alpha \otimes x := (\alpha \xi_1, \dots, \alpha \xi_n) \in X$, $\forall x = (\xi_1, \dots, \xi_n), y = (\eta_1, \dots, \eta_n) \in X$, $\forall \alpha \in \mathcal{F}$. Let $\vartheta := (\mathbf{0}, \dots, \mathbf{0}) \in X$. Then, it is straightforward to check that $\mathcal{F}^n := (X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathcal{F}$. $\diamond$

Example 6.22 Let $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$ be a field. $X = F^{m \times n} := \{m \times n\text{-dimensional } F \text{ valued matrices}\}$ with $m, n \in \mathbb{N}$. Define vector addition and scalar multiplication by $x \oplus y := (\xi_{ij} + \eta_{ij})_{m \times n} \in X$, $\alpha \otimes x := (\alpha \xi_{ij})_{m \times n} \in X$, $\forall x = (\xi_{ij})_{m \times n}, y = (\eta_{ij})_{m \times n} \in X$, $\forall \alpha \in \mathcal{F}$. Let $\vartheta := (\mathbf{0})_{m \times n} \in X$. Then, it is straightforward to check that $\mathcal{F}^{m \times n} := (X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathcal{F}$. $\diamond$

Example 6.23 Let $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$ be a field. $X = \{(\xi_k)_{k=1}^\infty \mid \xi_k \in \mathcal{F}, \forall k \in \mathbb{N}\}$. Define vector addition and scalar multiplication by $x \oplus y := (\xi_k + \eta_k)_{k=1}^\infty \in X$, $\alpha \otimes x := (\alpha \xi_k)_{k=1}^\infty \in X$, $\forall x = (\xi_k)_{k=1}^\infty, y = (\eta_k)_{k=1}^\infty \in X$, $\forall \alpha \in \mathcal{F}$. Let $\vartheta := (\mathbf{0}, \mathbf{0}, \dots) \in X$. Then, it is straightforward to check that $(X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathcal{F}$. $\diamond$

6.5 Product Spaces

Proposition 6.24 Let $\mathcal{X} := (X, \oplus_X, \otimes_X, \vartheta_X)$ and $\mathcal{Y} := (Y, \oplus_Y, \otimes_Y, \vartheta_Y)$ be vector spaces over the field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$. The Cartesian product of $\mathcal{X}$ and $\mathcal{Y}$, denoted by $\mathcal{X} \times \mathcal{Y}$, is the quadruple $(X \times Y, \oplus, \otimes, (\vartheta_X, \vartheta_Y))$,

where the vector addition $\oplus : (X \times Y) \times (X \times Y) \rightarrow X \times Y$ and the scalar multiplication $\otimes : \mathcal{F} \times (X \times Y) \rightarrow X \times Y$ are given by, $\forall (x_1, y_1), (x_2, y_2) \in X \times Y$, $\forall \alpha \in \mathcal{F}$, $(x_1, y_1) \oplus (x_2, y_2) := (x_1 \oplus_X x_2, y_1 \oplus_Y y_2)$ and $\alpha \otimes (x_1, y_1) := (\alpha \otimes_X x_1, \alpha \otimes_Y y_1)$. Then, $(\mathcal{X} \times \mathcal{Y}, \mathcal{F})$ is a vector space.

Proof Let $\vartheta := (\vartheta_X, \vartheta_Y) \in X \times Y$. $\forall (x_1, y_1), (x_2, y_2), (x_3, y_3) \in X \times Y$, $\forall \alpha, \beta \in \mathcal{F}$.

- (i) $(x_1, y_1) \oplus (x_2, y_2) = (x_1 \oplus_X x_2, y_1 \oplus_Y y_2) = (x_2 \oplus_X x_1, y_2 \oplus_Y y_1) = (x_2, y_2) \oplus (x_1, y_1)$;
- (ii) $((x_1, y_1) \oplus (x_2, y_2)) \oplus (x_3, y_3) = (x_1 \oplus_X x_2, y_1 \oplus_Y y_2) \oplus (x_3, y_3) = ((x_1 \oplus_X x_2) \oplus_X x_3, (y_1 \oplus_Y y_2) \oplus_Y y_3) = (x_1 \oplus_X (x_2 \oplus_X x_3), y_1 \oplus_Y (y_2 \oplus_Y y_3)) = (x_1, y_1) \oplus (x_2 \oplus_X x_3, y_2 \oplus_Y y_3) = (x_1, y_1) \oplus ((x_2, y_2) \oplus (x_3, y_3))$;
- (iii) $(x_1, y_1) \oplus \vartheta = (x_1 \oplus_X \vartheta_X, y_1 \oplus_Y \vartheta_Y) = (x_1, y_1)$;
- (iv) $\alpha \otimes ((x_1, y_1) \oplus (x_2, y_2)) = \alpha \otimes (x_1 \oplus_X x_2, y_1 \oplus_Y y_2) = (\alpha \otimes_X (x_1 \oplus_X x_2), \alpha \otimes_Y (y_1 \oplus_Y y_2)) = (\alpha \otimes_X x_1 \oplus_X \alpha \otimes_X x_2, \alpha \otimes_Y y_1 \oplus_Y \alpha \otimes_Y y_2) = (\alpha \otimes_X x_1, \alpha \otimes_Y y_1) \oplus (\alpha \otimes_X x_2, \alpha \otimes_Y y_2) = \alpha \otimes (x_1, y_1) \oplus \alpha \otimes (x_2, y_2)$;
- (v) $(\alpha + \beta) \otimes (x_1, y_1) = ((\alpha + \beta) \otimes_X x_1, (\alpha + \beta) \otimes_Y y_1) = (\alpha \otimes_X x_1 \oplus_X \beta \otimes_X x_1, \alpha \otimes_Y y_1 \oplus_Y \beta \otimes_Y y_1) = (\alpha \otimes_X x_1, \alpha \otimes_Y y_1) \oplus (\beta \otimes_X x_1, \beta \otimes_Y y_1) = \alpha \otimes (x_1, y_1) \oplus \beta \otimes (x_1, y_1)$;
- (vi) $(\alpha\beta) \otimes (x_1, y_1) = ((\alpha\beta) \otimes_X x_1, (\alpha\beta) \otimes_Y y_1) = (\alpha \otimes_X (\beta \otimes_X x_1), \alpha \otimes_Y (\beta \otimes_Y y_1)) = \alpha \otimes (\beta \otimes_X x_1, \beta \otimes_Y y_1) = \alpha \otimes (\beta \otimes (x_1, y_1))$;
- (vii) $\mathbf{0} \otimes (x_1, y_1) = (\mathbf{0} \otimes_X x_1, \mathbf{0} \otimes_Y y_1) = (\vartheta_X, \vartheta_Y) = \vartheta$; $\mathbf{1} \otimes (x_1, y_1) = (\mathbf{1} \otimes_X x_1, \mathbf{1} \otimes_Y y_1) = (x_1, y_1)$.

Hence, $\mathcal{X} \times \mathcal{Y}$ is a vector space over $\mathcal{F}$. □

With the above definition, it is easy to generalize to $\mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_n$, where $n \in \mathbb{N}$. We will also write $\mathcal{X}^n = \underbrace{\mathcal{X} \times \cdots \times \mathcal{X}}_n$, where $n \in \mathbb{N}$. When $n = 0$, $\prod_{i=1}^n \mathcal{X}_i$ is given by the vector space defined in Example 6.18.

6.6 Subspaces

Proposition 6.25 Let $\mathcal{X} := (X, \oplus, \otimes, \vartheta)$ be a vector space over the field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$, and $M \subseteq X$ with $M \neq \emptyset$. Then, $\mathcal{M} := (M, \oplus, \otimes, \vartheta)$ is a vector space over $\mathcal{F}$ (which will be called a subspace of $(\mathcal{X}, \mathcal{F})$) if, and only if, $\forall x, y \in M$, $\forall \alpha, \beta \in \mathcal{F}$, we have $\alpha \otimes x \oplus \beta \otimes y \in M$. We will also abuse the notation to say M is a subspace of $(\mathcal{X}, \mathcal{F})$. $\mathcal{M}$ is said to be a proper subspace of $(\mathcal{X}, \mathcal{F})$ if $M \subset X$.

Proof “If” Since $M \neq \emptyset$, then $\exists x_0 \in M$. $\vartheta = \vartheta \oplus \vartheta = \mathbf{0} \otimes x_0 \oplus \mathbf{0} \otimes x_0 \in M$. $\forall x, y \in M, \forall \alpha, \beta \in \mathcal{F}$. $x \oplus y = \mathbf{1} \otimes x \oplus \mathbf{1} \otimes y \in M$. Hence, M is closed under vector addition. $\alpha \otimes x = (\alpha + \mathbf{0}) \otimes x = \alpha \otimes x \oplus \mathbf{0} \otimes x \in M$. Hence, M is closed under scalar multiplication. Then, it is straightforward to show that $\mathcal{M}$ is a vector space over $\mathcal{F}$. “Only if” This is straightforward. This completes the proof of the proposition. $\square$

Example 6.26 We present the following list of examples of subspaces.

1. Let $(\mathcal{X}, \mathcal{F})$ be a vector space. Then, the singleton set $M = \{\vartheta\}$ is a subspace.
2. Consider the vector space $(\mathbb{R}^3, \mathbb{R})$. Any straight line or plane that passes through the origin is a subspace.
3. Consider the vector space $(\mathbb{R}^n, \mathbb{R})$, $n \in \mathbb{N}$. Let $a \in \mathbb{R}^n$. The set $M := \{x \in \mathbb{R}^n \mid \langle a, x \rangle = 0\}$ is a subspace.
4. Let $X := \{(\xi_k)_{k=1}^\infty \mid \xi_k \in \mathbb{R}, k \in \mathbb{N}\}$, $\oplus$ and $\otimes$ be the usual addition and scalar multiplication, and $\vartheta = (0, 0, \dots)$. By Example 6.23, $\mathcal{X} := (X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathbb{R}$. Let $M := \{(\xi_k)_{k=1}^\infty \in \mathcal{X} \mid \lim_{k \in \mathbb{N}} \xi_k \in \mathbb{R}\}$ is a subspace.
5. Let $X := \{f : (0, 1] \rightarrow \mathbb{R}^n\}$, $\oplus$ and $\otimes$ be the usual addition and scalar multiplication, and $\vartheta : (0, 1] \rightarrow \mathbb{R}^n$ be given by $\vartheta(t) = \vartheta_{\mathbb{R}^n}$, $\forall t \in (0, 1]$. By Example 6.20, $\mathcal{X} := (X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathbb{R}$. Let $M := \{f \in \mathcal{X} \mid f \text{ is continuous}\}$ is a subspace.

$\diamond$

To simplify notation in the theory, we will later simply discuss a vector space $(\mathcal{X}, \mathcal{F})$ without further reference to components of $\mathcal{X}$, where the operations are understood to be $\oplus_{\mathcal{X}}$ and $\otimes_{\mathcal{X}}$ and the null vector is understood to be $\vartheta_{\mathcal{X}}$. When it is clear from the context, we will neglect the subscript $\mathcal{X}$. Also, we will write $x_1 + x_2$ for $x_1 \oplus x_2$ and αx_1 for $\alpha \otimes x_1$, $\forall x_1, x_2 \in \mathcal{X}$, $\forall \alpha \in \mathcal{F}$.

Definition 6.27 Let $\mathcal{X}$ be a vector space over the field $\mathcal{F}$. $f : \mathcal{X} \rightarrow \mathcal{F}$ is said to be a functional.

Definition 6.28 Let $\mathcal{X}$ and $\mathcal{Y}$ be vector spaces over the field $\mathcal{F}$. $A : \mathcal{X} \rightarrow \mathcal{Y}$ is said to be linear if $\forall x_1, x_2 \in \mathcal{X}, \forall \alpha, \beta \in \mathcal{F}$, $A(\alpha x_1 + \beta x_2) = \alpha A(x_1) + \beta A(x_2)$. Then, A is called a (vector space) homomorphism or a linear operator. Furthermore, if it is bijective, then A is said to be a (vector space) isomorphism. The null space of A is $\mathcal{N}(A) := \{x \in \mathcal{X} \mid A(x) = \vartheta_{\mathcal{Y}}\}$. The range space of A is $\mathcal{R}(A) := \text{range}(A)$. $B : \mathcal{X} \rightarrow \mathcal{Y}$ is said to be an affine operator if $B(x) = A(x) + y_0$, $\forall x \in \mathcal{X}$, where $A : \mathcal{X} \rightarrow \mathcal{Y}$ is a linear operator and $y_0 \in \mathcal{Y}$.

Example 6.29 A row vector $v \in \mathbb{R}^{1 \times n}$ is a linear functional on $\mathbb{R}^n$. A matrix $A \in \mathbb{R}^{m \times n}$ is a linear function of $\mathbb{R}^n$ to $\mathbb{R}^m$. $\diamond$

For linear operators, we will adopt the following convention. Let $A : \mathcal{X} \rightarrow \mathcal{Y}$, $B : \mathcal{Y} \rightarrow \mathcal{Z}$, and $x \in \mathcal{X}$, where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are vector spaces and A and B are linear operators. We will write Ax for $A(x)$ and BA for $B \circ A$. Clearly, $\mathcal{N}(A)$ is a subspace of $\mathcal{X}$ and $\mathcal{R}(A)$ is a subspace of $\mathcal{Y}$. When A and B are bijective, we will denote A_{inv} by A^{-1} and B_{inv} by B^{-1} , where A^{-1} and B^{-1} are linear operators, then BA is also bijective and $(BA)^{-1} = A^{-1}B^{-1}$.

Definition 6.30 Let $(\mathcal{X}, \mathcal{F})$ be a vector space, $\alpha \in \mathcal{F}$, and $S, T \subseteq \mathcal{X}$. The sets αS and $S + T$ are defined by

$$\alpha S := \{\alpha s \in \mathcal{X} \mid s \in S\}; \quad S + T := \{s + t \in \mathcal{X} \mid s \in S, t \in T\}$$

This concept is illustrated in Figure 6.1.

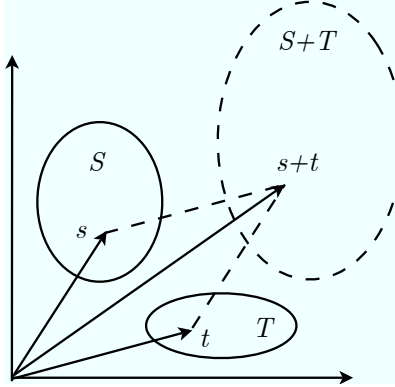


Figure 6.1: The sum of two sets.

We should note that $S + T = T + S$, $\emptyset + S = \emptyset$, $\{\vartheta\} + S = S$, and $\alpha\emptyset = \emptyset$. Thus, $S - T := S + (-T)$.

Proposition 6.31 Let M and N be subspaces of a vector space $(\mathcal{X}, \mathcal{F})$ and $\bar{\alpha} \in \mathcal{F}$. Then, $M \cap N$, $M + N$, and $\bar{\alpha}M$ are subspaces of $(\mathcal{X}, \mathcal{F})$.

Proof Since M and N are subspaces, then $\vartheta \in M$ and $\vartheta \in N$. Hence, $\vartheta \in M \cap N \neq \emptyset$, $\vartheta \in M + N \neq \emptyset$, and $\vartheta = \bar{\alpha}\vartheta \in \bar{\alpha}M \neq \emptyset$.

$\forall x, y \in M \cap N, \forall \alpha, \beta \in \mathcal{F}, \alpha x + \beta y \in M$ and $\alpha x + \beta y \in N$. Then, $\alpha x + \beta y \in M \cap N$. Hence, $M \cap N$ is a subspace.

$\forall x, y \in M + N, \forall \alpha, \beta \in \mathcal{F}$, we have $x = x_1 + x_2$ and $y = y_1 + y_2$, where $x_1, y_1 \in M$ and $x_2, y_2 \in N$. Then, $\alpha x_1 + \beta y_1 \in M$ and $\alpha x_2 + \beta y_2 \in N$. This implies that $\alpha x + \beta y = (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) \in M + N$. Hence, $M + N$ is a subspace.

$\forall x, y \in \bar{\alpha}M, \forall \alpha, \beta \in \mathcal{F}$, we have $x = \bar{\alpha}\bar{x}$ and $y = \bar{\alpha}\bar{y}$, where $\bar{x}, \bar{y} \in M$. Then, $\alpha x + \beta y = \alpha \bar{\alpha}\bar{x} + \beta \bar{\alpha}\bar{y} = \bar{\alpha}(\alpha\bar{x} + \beta\bar{y}) \in \bar{\alpha}M$. Hence, $\bar{\alpha}M$ is a subspace.

This completes the proof of the proposition. $\square$

We should note that $M \cup N$ is in general not a subspace.

Definition 6.32 A linear combination of vectors $x_1, \dots, x_n$, where $n \in \mathbb{Z}_+$, in a vector space $(\mathcal{X}, \mathcal{F})$ is a sum of the form $\sum_{i=1}^n \alpha_i x_i := \alpha_1 x_1 + \dots + \alpha_n x_n$, where $\alpha_1, \dots, \alpha_n \in \mathcal{F}$.

Note that $+$ is defined for two vectors. To sum n vectors, one must add two at a time. By the definition of vector space, the simplified notion is not ambiguous. When $n = 0$, we take the sum to be $\vartheta_{\mathcal{X}}$.

Definition 6.33 Let $(\mathcal{X}, \mathcal{F})$ be a vector space and $S \subseteq \mathcal{X}$.

$$\text{span}(S) := \{\text{linear combination of vectors in } S\}$$

is called the subspace generated by S .

Proposition 6.34 Let $(\mathcal{X}, \mathcal{F})$ be a vector space and $S \subseteq \mathcal{X}$. Then, $\text{span}(S)$ is the smallest subspace containing S .

Proof Clearly, $\vartheta \in \text{span}(S) \neq \emptyset$. $\forall x, y \in \text{span}(S), \forall \alpha, \beta \in \mathcal{F}$. It is easy to show that $\alpha x + \beta y \in \text{span}(S)$. Hence, $\text{span}(S)$ is a subspace. Clearly $\text{span}(S) \supseteq S$.

$\forall M \subseteq \mathcal{X}$ such that $S \subseteq M$ and M is a subspace. Clearly, $\vartheta \in M$ by the proof of Proposition 6.25. $\forall y$ equals to a linear combination of vectors in S . Then, $y = \sum_{i=1}^n \alpha_i x_i$ with $n \in \mathbb{Z}_+, x_1, \dots, x_n \in S$, and $\alpha_1, \dots, \alpha_n \in \mathcal{F}$. This implies that $x_1, \dots, x_n \in M$ and $y \in M$. Hence, $\text{span}(S) \subseteq M$.

Hence, $\text{span}(S)$ is the smallest subspace containing S . This completes the proof of the proposition. $\square$

Definition 6.35 Let $(\mathcal{X}, \mathcal{F})$ be a vector space, $M \subseteq \mathcal{X}$ be a subspace, and $x_0 \in \mathcal{X}$. Then, $V := \{x_0\} + M$ is called a linear variety.

The translation of a subspace is a linear variety. We will abuse the notation to write $x_0 + M$ for $\{x_0\} + M$. $\forall \bar{x} \in V, V - \bar{x} := V - \{\bar{x}\}$ is a subspace.

Definition 6.36 Let $(\mathcal{X}, \mathcal{F})$ be a vector space, $S \subseteq \mathcal{X}$, and $S \neq \emptyset$. The linear variety generated by S , denoted by $v(S)$, is defined as the intersection of all linear varieties in $\mathcal{X}$ that contain S .

Proposition 6.37 Let $(\mathcal{X}, \mathcal{F})$ be a vector space and $\emptyset \neq S \subseteq \mathcal{X}$. Then, $v(S)$ is a linear variety given by $v(S) = x_0 + \text{span}(S - x_0)$, where x_0 is any vector in S .

Proof Note that $S = S - x_0 + x_0 \subseteq x_0 + \text{span}(S - x_0)$ and $x_0 + \text{span}(S - x_0)$ is a linear variety. Hence, $v(S) \subseteq x_0 + \text{span}(S - x_0)$.

$\forall V \subseteq \mathcal{X}$ such that $S \subseteq V$ and V is a linear variety. Then, $x_0 \in V$ and $V - x_0$ is a subspace. Clearly, $S - x_0 \subseteq V - x_0$, which implies that, by Proposition 6.34, $\text{span}(S - x_0) \subseteq V - x_0$. Therefore, $x_0 + \text{span}(S - x_0) \subseteq V$. Hence, $x_0 + \text{span}(S - x_0) \subseteq v(S)$. Therefore, $v(S) = x_0 + \text{span}(S - x_0)$.

This completes the proof of the proposition. $\square$

6.7 Convex Sets

Denote $\mathbb{K}$ to be either $\mathbb{R}$ or $\mathbb{C}$.

Definition 6.38 Let $(\mathcal{X}, \mathbb{K})$ be a vector space and $C \subseteq \mathcal{X}$. C is said to be convex if, $\forall x_1, x_2 \in C, \forall \alpha \in [0, 1] \subset \mathbb{R}$, we have $\alpha x_1 + (1 - \alpha)x_2 \in C$.

Subspaces and linear varieties are convex, so is $\emptyset$.

Proposition 6.39 Let $(\mathcal{X}, \mathbb{K})$ be a vector space and $K, G \subseteq \mathcal{X}$ be convex sets. Then,

1. λK is convex, $\forall \lambda \in \mathbb{K}$;
2. $K + G$ is convex.

Proof $\forall \lambda \in \mathbb{K}, \forall x_1, x_2 \in \lambda K, \forall \alpha \in [0, 1] \subset \mathbb{R}, \exists k_1, k_2 \in K$ such that $x_1 = \lambda k_1$ and $x_2 = \lambda k_2$. Then, $\alpha x_1 + (1 - \alpha)x_2 = \alpha \lambda k_1 + (1 - \alpha)\lambda k_2 = \lambda(\alpha k_1 + (1 - \alpha)k_2)$. Since K is convex, then $\alpha k_1 + (1 - \alpha)k_2 \in K$. This implies that $\alpha x_1 + (1 - \alpha)x_2 \in \lambda K$. Hence, λK is convex.

$\forall x_1, x_2 \in K + G, \forall \alpha \in [0, 1] \subset \mathbb{R}, \exists k_1, k_2 \in K$ and $\exists g_1, g_2 \in G$ such that $x_i = k_i + g_i, i = 1, 2$. Note that

$$\begin{aligned} \alpha x_1 + (1 - \alpha)x_2 &= \alpha(k_1 + g_1) + (1 - \alpha)(k_2 + g_2) \\ &= \alpha k_1 + \alpha g_1 + (1 - \alpha)k_2 + (1 - \alpha)g_2 \\ &= (\alpha k_1 + (1 - \alpha)k_2) + (\alpha g_1 + (1 - \alpha)g_2) \end{aligned}$$

Since K and G are convex, then $(\alpha k_1 + (1 - \alpha)k_2) \in K$ and $(\alpha g_1 + (1 - \alpha)g_2) \in G$. This implies that $\alpha x_1 + (1 - \alpha)x_2 \in K + G$. Hence, $K + G$ is convex.

This completes the proof of the proposition. $\square$

Proposition 6.40 Let $(\mathcal{X}, \mathbb{K})$ be a vector space and $\{C_\lambda\}_{\lambda \in \Lambda}$ be a collection of convex subsets of $\mathcal{X}$. Then, $C := \bigcap_{\lambda \in \Lambda} C_\lambda$ is convex.

Proof $\forall x_1, x_2 \in C, \forall \alpha \in [0, 1] \subset \mathbb{R}, \forall \lambda \in \Lambda, x_1, x_2 \in C_\lambda$. Since C_λ is convex, then $\alpha x_1 + (1 - \alpha)x_2 \in C_\lambda$. This implies that $\alpha x_1 + (1 - \alpha)x_2 \in C$. Hence, C is convex. This completes the proof of the proposition. $\square$

Definition 6.41 Let $(\mathcal{X}, \mathbb{K})$ be a vector space and $S \subseteq \mathcal{X}$. The convex hull generated by S , denoted by $\text{co}(S)$, is the smallest convex set containing S .

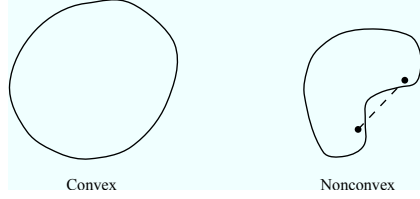


Figure 6.2: Convex and nonconvex sets.

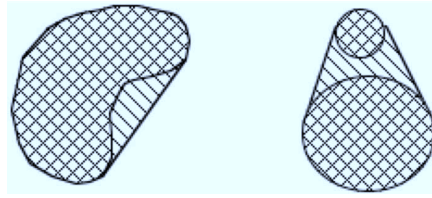


Figure 6.3: Convex hulls.

Justification of the existence of convex hull rests with Proposition 6.40.

Definition 6.42 Let $(\mathcal{X}, \mathbb{K})$ be a vector space. A convex combination of vectors $x_1, \dots, x_n \in \mathcal{X}$, where $n \in \mathbb{N}$, is a linear combination $\sum_{i=1}^n \alpha_i x_i$ with $\alpha_i \in [0, 1] \subset \mathbb{R}$, $\forall i = 1, \dots, n$, and $\sum_{i=1}^n \alpha_i = 1$.

Proposition 6.43 Let $(\mathcal{X}, \mathbb{K})$ be a vector space and $S \subseteq \mathcal{X}$. Then,

$$\text{co}(S) = \{\text{convex combinations of vectors in } S\}$$

Proof We need the follow result.

Claim 6.43.1 Let $G \subseteq \mathcal{X}$ be a convex subset. Then any convex combination of vectors in G belongs to G .

Proof of claim: We need to show: $\forall n \in \mathbb{N}$, $\forall x_1, \dots, x_n \in G$, $\forall \alpha_1, \dots, \alpha_n \in \mathbb{R}$, such that $\alpha_i \in [0, 1] \subset \mathbb{R}$, $i = 1, \dots, n$, $\sum_{i=1}^n \alpha_i = 1$ implies $\sum_{i=1}^n \alpha_i x_i \in G$. We will prove this by mathematical induction on n .

1° Consider $n = 1$. Then, $\alpha_1 = 1$ and $\sum_{i=1}^n \alpha_i x_i = x_1 \in G$. The result holds.

2° Assume that the result holds for $n = k \in \mathbb{N}$.

3° Consider the case $n = k+1$. Without loss of generality, assume $\alpha_1 > 0$.

By the induction hypothesis, we have $\sum_{i=1}^k \frac{\alpha_i}{\alpha_1 + \dots + \alpha_k} x_i \in G$. Then,

$$\sum_{i=1}^{k+1} \alpha_i x_i = (\alpha_1 + \dots + \alpha_k) \sum_{i=1}^k \frac{\alpha_i}{\alpha_1 + \dots + \alpha_k} x_i + \alpha_{k+1} x_{k+1}$$

$$= \alpha_{k+1}x_{k+1} + (1 - \alpha_{k+1}) \sum_{i=1}^k \frac{\alpha_i}{\alpha_1 + \dots + \alpha_k} x_i$$

By the convexity of G , we have $\sum_{i=1}^{k+1} \alpha_i x_i \in G$. Hence, the result holds for $n = k + 1$.

This completes the induction process and the proof of the claim. $\square$

$\forall x_1, x_2 \in K := \{\text{convex combinations of vectors in } S\}, \forall \alpha \in [0, 1] \subset \mathbb{R}$. By the definition of K , $\forall i = 1, 2, \exists n_i \in \mathbb{N}, \exists y_{i,1}, \dots, y_{i,n_i} \in S, \exists \alpha_{i,1}, \dots, \alpha_{i,n_i} \in [0, 1] \subset \mathbb{R}$, such that $\sum_{j=1}^{n_i} \alpha_{i,j} = 1$ and $x_i = \sum_{j=1}^{n_i} \alpha_{i,j} y_{i,j}$. Then,

$$\begin{aligned} \alpha x_1 + (1 - \alpha) x_2 &= \alpha \sum_{j=1}^{n_1} \alpha_{1,j} y_{1,j} + (1 - \alpha) \sum_{j=1}^{n_2} \alpha_{2,j} y_{2,j} \\ &= \sum_{j=1}^{n_1} \alpha \alpha_{1,j} y_{1,j} + \sum_{j=1}^{n_2} (1 - \alpha) \alpha_{2,j} y_{2,j} \end{aligned}$$

Note that $\alpha \alpha_{1,j} \geq 0, j = 1, \dots, n_1$, and $(1 - \alpha) \alpha_{2,j} \geq 0, j = 1, \dots, n_2$, and $\sum_{j=1}^{n_1} \alpha \alpha_{1,j} + \sum_{j=1}^{n_2} (1 - \alpha) \alpha_{2,j} = \alpha \sum_{j=1}^{n_1} \alpha_{1,j} + (1 - \alpha) \sum_{j=1}^{n_2} \alpha_{2,j} = \alpha + (1 - \alpha) = 1$.

Hence, $\alpha x_1 + (1 - \alpha) x_2 \in K$. This shows that K is convex. Clearly, $S \subseteq K$.

On the other hand, fix any convex set G in the vector space, satisfying $S \subseteq G$. $\forall p \in K$, by Claim 6.43.1, $p \in G$ since $S \subseteq G$. Then, $K \subseteq G$.

The above implies that K is the smallest convex set containing S . Hence, $K = \text{co}(S)$. This completes the proof of the proposition. $\square$

Definition 6.44 Let $(\mathcal{X}, \mathbb{K})$ be a vector space and $C \subseteq \mathcal{X}$. C is said to be a cone with vertex at origin if $\vartheta \in C$ and, $\forall x \in C, \forall \alpha \in [0, \infty) \subset \mathbb{R}$, we have $\alpha x \in C$. C is said to be a cone with vertex $p \in \mathcal{X}$ if $C = p + D$, where D is a cone with vertex at origin. C is said to be a conic segment if $\vartheta \in C$ and, $\forall x \in C, \forall \alpha \in [0, 1] \subset \mathbb{R}$, we have $\alpha x \in C$.

If vertex is not explicitly mentioned, it is assumed to be at origin.

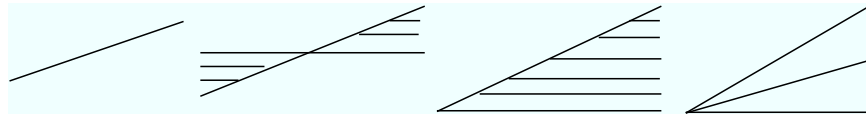


Figure 6.4: Cones.

Convex cones: arises in connection with positive vectors. In $\mathbb{R}^n$ with $n \in \mathbb{N}$, the positive cone may be defined as

$$P = \{x = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n \mid \xi_i \geq 0, i = 1, \dots, n\}$$

6.8 Linear Independence and Dimensions

Definition 6.45 Let $(\mathcal{X}, \mathcal{F})$ be a vector space, $x \in \mathcal{X}$, and $S \subseteq \mathcal{X}$. The vector x is said to be linearly dependent upon S if $x \in \text{span}(S)$. Otherwise, x is said to be linearly independent of S . S is said to be a linearly independent set if, $\forall y \in S$, y is linearly independent of $S \setminus \{y\}$.

Note that $\emptyset$ is a linearly independent set; $\{x\}$ is a linearly independent set if, and only if, $x \neq \vartheta$; and $\{x_1, x_2\}$ is a linearly independent set if, and only if, x_1 and x_2 do not lie on a common line through the origin.

Theorem 6.46 Let $\mathcal{X}$ be a vector space over the field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$ and $S \subseteq \mathcal{X}$. Then, S is a linearly independent set if, and only if, $\forall n \in \mathbb{N}$, $\forall \alpha_1, \dots, \alpha_n \in \mathcal{F}$, $\forall x_1, \dots, x_n \in \mathcal{X}$ which are distinct, we have $\sum_{i=1}^n \alpha_i x_i = \vartheta$ implies that $\alpha_i = \mathbf{0}$, $i = 1, \dots, n$.

Proof “Sufficiency” We will prove it using an argument of contradiction. Suppose S is not a linearly independent set. Then, $\exists y \in S$ such that y is linearly dependent upon $S \setminus \{y\}$. So, $y \in \text{span}(S \setminus \{y\})$. $\exists n \in \mathbb{N}$, $\exists x_2, \dots, x_n \in S \setminus \{y\}$, and $\exists \alpha_2, \dots, \alpha_n \in \mathcal{F}$ such that $y = \sum_{i=2}^n \alpha_i x_i$ (when $n = 1$, then $y = \vartheta$). Without loss of generality, we may assume that $x_2, \dots, x_n$ are distinct. Let $x_1 = y$ and $\alpha_1 = -\mathbf{1} \neq \mathbf{0}$. Then, we have $\sum_{i=1}^n \alpha_i x_i = \vartheta$ with $\alpha_1 \neq \mathbf{0}$ and $x_1, \dots, x_n$ are distinct. This is a contradiction. Hence, the sufficiency result holds.

“Necessity” We again prove this by an argument of contradiction. Suppose the result does not hold. $\exists n \in \mathbb{N}$, $\exists \alpha_1, \dots, \alpha_n \in \mathcal{F}$, and $\exists x_1, \dots, x_n \in S$ which are distinct such that $\sum_{i=1}^n \alpha_i x_i = \vartheta$ and $\exists i_0 \in \{1, \dots, n\}$ such that $\alpha_{i_0} \neq \mathbf{0}$. Without loss of generality, we may assume $i_0 = 1$. Then, we have $\alpha_1 x_1 = -\sum_{i=2}^n \alpha_i x_i$. Hence, $x_1 = \sum_{i=2}^n (-\alpha_1^{-1} \alpha_i) x_i$. Note that $x_i \neq x_1$ implies that $x_i \in S \setminus \{x_1\}$, $i = 2, \dots, n$. Hence, x_1 is linearly dependent upon $S \setminus \{x_1\}$. This is a contradiction. Then, the necessity result holds.

This completes the proof of this theorem. $\square$

Corollary 6.47 Let $\mathcal{X}$ be a vector space over the field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$, $\{x_1, \dots, x_n\} \subseteq \mathcal{X}$ be a linearly independent set, x_i 's are distinct, $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathcal{F}$, and $n \in \mathbb{Z}_+$. Assume that $\sum_{i=1}^n \alpha_i x_i = \sum_{i=1}^n \beta_i x_i$. Then, $\alpha_i = \beta_i$, $i = 1, \dots, n$.

Proof The assumption implies $\sum_{i=1}^n (\alpha_i - \beta_i) x_i = \vartheta$. When $n = 0$, clearly the result holds. When $n \in \mathbb{N}$, by Theorem 6.46, we have $\alpha_i - \beta_i = \mathbf{0}$, $i = 1, \dots, n$. This completes the proof of the corollary. $\square$

Definition 6.48 Let $\mathcal{X}$ be a vector space over the field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$, $n \in \mathbb{Z}_+$, $x_1, \dots, x_n \in \mathcal{X}$. The vectors $x_1, \dots, x_n$ are linearly independent if, $\forall \alpha_1, \dots, \alpha_n \in \mathcal{F}$, $\sum_{i=1}^n \alpha_i x_i = \vartheta$ implies $\alpha_1 = \dots = \alpha_n = \mathbf{0}$. Otherwise, these vectors are said to be linearly dependent.

Lemma 6.49 *Let $\mathcal{X}$ be a vector space over the field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$. Then,*

1. $x_1, \dots, x_n \in \mathcal{X}$ are linearly independent, where $n \in \mathbb{Z}_+$, if, and only if, they are distinct and the set $\{x_1, \dots, x_n\}$ is a linearly independent set.
2. $S \subseteq \mathcal{X}$ is a linearly independent set if, and only if, $\forall n \in \mathbb{N}$, $\forall x_1, \dots, x_n \in S$ which are distinct implies that $x_1, \dots, x_n$ are linearly independent.

Proof 1. “Sufficiency” When $n = 0$, $x_1, \dots, x_n$ are linearly independent and $\emptyset$ is a linearly independent set. Hence, the result holds. When $n \in \mathbb{N}$, this is straightforward from Theorem 6.46. “Necessity” We will prove this using an argument of contradiction. Suppose the result does not hold. We will distinguish two exhaustive cases: Case 1: $x_1, \dots, x_n$ are not distinct; Case 2: the set $\{x_1, \dots, x_n\}$ is not a linearly independent set. Case 1. Without loss of generality, assume $x_1 = x_2$. Set $\alpha_1 = \mathbf{1}$, $\alpha_2 = -\mathbf{1}$ and the rest of α_i ’s to $\mathbf{0}$. Then, we have $\sum_{i=1}^n \alpha_i x_i = \vartheta$ and hence $x_1, \dots, x_n$ is linearly dependent. This is a contradiction. Case 2. By Theorem 6.46, $\exists m \in \mathbb{N}$, $\exists \alpha_1, \dots, \alpha_m \in \mathcal{F}$ which are not all $\mathbf{0}$ ’s, $\exists y_1, \dots, y_m \in \{x_1, \dots, x_n\}$ which are distinct, such that $\sum_{i=1}^m \alpha_i y_i = \vartheta$. Clearly, $m \leq n$, otherwise y_i ’s are not distinct. Without loss of generality, assume $y_1 = x_1, \dots, y_m = x_m$. Set $\alpha_{m+1} = \dots = \alpha_n = \mathbf{0}$. Then, we have $\sum_{i=1}^n \alpha_i x_i = \vartheta$ with $\alpha_1, \dots, \alpha_n$ not all $\mathbf{0}$ ’s. This is a contradiction. Thus, we have arrived at a contradiction in every case. Hence, the necessity result holds.

2. This is straightforward from Theorem 6.46.

This completes the proof of the lemma. $\square$

Definition 6.50 *Let $(\mathcal{X}, \mathcal{F})$ be a vector space and $S \subseteq \mathcal{X}$ be a linearly independent set with $n \in \mathbb{Z}_+$ elements. S is said to be a basis of the vector space if $\text{span}(S) = \mathcal{X}$. In this case, the vector space is said to be finite dimensional with dimension n . All other vector spaces are said to be infinite dimensional.*

Theorem 6.51 *Let $\mathcal{X}$ be a finite dimensional vector space over the field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$. Then, the dimension $n \in \{0\} \cup \mathbb{N}$ is unique. Furthermore, any linearly independent set of n vectors form a basis of the vector space.*

Proof Let $n \in \mathbb{Z}_+$ be the minimum of dimensions for $\mathcal{X}$. Then, there exists a set $S_1 \subseteq \mathcal{X}$ with n elements such that S_1 is a basis for the vector space. We will show that $\forall y_1, \dots, y_m \in \mathcal{X}$ with $m > n$, then $y_1, \dots, y_m$ are linearly dependent. This implies that any subset with more than n elements cannot be a linearly independent set by Lemma 6.49. Henceforth, the dimension of the vector space is unique.

$\forall y_1, \dots, y_m \in \mathcal{X}$ with $m > n$. There are two exhaustive and mutually exclusive cases: Case 1: $n = 0$; Case 2: $n > 0$. Case 1: $n = 0$. Then, $S_1 = \emptyset$ and $\mathcal{X}$ contains a single vector ϑ . Hence, $y_1 = \dots = y_m = \vartheta$. Clearly, $y_1, \dots, y_m$ are linearly dependent. This case is proven. Case 2: $n > 0$. Take $S_1 = \{x_1, \dots, x_n\}$. Since S_1 is a basis for the vector space, then, $\forall i \in \{1, \dots, m\}$, $y_i \in \text{span}(S_1)$, that is, $\exists \alpha_{ij} \in \mathcal{F}$, $j = 1, \dots, n$, such that $y_i = \sum_{j=1}^n \alpha_{ij} x_j$. Consider the $m \times n$ -dimensional matrix $A := (\alpha_{ij})_{m \times n}$. Since $m > n$, then $\text{rank}(A) < m$, which implies that the row vectors of A are linearly dependent. $\exists \beta_1, \dots, \beta_m \in \mathcal{F}$, which are not all $\mathbf{0}$'s, such that $\sum_{i=1}^m \beta_i \alpha_{ij} = \mathbf{0}$, $j = 1, \dots, n$. This implies $\sum_{i=1}^m \beta_i y_i = \sum_{i=1}^m \sum_{j=1}^n \beta_i \alpha_{ij} x_j = \sum_{j=1}^n (\sum_{i=1}^m \beta_i \alpha_{ij}) x_j = \vartheta$. Hence, $y_1, \dots, y_m$ are linearly dependent. This case is also proven.

Hence, the dimension of the vector space is unique.

$\forall S_2 \subseteq \mathcal{X}$ with n vectors that is a linearly independent set, we will show that S_2 is a basis for the vector space. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $n = 0$; Case 2: $n > 0$. Case 1: $n = 0$. Then, $S_2 = \emptyset$ and $\mathcal{X}$ is a singleton set consisting of the null vector, which equals to $\text{span}(S_2)$. Hence, S_2 is a basis for the vector space. Case 2: $n > 0$. Take $S_2 = \{z_1, \dots, z_n\}$. $\forall x \in \mathcal{X}$, $x, z_1, \dots, z_n$ (for a total of $n + 1$ vectors), by the preceding proof, are linearly dependent. Then, $\exists \alpha, \beta_1, \dots, \beta_n \in \mathcal{F}$, which are not all $\mathbf{0}$'s, such that $\alpha x + \sum_{i=1}^n \beta_i z_i = \vartheta$. Suppose $\alpha = \mathbf{0}$. Then, $\sum_{i=1}^n \beta_i z_i = \vartheta$. Since S_2 is a linearly independent set, by Theorem 6.46, then $\beta_1 = \dots = \beta_n = \mathbf{0}$. This contradicts with the fact that $\alpha, \beta_1, \dots, \beta_n$ are not all $\mathbf{0}$'s. Hence, $\alpha \neq \mathbf{0}$. Then, $x = \sum_{i=1}^n (-\alpha^{-1} \beta_i) z_i$ and $x \in \text{span}(S_2)$. Therefore, we have $\mathcal{X} = \text{span}(S_2)$. Hence, S_2 is a basis of the vector space.

This completes the proof of the theorem. $\square$

Finite-dimensional spaces are simpler to analyze. Many results of finite-dimensional spaces may be generalized to infinite-dimensional spaces. We endeavor to stress the similarity between the finite- and infinite-dimensional spaces.

Chapter 7

Banach Spaces

7.1 Normed Linear Spaces

Vector spaces admit algebraic properties, but they lack topological properties.

Denote $\mathbb{K}$ to be either $\mathbb{R}$ or $\mathbb{C}$.

Definition 7.1 Let $(\mathcal{X}, \mathbb{K})$ be a vector space. A norm on the vector space is a real-valued function $\|\cdot\| : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$ that satisfies the following properties: $\forall x, y \in \mathcal{X}, \forall \alpha \in \mathbb{K}$,

$$(i) \quad 0 \leq \|x\| < +\infty \text{ and } \|x\| = 0 \Leftrightarrow x = \vartheta;$$

$$(ii) \quad \|x + y\| \leq \|x\| + \|y\|; \quad (\text{triangle inequality})$$

$$(iii) \quad \|\alpha x\| = |\alpha| \|x\|.$$

A real (complex) normed linear space is a vector space over the field $\mathbb{R}$ (or $\mathbb{C}$) together with a norm defined on it. A normed linear space consisting of the triple $(\mathcal{X}, \mathbb{K}, \|\cdot\|)$.

To simplify notation in the theory, we may later simply discuss a normed linear space $\mathcal{X} := (\mathcal{X}, \mathbb{K}, \|\cdot\|)$ without further reference to components of $\mathcal{X}$, where the operations are understood to be $\oplus_{\mathcal{X}}$ and $\otimes_{\mathcal{X}}$, the null vector is understood to be $\vartheta_{\mathcal{X}}$, and the norm is understood to be $\|\cdot\|_{\mathcal{X}}$. When it is clear from the context, we will neglect the subscript $\mathcal{X}$.

Example 7.2 Let $n \in \mathbb{Z}_+$. The Euclidean space $(\mathbb{R}^n, \mathbb{R}, |\cdot|)$ is a normed linear space, where the norm is defined by $|(\xi_1, \dots, \xi_n)| = \sqrt{\xi_1^2 + \dots + \xi_n^2}$, $\forall (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$. Note that we specifically denote the Euclidean norm as $|\cdot|$, rather than $\|\cdot\|$, to distinguish it from other norms on $\mathbb{R}^n$. $\diamond$

Example 7.3 Let $n \in \mathbb{Z}_+$. The space $(\mathbb{C}^n, \mathbb{C}, |\cdot|)$ is a normed linear space, where the norm is defined by $|(\xi_1, \dots, \xi_n)| = \sqrt{\sum_{i=1}^n |\xi_i|^2}$, $\forall (\xi_1, \dots, \xi_n) \in \mathbb{C}^n$.

$\mathbb{C}^n$. Note that we specifically denote the norm as $|\cdot|$, rather than $\|\cdot\|$, to distinguish it from other norms on $\mathbb{C}^n$. $\diamond$

Example 7.4 Let $X := \{f : [a, b] \rightarrow \mathbb{R}\}$, where $a, b \in \mathbb{R}$ with $a \leq b$, $\oplus$ and $\otimes$ be the usual vector addition and scalar multiplication, and $\vartheta : [a, b] \rightarrow \mathbb{R}$ be given by $\vartheta(t) = 0, \forall t \in [a, b]$. By Example 6.20, $\mathcal{X} := (X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathbb{R}$. Let $M := \{f \in \mathcal{X} \mid f \text{ is continuous}\}$ is a subspace by Proposition 6.25. Let $\mathcal{M} := (M, \oplus, \otimes, \vartheta)$. Then, $(\mathcal{M}, \mathbb{R})$ is a vector space. Introduce a norm on this space by $\|x\| = \max_{a \leq t \leq b} |x(t)|, \forall x \in \mathcal{M}$. Now, we verify the properties of the norm. $\forall x, y \in \mathcal{M}, \forall \alpha \in \mathbb{R}$.

- (i) Since $[a, b]$ is a nonempty compact set and x is continuous, then, $\|x\| = |x(t)| \in [0, \infty) \subset \mathbb{R}$, for some $t \in [a, b]$. Clearly, $\|x\| = 0 \Leftrightarrow x = \vartheta$.
- (ii) $\|x + y\| = \max_{a \leq t \leq b} |x(t) + y(t)| \leq \max_{a \leq t \leq b} (|x(t)| + |y(t)|) \leq \max_{a \leq t \leq b} |x(t)| + \max_{a \leq t \leq b} |y(t)| = \|x\| + \|y\|.$
- (iii) $\|\alpha x\| = \max_{a \leq t \leq b} |\alpha x(t)| = \max_{a \leq t \leq b} |\alpha| |x(t)| = |\alpha| \cdot \max_{a \leq t \leq b} |x(t)| = |\alpha| \|x\|$, where we have made use of Proposition 3.81 and the fact that $\|x\| \in \mathbb{R}$ in the third equality.

Hence, $\mathcal{C}([a, b]) := (\mathcal{M}, \mathbb{R}, \|\cdot\|)$ is a normed linear space. $\diamond$

Example 7.5 Let $X := \{f : [a, b] \rightarrow \mathbb{R}\}$, where $a, b \in \mathbb{R}$ with $a < b$, $\oplus$ and $\otimes$ be the usual vector addition and scalar multiplication, and $\vartheta : [a, b] \rightarrow \mathbb{R}$ be given by $\vartheta(t) = 0, \forall t \in [a, b]$. By Example 6.20, $\mathcal{X} := (X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathbb{R}$. Let $M := \{f \in \mathcal{X} \mid f \text{ is continuously differentiable}\}$ is a subspace by Proposition 6.25. Let $\mathcal{M} := (M, \oplus, \otimes, \vartheta)$. Then, $(\mathcal{M}, \mathbb{R})$ is a vector space. Introduce a norm on this space by $\|x\| = \max_{a \leq t \leq b} |x(t)| + \max_{a \leq t \leq b} |x^{(1)}(t)|, \forall x \in \mathcal{M}$. Now, we verify the properties of the norm. $\forall x, y \in \mathcal{M}, \forall \alpha \in \mathbb{R}$.

- (i) Since $[a, b]$ is a nonempty compact set and x is continuously differentiable, then, $\|x\| = |x(t_1)| + |x^{(1)}(t_2)| \in [0, \infty) \subset \mathbb{R}$, for some $t_1, t_2 \in [a, b]$. Clearly, $\|x\| = 0 \Leftrightarrow x = \vartheta$.
- (ii) $\|x + y\| = \max_{a \leq t \leq b} |x(t) + y(t)| + \max_{a \leq t \leq b} |x^{(1)}(t) + y^{(1)}(t)| \leq \max_{a \leq t \leq b} |x(t)| + \max_{a \leq t \leq b} |y(t)| + \max_{a \leq t \leq b} |x^{(1)}(t)| + \max_{a \leq t \leq b} |y^{(1)}(t)| = \|x\| + \|y\|.$
- (iii) $\|\alpha x\| = \max_{a \leq t \leq b} |\alpha x(t)| + \max_{a \leq t \leq b} |\alpha x^{(1)}(t)| = |\alpha| \|x\|.$

Hence, $\mathcal{C}_1([a, b]) := (\mathcal{M}, \mathbb{R}, \|\cdot\|)$ is a normed linear space. $\diamond$

Example 7.6 Let $X := \{(\xi_k)_{k=1}^\infty \mid \xi_k \in \mathbb{R}, \forall k \in \mathbb{N}\}$, $\oplus$ and $\otimes$ be the usual vector addition and scalar multiplication, and $\vartheta := (0, 0, \dots) \in X$. By Example 6.23, $\mathcal{X} := (X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathbb{R}$. For $p \in [1, \infty) \subset$

$\mathbb{R}$, let $M_p := \{(\xi_k)_{k=1}^\infty \in X \mid \sum_{k=1}^\infty |\xi_k|^p < +\infty\}$. Define the norm $\|x\|_p = (\sum_{k=1}^\infty |\xi_k|^p)^{1/p}$, $\forall x = (\xi_k)_{k=1}^\infty \in M_p$. Let $M_\infty := \{(\xi_k)_{k=1}^\infty \in X \mid \sup_{k \geq 1} |\xi_k| < +\infty\}$. Define the norm $\|x\|_\infty = \sup_{k \geq 1} |\xi_k|$, $\forall x = (\xi_k)_{k=1}^\infty \in M_\infty$. $\forall p \in [1, +\infty] \subset \mathbb{R}_e$, we will show that $\mathcal{M}_p := (M_p, \oplus, \otimes, \vartheta)$ is a subspace in $(\mathcal{X}, \mathbb{R})$ and $l_p := (\mathcal{M}_p, \mathbb{R}, \|\cdot\|_p)$ is a normed linear space.

$\forall p \in [1, +\infty] \subset \mathbb{R}_e$. $\vartheta \in l_p \neq \emptyset$. $\forall x, y \in \mathcal{M}_p$, $\forall \alpha, \beta \in \mathbb{R}$. It is easy to check that $\|\alpha x\|_p = |\alpha| \|x\|_p$. Then, by Theorem 7.9, $\|\alpha x + \beta y\|_p \leq \|\alpha x\|_p + \|\beta y\|_p = |\alpha| \|x\|_p + |\beta| \|y\|_p < +\infty$. This implies that $\alpha x + \beta y \in \mathcal{M}_p$. Hence, $\mathcal{M}_p$ is a subspace of $(\mathcal{X}, \mathbb{R})$. It is easy to check that $\|x\|_p \in [0, \infty) \subset \mathbb{R}$, $\forall x \in l_p$, and $\|x\|_p = 0 \Leftrightarrow x = \vartheta$. Therefore, l_p is a normed linear space. $\diamond$

Lemma 7.7 $\forall a, b \in [0, \infty) \subset \mathbb{R}$, $\forall \lambda \in (0, 1) \subset \mathbb{R}$, we have

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda)b$$

where equality holds if, and only if, $a = b$.

Proof Define $f : [0, \infty) \rightarrow \mathbb{R}$ by $f(t) = t^\lambda - \lambda t + \lambda - 1$, $\forall t \in [0, \infty) \subset \mathbb{R}$. Then, f is continuous, and is differentiable on $(0, \infty)$. $f^{(1)}(t) = \lambda t^{\lambda-1} - \lambda$, $\forall t \in (0, \infty)$. Then, $f^{(1)}(t) > 0$, $\forall t \in (0, 1)$ and $f^{(1)}(t) < 0$, $\forall t \in (1, \infty)$. Then, $f(t) \leq f(1) = 0$, $\forall t \in [0, \infty)$, where equality holds if, and only if, $t = 1$.

We will distinguish two exhaustive and mutually exclusive cases: Case 1: $b = 0$; Case 2: $b > 0$. Case 1: $b = 0$. We have $a^\lambda b^{1-\lambda} = 0 \leq \lambda a = \lambda a + (1-\lambda)b$, where equality holds if, and only if, $a = b = 0$. This case is proved. Case 2: $b > 0$. Since $f(a/b) \leq 0$, we immediately obtain the desired inequality, where equality holds if, and only if, $a/b = 1 \Leftrightarrow a = b$. This case is also proved.

This completes the proof of the lemma. $\square$

Theorem 7.8 (Hölder's Inequality) Let $p \in [1, +\infty) \subset \mathbb{R}$ and $q \in (1, +\infty] \subset \mathbb{R}_e$ with $1/p + 1/q = 1$. Then, $\forall x = (\xi_k)_{k=1}^\infty \in l_p$, $\forall y = (\eta_k)_{k=1}^\infty \in l_q$, we have

$$\sum_{k=1}^\infty |\xi_k \eta_k| \leq \|x\|_p \|y\|_q$$

When $q < \infty$, equality holds if, and only if, $\exists \alpha, \beta \in \mathbb{R}$, which are not both zeros, such that $\alpha |\xi_k|^p = \beta |\eta_k|^q$, $k = 1, 2, \dots$. When $q = \infty$, equality holds if, and only if, $|\eta_k| = \|y\|_\infty$ for any $k \in \mathbb{N}$ with $|\xi_k| > 0$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $q = \infty$; Case 2: $1 < q < +\infty$. Case 1: $q = \infty$. Then, $p = 1$. We have

$$\sum_{k=1}^\infty |\xi_k \eta_k| = \sum_{k=1}^\infty |\xi_k| |\eta_k| \leq \sum_{k=1}^\infty |\xi_k| \|y\|_\infty = \|x\|_1 \|y\|_\infty$$

where equality holds if, and only if, $|\eta_k| = \|y\|_\infty$ for any $k \in \mathbb{N}$ with $|\xi_k| > 0$. This case is proved.

Case 2: $1 < q < +\infty$. Then, $1 < p < +\infty$. We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: $\|x\|_p \|y\|_q = 0$; Case 2b: $\|x\|_p \|y\|_q > 0$. Case 2a: $\|x\|_p \|y\|_q = 0$. Without loss of generality, assume $\|y\|_q = 0$. Then, $y = \vartheta$ and

$$\sum_{k=1}^{\infty} |\xi_k \eta_k| = 0 = \|x\|_p \|y\|_q$$

Equality holds $\Rightarrow \alpha = 0$ and $\beta = 1$ and $\alpha |\xi_k|^p = \beta |\eta_k|^q$, $k = 1, 2, \dots$. This subcase is proved. Case 2b: $\|x\|_p \|y\|_q > 0$. Then, $\|x\|_p > 0$ and $\|y\|_q > 0$. $\forall k \in \mathbb{N}$, by Lemma 7.7 with $a = \left(\frac{|\xi_k|}{\|x\|_p}\right)^p$, $b = \left(\frac{|\eta_k|}{\|y\|_q}\right)^q$, and $\lambda = 1/p$, we have

$$\frac{|\xi_k \eta_k|}{\|x\|_p \|y\|_q} \leq \frac{1}{p} \frac{|\xi_k|^p}{\|x\|_p^p} + \frac{1}{q} \frac{|\eta_k|^q}{\|y\|_q^q}$$

with equality if, and only if, $\frac{|\xi_k|^p}{\|x\|_p^p} = \frac{|\eta_k|^q}{\|y\|_q^q}$. Summing the above inequality for all $k \in \mathbb{N}$, we have

$$\frac{\sum_{k=1}^{\infty} |\xi_k \eta_k|}{\|x\|_p \|y\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

with equality if, and only if, $\frac{|\xi_k|^p}{\|x\|_p^p} = \frac{|\eta_k|^q}{\|y\|_q^q}$, $k = 1, 2, \dots$. Equality $\Rightarrow \alpha = 1/\|x\|_p^p$ and $\beta = 1/\|y\|_q^q$ and $\alpha |\xi_k|^p = \beta |\eta_k|^q$, $k = 1, 2, \dots$. On the other hand, if $\exists \alpha, \beta \in \mathbb{R}$, which are not both zeros, such that $\alpha |\xi_k|^p = \beta |\eta_k|^q$, $k = 1, 2, \dots$, then, without loss of generality, assume $\beta \neq 0$. Let $\alpha_1 = \alpha/\beta$. Then, $\alpha_1 |\xi_k|^p = |\eta_k|^q$, $k = 1, 2, \dots$. Hence, $\alpha_1 \|x\|_p^p = \|y\|_q^q$, which further implies that $\alpha_1 = \|y\|_q^q / \|x\|_p^p$. Hence, $\frac{|\xi_k|^p}{\|x\|_p^p} = \frac{|\eta_k|^q}{\|y\|_q^q}$, $k = 1, 2, \dots$. This implies equality. Therefore, equality if, and only if, $\exists \alpha, \beta \in \mathbb{R}$, which are not both zeros, such that $\alpha |\xi_k|^p = \beta |\eta_k|^q$, $k = 1, 2, \dots$. This subcase is proved.

This completes the proof of the theorem. $\square$

When $p = 2 = q$, the Hölder's inequality becomes the well-known *Cauchy-Schwarz Inequality*:

$$\sum_{k=1}^{\infty} |\xi_k \eta_k| \leq \left(\sum_{k=1}^{\infty} |\xi_k|^2 \right)^{1/2} \left(\sum_{k=1}^{\infty} |\eta_k|^2 \right)^{1/2}$$

Theorem 7.9 (Minkowski's Inequality) *Let $p \in [1, \infty] \subset \mathbb{R}_e$. $\forall x, y \in l_p$, then, $x + y \in l_p$ and*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

When $1 < p < \infty$, equality holds if, and only if, $\exists \alpha, \beta \in [0, \infty) \subset \mathbb{R}$, which are not both zeros, such that $\alpha x = \beta y$.

Proof $\forall x = (\xi_k)_{k=1}^\infty, y = (\eta_k)_{k=1}^\infty \in l_p$. We will distinguish three exhaustive and mutually exclusive cases: Case 1: $p = 1$; Case 2: $p = \infty$; Case 3: $1 < p < \infty$.

Case 1: $p = 1$. $\|x + y\|_1 = \sum_{k=1}^\infty |\xi_k + \eta_k| \leq \sum_{k=1}^\infty (|\xi_k| + |\eta_k|) = \|x\|_1 + \|y\|_1 < +\infty$. Equality holds if, and only if, $\xi_k \eta_k \geq 0$, $k = 1, 2, \dots$. Hence, $x + y \in l_1$.

Case 2: $p = \infty$. $\|x + y\|_\infty = \sup_{k \in \mathbb{N}} |\xi_k + \eta_k| \leq \sup_{k \in \mathbb{N}} (|\xi_k| + |\eta_k|) \leq \sup_{k \in \mathbb{N}} |\xi_k| + \sup_{k \in \mathbb{N}} |\eta_k| = \|x\|_\infty + \|y\|_\infty < +\infty$. Hence, $x + y \in l_\infty$.

Case 3: $1 < p < \infty$. We will further distinguish two exhaustive and mutually exclusive cases: Case 3a: $\|x\|_p \|y\|_p = 0$; Case 3b: $\|x\|_p \|y\|_p > 0$. Case 3a: $\|x\|_p \|y\|_p = 0$. Without loss of generality, assume $\|y\|_p = 0$. Then, $y = \vartheta$ and $\eta_k = 0$, $k = 1, 2, \dots$. Hence, $x + y = x \in l_p$ and $\|x + y\|_p = \|x\|_p = \|x\|_p + \|y\|_p$. Let $\alpha = 0$ and $\beta = 1$, we have $\alpha x = \beta y$. Hence, equality holds if, and only if, $\exists \alpha, \beta \in [0, \infty) \subset \mathbb{R}$, which are not both zeros, such that $\alpha x = \beta y$. This subcase is proved.

Case 3b: $\|x\|_p \|y\|_p > 0$. Then, $\|x\|_p > 0$ and $\|y\|_p > 0$. Let $\lambda = \frac{\|x\|_p}{\|x\|_p + \|y\|_p} \in (0, 1) \subset \mathbb{R}$. $\forall k \in \mathbb{N}$.

$$\frac{|\xi_k + \eta_k|^p}{(\|x\|_p + \|y\|_p)^p} \leq \left(\frac{|\xi_k| + |\eta_k|}{\|x\|_p + \|y\|_p} \right)^p = \left(\lambda \frac{|\xi_k|}{\|x\|_p} + (1 - \lambda) \frac{|\eta_k|}{\|y\|_p} \right)^p$$

Since $1 < p < \infty$, then the function t^p is strictly convex on $[0, \infty) \subset \mathbb{R}$. Then, we have

$$\frac{|\xi_k + \eta_k|^p}{(\|x\|_p + \|y\|_p)^p} \leq \lambda \frac{|\xi_k|^p}{\|x\|_p^p} + (1 - \lambda) \frac{|\eta_k|^p}{\|y\|_p^p}$$

where equality holds $\Leftrightarrow \xi_k \eta_k \geq 0$ and $\frac{|\xi_k|}{\|x\|_p} = \frac{|\eta_k|}{\|y\|_p} \Leftrightarrow \frac{\xi_k}{\|x\|_p} = \frac{\eta_k}{\|y\|_p}$.

Summing the above inequalities for all $k \in \mathbb{N}$, we have

$$\begin{aligned} \frac{\|x + y\|_p^p}{(\|x\|_p + \|y\|_p)^p} &= \sum_{k=1}^\infty \frac{|\xi_k + \eta_k|^p}{(\|x\|_p + \|y\|_p)^p} \\ &\leq \lambda \frac{\sum_{k=1}^\infty |\xi_k|^p}{\|x\|_p^p} + (1 - \lambda) \frac{\sum_{k=1}^\infty |\eta_k|^p}{\|y\|_p^p} = 1 \end{aligned}$$

Therefore,

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p < +\infty$$

where equality holds $\Leftrightarrow \frac{\xi_k}{\|x\|_p} = \frac{\eta_k}{\|y\|_p}$, $k = 1, 2, \dots \Leftrightarrow (1/\|x\|_p)x = (1/\|y\|_p)y$. Equality implies that $\alpha = 1/\|x\|_p$ and $\beta = 1/\|y\|_p$ and $\alpha x = \beta y$. On the other hand, if $\exists \alpha, \beta \in [0, \infty) \subset \mathbb{R}$, which are not both zeros, such that $\alpha x = \beta y$, then, without loss of generality, assume $\beta \neq 0$. Then, $y = \alpha_1 x$ with $\alpha_1 = \alpha/\beta$. Easy to show that $\|y\|_p = \alpha_1 \|x\|_p$. Then, $\alpha_1 = \|y\|_p / \|x\|_p$. This implies that $(1/\|x\|_p)x = (1/\|y\|_p)y$, which further implies equality. Hence, equality holds if, and only if, $\exists \alpha, \beta \in [0, \infty)$, which are not both zeros, such that $\alpha x = \beta y$. Clearly, $x + y \in l_p$. This subcase is proved.

This completes the proof of the theorem. $\square$

Example 7.10 Let $\mathcal{X}$ be a real (complex) normed linear space. Let $Y := \{(\xi_k)_{k=1}^\infty \mid \xi_k \in \mathcal{X}, \forall k \in \mathbb{N}\}$. By Example 6.20, $(Y, \oplus, \otimes, \vartheta)$ is a vector space over $\mathbb{K}$, where $\oplus$, $\otimes$, and ϑ are defined as in the example. For $p \in [1, \infty) \subset \mathbb{R}$, let $M_p := \{(\xi_k)_{k=1}^\infty \in Y \mid \sum_{k=1}^\infty \|\xi_k\|_{\mathcal{X}}^p < +\infty\}$. Define the norm $\|x\|_p = (\sum_{k=1}^\infty \|\xi_k\|_{\mathcal{X}}^p)^{1/p}$, $\forall x = (\xi_k)_{k=1}^\infty \in M_p$. Let $M_\infty := \{(\xi_k)_{k=1}^\infty \in Y \mid \sup_{k \geq 1} \|\xi_k\|_{\mathcal{X}} < +\infty\}$. Define the norm $\|x\|_\infty = \sup_{k \geq 1} \|\xi_k\|_{\mathcal{X}}$, $\forall x = (\xi_k)_{k=1}^\infty \in M_\infty$.

$\forall p \in [1, \infty] \subset \mathbb{R}_e$, $\forall x = (\xi_k)_{k=1}^\infty, y = (\eta_k)_{k=1}^\infty \in M_p$.

$$\begin{aligned} & \|x + y\|_p \\ &= \begin{cases} \left(\sum_{k=1}^\infty \|\xi_k + \eta_k\|_{\mathcal{X}}^p \right)^{1/p} \leq \left(\sum_{k=1}^\infty (\|\xi_k\|_{\mathcal{X}} + \|\eta_k\|_{\mathcal{X}})^p \right)^{1/p} & p \in [1, \infty) \\ \sup_{k \geq 1} \|\xi_k + \eta_k\|_{\mathcal{X}} \leq \sup_{k \geq 1} (\|\xi_k\|_{\mathcal{X}} + \|\eta_k\|_{\mathcal{X}}) & p = \infty \end{cases} \\ &\leq \begin{cases} \left(\sum_{k=1}^\infty \|\xi_k\|_{\mathcal{X}}^p \right)^{1/p} + \left(\sum_{k=1}^\infty \|\eta_k\|_{\mathcal{X}}^p \right)^{1/p} & p \in [1, \infty) \\ \sup_{k \geq 1} \|\xi_k\|_{\mathcal{X}} + \sup_{k \geq 1} \|\eta_k\|_{\mathcal{X}} & p = \infty \end{cases} \\ &= \|x\|_p + \|y\|_p < +\infty \end{aligned}$$

where we have made use of the Minkowski's Inequality. In the preceding inequality, when $1 < p < \infty$, equality holds if, and only if, $\|\xi_k + \eta_k\|_{\mathcal{X}} = \|\xi_k\|_{\mathcal{X}} + \|\eta_k\|_{\mathcal{X}}$, $\forall k \in \mathbb{N}$, and $\exists \alpha, \beta \in [0, \infty) \subset \mathbb{R}$, which are not both zeros, such that $\alpha \|\xi_k\|_{\mathcal{X}} = \beta \|\eta_k\|_{\mathcal{X}}$, $\forall k \in \mathbb{N}$.

$\forall p \in [1, \infty] \subset \mathbb{R}_e$. Note that $\vartheta = (\vartheta_{\mathcal{X}}, \vartheta_{\mathcal{X}}, \dots) \in M_p \neq \emptyset$. $\forall x = (\xi_k)_{k=1}^\infty \in M_p$, $\forall \alpha \in \mathbb{K}$, we have

$$\| \alpha x \|_p = \begin{cases} \left(\sum_{k=1}^\infty \|\alpha \xi_k\|_{\mathcal{X}}^p \right)^{1/p} = \left(\sum_{k=1}^\infty |\alpha|^p \|\xi_k\|_{\mathcal{X}}^p \right)^{1/p} & p \in [1, \infty) \\ \sup_{k \geq 1} \|\alpha \xi_k\|_{\mathcal{X}} = \sup_{k \geq 1} |\alpha| \|\xi_k\|_{\mathcal{X}} & p = \infty \end{cases}$$

$$= \begin{cases} |\alpha| \left(\sum_{k=1}^{\infty} \|\xi_k\|_{\mathcal{X}}^p \right)^{1/p} & p \in [1, \infty) \\ |\alpha| \sup_{k \geq 1} \|\xi_k\|_{\mathcal{X}} & p = \infty \end{cases} = |\alpha| \|x\|_p < +\infty$$

where we have made use of Proposition 3.81. Then, $\forall x, y \in M_p, \forall \alpha, \beta \in \mathbb{K}$, we have $\|\alpha x + \beta y\|_p \leq \|\alpha x\|_p + \|\beta y\|_p = |\alpha| \|x\|_p + |\beta| \|y\|_p < +\infty$. Then, $\alpha x + \beta y \in M_p$. Hence, $\mathcal{M}_p := (M_p, \oplus, \otimes, \vartheta)$ is a subspace in $(Y, \oplus, \otimes, \vartheta)$. It is easy to check that, $\forall x \in \mathcal{M}_p, \|x\|_p \in [0, \infty) \subset \mathbb{R}$, and $\|x\|_p = 0 \Leftrightarrow x = (\vartheta_{\mathcal{X}}, \vartheta_{\mathcal{X}}, \dots) = \vartheta$. Therefore, $l_p(\mathcal{X}) := (\mathcal{M}_p, \mathbb{K}, \|\cdot\|_p)$ is a real (complex) normed linear space, $\forall p \in [1, \infty] \subset \mathbb{R}_e$. $\diamond$

Example 7.11 Let $X := \{f : [a, b] \rightarrow \mathbb{R}\}$, where $a, b \in \mathbb{R}$ with $a < b$, $\oplus$ and $\otimes$ be the usual vector addition and scalar multiplication, and $\vartheta : [a, b] \rightarrow \mathbb{R}$ be given by $\vartheta(t) = 0, \forall t \in [a, b]$. By Example 6.20, $\mathcal{X} := (X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathbb{R}$. Let $M := \{f \in \mathcal{X} \mid f \text{ is continuous}\}$ is a subspace by Proposition 6.25. Let $\mathcal{M} := (M, \oplus, \otimes, \vartheta)$. Then, $(\mathcal{M}, \mathbb{R})$ is a vector space. Introduce a norm on this space by $\|x\| = \int_a^b |x(t)| dt, \forall x \in \mathcal{M}$. Now, we verify the properties of the norm. $\forall x, y \in \mathcal{M}, \forall \alpha \in \mathbb{R}$.

(i) Since x is continuous, then, $|x(\cdot)|$ is continuous and therefore integrable on $[a, b]$. Hence, $0 \leq \|x\| < +\infty$. $x = \vartheta \Rightarrow \|x\| = 0$. On the other hand, $x \neq \vartheta \Rightarrow \exists t_1 \in [a, b]$ such that $|x(t_1)| > 0$. By continuity of $x, \exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $|x(t)| > |x(t_1)|/2, \forall t \in [t_1 - \delta, t_1 + \delta] \cap [a, b] \Rightarrow \|x\| > 0$. Hence, $\|x\| = 0 \Leftrightarrow x = \vartheta$.

(ii) $\|x + y\| = \int_a^b |x(t) + y(t)| dt \leq \int_a^b (|x(t)| + |y(t)|) dt = \|x\| + \|y\|$.

(iii) $\|\alpha x\| = \int_a^b |\alpha x(t)| dt = |\alpha| \|x\|$.

Hence, $(\mathcal{M}, \mathbb{R}, \|\cdot\|)$ is a normed linear space. $\diamond$

Definition 7.12 Let X be a set, $\mathcal{Y}$ be a normed linear space, $f : X \rightarrow \mathcal{Y}$. f is said to be bounded if $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|f(x)\| \leq M, \forall x \in X$. $S \subseteq \mathcal{Y}$ is said to be bounded if $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|s\| \leq M, \forall s \in S$.

Proposition 7.13 Let $\mathcal{X} := (\mathcal{X}, \mathbb{K}, \|\cdot\|)$ be a normed linear space and $M \subseteq \mathcal{X}$ be a subspace. Then, $(M, \mathbb{K}, \|\cdot\|)$ is also a normed linear space.

Proof This is straightforward and is therefore omitted. $\square$

7.2 The Natural Metric

A normed linear space is actually a metric space.

Proposition 7.14 *A normed linear space $\mathcal{X} := (\mathcal{X}, \mathbb{K}, \|\cdot\|)$ admits the natural metric $\rho : \mathcal{X} \times \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$ given by $\rho(x, y) = \|x - y\|$, $\forall x, y \in \mathcal{X}$.*

Proof $\forall x, y, z \in \mathcal{X}$, we have (i) $\rho(x, y) \in [0, \infty)$; (ii) $\rho(x, y) = \|x - y\| = 0 \Leftrightarrow x - y = \vartheta \Leftrightarrow x = y$; (iii) $\rho(x, y) = \|x - y\| = \|(-1)(x - y)\| = \|y - x\| = \rho(y, x)$; and (iv) Note that

$$\begin{aligned} \rho(x, z) &= \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| \\ &= \rho(x, y) + \rho(y, z) \end{aligned}$$

Hence, ρ is a metric. This completes the proof of the proposition. $\square$

Now, we can talk about topological properties and metric properties of a normed linear space. When we refer to these properties of a normed linear space, we are referring to the above metric specifically.

Proposition 7.15 *Let $\mathcal{X}$ be a normed linear space and $C \subseteq \mathcal{X}$ be convex. Then, $\overline{C}$ and C° are convex.*

Proof $\forall x_1, x_2 \in \overline{C}$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we need to show that $\alpha x_1 + (1 - \alpha)x_2 \in \overline{C}$. By Proposition 3.3, $\forall r \in (0, \infty) \subset \mathbb{R}$, $\mathcal{B}(x_i, r) \cap C \neq \emptyset$, $i = 1, 2$. Let $p_i \in \mathcal{B}(x_i, r) \cap C$, $i = 1, 2$. Then, by convexity of C , we have $\alpha p_1 + (1 - \alpha)p_2 \in C$. Note that

$$\begin{aligned} &\|(\alpha x_1 + (1 - \alpha)x_2) - (\alpha p_1 + (1 - \alpha)p_2)\| \\ &= \|\alpha(x_1 - p_1) + (1 - \alpha)(x_2 - p_2)\| \\ &\leq \alpha \|x_1 - p_1\| + (1 - \alpha) \|x_2 - p_2\| < r \end{aligned}$$

Then, $\alpha p_1 + (1 - \alpha)p_2 \in \mathcal{B}(\alpha x_1 + (1 - \alpha)x_2, r) \cap C \neq \emptyset$. Hence, $\alpha x_1 + (1 - \alpha)x_2 \in \overline{C}$ by the arbitrariness of r and Proposition 3.3. Then, $\overline{C}$ is convex.

$\forall x_1, x_2 \in C^\circ$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we need to show that $\alpha x_1 + (1 - \alpha)x_2 \in C^\circ$. $\exists r_i \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(x_i, r_i) \subseteq C$, $i = 1, 2$. Let $r := \min\{r_1, r_2\} > 0$. $\forall p \in \mathcal{B}(\alpha x_1 + (1 - \alpha)x_2, r)$, let $w := p - \alpha x_1 - (1 - \alpha)x_2$. Then, $\|w\| < r$ and $x_i + w \in \mathcal{B}(x_i, r_i) \subseteq C$, $i = 1, 2$. By the convexity of C , we have $C \ni \alpha(x_1 + w) + (1 - \alpha)(x_2 + w) = p$. Hence, we have $\mathcal{B}(\alpha x_1 + (1 - \alpha)x_2, r) \subseteq C$ and $\alpha x_1 + (1 - \alpha)x_2 \in C^\circ$. Hence, C° is convex.

This completes the proof of the proposition. $\square$

Proposition 7.16 *Let $\mathcal{X}$ be a normed linear space, $x_0 \in \mathcal{X}$, $S \subseteq \mathcal{X}$, and $P = x_0 + S$. Then, $\overline{P} = x_0 + \overline{S}$ and $P^\circ = x_0 + S^\circ$.*

Proof $\forall x \in \overline{P}$, by Proposition 3.3, $\forall r \in (0, \infty) \subset \mathbb{R}$, $\exists p_0 \in P \cap \mathcal{B}(x, r)$. Then, $p_0 - x_0 \in S \cap \mathcal{B}(x - x_0, r) \neq \emptyset$. Hence, by Proposition 3.3, $x - x_0 \in \overline{S}$ and $x \in x_0 + \overline{S}$. Hence, $\overline{P} \subseteq x_0 + \overline{S}$.

On the other hand, $\forall x \in x_0 + \overline{S}$, we have $x - x_0 \in \overline{S}$ and, by Proposition 3.3, $\forall r \in (0, \infty) \subset \mathbb{R}$, $\exists s_0 \in S \cap \mathcal{B}(x - x_0, r) \neq \emptyset$. Then, $x_0 + s_0 \in P \cap \mathcal{B}(x, r) \neq \emptyset$. By Proposition 3.3, $x \in \overline{P}$. Hence, $x_0 + \overline{S} \subseteq \overline{P}$.

Therefore, we have $\overline{P} = x_0 + \overline{S}$.

$\forall x \in P^\circ, \exists r \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(x, r) \subseteq P$. Then, $\mathcal{B}(x - x_0, r) = \mathcal{B}(x, r) - x_0 \subseteq P - x_0 = S$. Hence, $x - x_0 \in S^\circ, x \in x_0 + S^\circ$ and $P^\circ \subseteq x_0 + S^\circ$.

On the other hand, $\forall x \in x_0 + S^\circ$, we have $x - x_0 \in S^\circ$. Then, $\exists r \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(x - x_0, r) \subseteq S$. Hence, $\mathcal{B}(x, r) = \mathcal{B}(x - x_0, r) + x_0 \subseteq x_0 + S = P$. Therefore, $x \in P^\circ$ and $x_0 + S^\circ \subseteq P^\circ$.

Hence, we have $P^\circ = x_0 + S^\circ$. This completes the proof of the proposition. $\square$

Proposition 7.17 *Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$. Then, the following statements hold.*

- (i) *If $M \subseteq \mathcal{X}$ is a subspace, then $\overline{M}$ is a subspace.*
- (ii) *If $V \subseteq \mathcal{X}$ is a linear variety, then $\overline{V}$ is a linear variety.*
- (iii) *If $C \subseteq \mathcal{X}$ is a cone with vertex $p \in \mathcal{X}$, then $\overline{C}$ is a cone with vertex p .*
- (iv) *Let $x_0 \in \mathcal{X}$ and $r \in (0, \infty) \subset \mathbb{R}$. Then, $\overline{\mathcal{B}(x_0, r)} = \overline{\mathcal{B}}(x_0, r)$.*

Proof (i) Clearly, we have $\vartheta \in M \subseteq \overline{M} \neq \emptyset$. $\forall x, y \in \overline{M}, \forall \alpha \in \mathbb{K}$, we will show that $x + y, \alpha x \in \overline{M}$. Then, $\overline{M}$ is a subspace. $\forall r \in (0, \infty) \subset \mathbb{R}$, by Proposition 3.3, $\exists x_0 \in \mathcal{B}(x, r/2) \cap M \neq \emptyset$ and $\exists y_0 \in \mathcal{B}(y, r/2) \cap M \neq \emptyset$. Since M is a subspace, then $x_0 + y_0 \in M$ and $x_0 + y_0 \in \mathcal{B}(x + y, r)$. Hence, we have $M \cap \mathcal{B}(x + y, r) \neq \emptyset$, which implies that $x + y \in \overline{M}$. $\forall r \in (0, \infty) \subset \mathbb{R}$, by Proposition 3.3, $\exists x_0 \in \mathcal{B}(x, r/(1 + |\alpha|)) \cap M \neq \emptyset$. Since M is a subspace, then $\alpha x_0 \in M$ and $\alpha x_0 \in \mathcal{B}(\alpha x, r)$. Hence, we have $M \cap \mathcal{B}(\alpha x, r) \neq \emptyset$, which implies that $\alpha x \in \overline{M}$. Hence, $\overline{M}$ is a subspace.

(ii) Note that $V = x_0 + M$, where $x_0 \in \mathcal{X}$ and $M \subseteq \mathcal{X}$ is a subspace. By Proposition 7.16, $\overline{V} = x_0 + \overline{M}$. By (i), $\overline{M}$ is a subspace. Then, $\overline{V}$ is a linear variety.

(iii) First, we will prove the special case $p = \vartheta$. Clearly, $\vartheta \in C \subseteq \overline{C}$. $\forall x \in \overline{C}, \forall \alpha \in [0, \infty) \subset \mathbb{R}$, we will show that $\alpha x \in \overline{C}$. $\forall r \in (0, \infty) \subset \mathbb{R}$, by Proposition 3.3, $\exists x_0 \in C \cap \mathcal{B}(x, r/(1 + |\alpha|)) \neq \emptyset$. Then, we have $\alpha x_0 \in C \cap \mathcal{B}(\alpha x, r) \neq \emptyset$. By Proposition 3.3, $\alpha x \in \overline{C}$. Hence, $\overline{C}$ is a cone with vertex at the origin.

For general $p \in \mathcal{X}$, we have $C = p + C_0$, where C_0 is a cone with vertex at the origin. By Proposition 7.16, $\overline{C} = p + \overline{C_0}$. By the special case we have shown, $\overline{C_0}$ is a cone with vertex at the origin. Hence, $\overline{C}$ is a cone with vertex p .

(iv) First, we will prove the special case $x_0 = \vartheta$. By Proposition 4.3, $\overline{\mathcal{B}}(\vartheta, r)$ is a closed set and $\overline{\mathcal{B}}(\vartheta, r) \supseteq \mathcal{B}(\vartheta, r)$. Then, we have $\overline{\mathcal{B}}(\vartheta, r) \subseteq \overline{\mathcal{B}}(\vartheta, r)$. On the other hand, $\forall x \in \overline{\mathcal{B}}(\vartheta, r)$, define $(x_k)_{k=1}^\infty \subseteq \mathcal{X}$ by $x_k := \frac{k}{1+k}x, \forall k \in \mathbb{N}$. Then, $\|x_k\| = \frac{k}{1+k}\|x\| \leq \frac{k}{1+k}r < r, \forall k \in \mathbb{N}$. Therefore, $(x_k)_{k=1}^\infty \subseteq \mathcal{B}(\vartheta, r)$. It is obvious that $\lim_{k \in \mathbb{N}} x_k = x$, then, by

Proposition 4.13, $x \in \overline{\mathcal{B}(\vartheta, r)}$. Then, we have $\overline{\mathcal{B}(\vartheta, r)} \subseteq \overline{\mathcal{B}(\vartheta, r)}$. Hence, the result holds for $x_0 = \vartheta$.

For arbitrary $x_0 \in \mathcal{X}$, by Proposition 7.16, $\overline{\mathcal{B}(x_0, r)} = \overline{x_0 + \mathcal{B}(\vartheta, r)} = x_0 + \overline{\mathcal{B}(\vartheta, r)} = x_0 + \overline{\mathcal{B}(\vartheta, r)} = \overline{\mathcal{B}(x_0, r)}$. Hence, the result holds.

This completes the proof of the proposition. $\square$

Proposition 7.18 *Let $\mathcal{X}$ be a normed linear space, $P \subseteq \mathcal{X}$, and $P \neq \emptyset$. The closed linear variety generated by P , denoted by $V(P)$, is the intersection of all closed linear varieties containing P . Then, $V(P) = \overline{v(P)}$.*

Proof By Proposition 7.17, $\overline{v(P)}$ is a closed linear variety containing P . Then, we have $V(P) \subseteq \overline{v(P)}$. On the other hand, $V(P)$ is a closed linear variety containing P , and then $v(P) \subseteq V(P)$. Hence, $\overline{v(P)} \subseteq V(P)$. Therefore, the result holds. This completes the proof of the proposition. $\square$

The justification of the definition of $V(P)$ is that intersection of linear varieties is a linear variety when the intersection is nonempty.

Definition 7.19 *Let $\mathcal{X}$ be a normed linear space and $P \subseteq \mathcal{X}$ be nonempty. $x \in P$ is said to be a relative interior point of P if it is an interior point of P relative to the subset topology of $V(P)$. The set of all relative interior points of P is called the relative interior of P , and denoted by $^\circ P$. P is said to be relatively open if P is open in the subset topology of $V(P)$.*

Proposition 7.20 *Let $\mathcal{X}$ be a normed linear space and $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq \mathcal{X}$ be a net. Then, $\lim_{\alpha \in \mathcal{A}} x_\alpha = x \in \mathcal{X}$ if, and only if, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \alpha_0 \in \mathcal{A}$, $\forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$, we have $\|x - x_\alpha\| < \epsilon$.*

Proposition 7.21 *Let $\mathcal{X}$ be a normed linear space. Then, $\|\cdot\|$ is a uniformly continuous function on $\mathcal{X}$.*

Proof This follows directly from Propositions 4.30 and 7.14. $\square$

7.3 Product Spaces

Proposition 7.22 *Let $\mathcal{X} := (\mathcal{X}, \mathbb{K}, \|\cdot\|_{\mathcal{X}})$ and $\mathcal{Y} := (\mathcal{Y}, \mathbb{K}, \|\cdot\|_{\mathcal{Y}})$ be normed linear spaces. By Proposition 6.24, $\mathcal{X} \times \mathcal{Y}$ is a vector space over $\mathbb{K}$. Define a function $\|\cdot\| : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\|(x, y)\| := (\|x\|_{\mathcal{X}}^2 + \|y\|_{\mathcal{Y}}^2)^{1/2}$, $\forall (x, y) \in \mathcal{X} \times \mathcal{Y}$. Then, $(\mathcal{X} \times \mathcal{Y}, \mathbb{K}, \|\cdot\|)$ is a normed linear space. This normed linear space will be called the Cartesian product of $\mathcal{X}$ and $\mathcal{Y}$ and be denoted by $\mathcal{X} \times \mathcal{Y}$.*

Proof $\forall (x_1, y_1), (x_2, y_2) \in \mathcal{X} \times \mathcal{Y}$, $\forall \alpha \in \mathbb{K}$, we have $\|(x_1, y_1)\| = (\|x_1\|_{\mathcal{X}}^2 + \|y_1\|_{\mathcal{Y}}^2)^{1/2} \in [0, \infty) \subset \mathbb{R}$ and $\|(x_1, y_1)\| = 0 \Leftrightarrow \|x_1\|_{\mathcal{X}} = 0$ and $\|y_1\|_{\mathcal{Y}} = 0 \Leftrightarrow x_1 = \vartheta_{\mathcal{X}}$ and $y_1 = \vartheta_{\mathcal{Y}} \Leftrightarrow (x_1, y_1) = \vartheta$; $\|(x_1, y_1) + (x_2, y_2)\| = \|(x_1 + x_2, y_1 + y_2)\| = (\|x_1 + x_2\|_{\mathcal{X}}^2 + \|y_1 + y_2\|_{\mathcal{Y}}^2)^{1/2} \leq ((\|x_1\|_{\mathcal{X}} + \|x_2\|_{\mathcal{X}})^2 + (\|y_1\|_{\mathcal{Y}} + \|y_2\|_{\mathcal{Y}})^2)^{1/2} \leq (\|x_1\|_{\mathcal{X}} + \|x_2\|_{\mathcal{X}}) + (\|y_1\|_{\mathcal{Y}} + \|y_2\|_{\mathcal{Y}}) = \|x_1\|_{\mathcal{X}} + \|x_2\|_{\mathcal{X}} + \|y_1\|_{\mathcal{Y}} + \|y_2\|_{\mathcal{Y}} = \|(x_1, y_1)\| + \|(x_2, y_2)\|$

$(\|y_1\|_{\mathcal{Y}} + \|y_2\|_{\mathcal{Y}})^2)^{1/2} \leq (\|x_1\|_{\mathcal{X}}^2 + \|y_1\|_{\mathcal{Y}}^2)^{1/2} + (\|x_2\|_{\mathcal{X}}^2 + \|y_2\|_{\mathcal{Y}}^2)^{1/2} = \| (x_1, y_1) \| + \| (x_2, y_2) \|$, where the first inequality follows from the fact that $\mathcal{X}$ and $\mathcal{Y}$ are normed linear spaces and the second inequality follows from the Minkowski's Inequality; $\| \alpha (x_1, y_1) \| = \| (\alpha x_1, \alpha y_1) \| = (\| \alpha x_1 \|_{\mathcal{X}}^2 + \| \alpha y_1 \|_{\mathcal{Y}}^2)^{1/2} = (|\alpha|^2 \|x_1\|_{\mathcal{X}}^2 + |\alpha|^2 \|y_1\|_{\mathcal{Y}}^2)^{1/2} = |\alpha| (\|x_1\|_{\mathcal{X}}^2 + \|y_1\|_{\mathcal{Y}}^2)^{1/2} = |\alpha| \| (x_1, y_1) \|$. Hence, $\| \cdot \|$ is a norm on $\mathcal{X} \times \mathcal{Y}$. This completes the proof of the proposition. $\square$

Clearly, the natural metric for the Cartesian product $\mathcal{X} \times \mathcal{Y}$ is the Cartesian metric defined in Definition 4.28.

The above proposition may also be generalized to the case of $\mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_n$, where $n \in \mathbb{N}$ and $\mathcal{X}_i$'s are normed linear spaces over the same field $\mathbb{K}$. When $n = 0$, it should be noted that $\prod_{\alpha \in \emptyset} \mathcal{X}_{\alpha} = (\{\emptyset =: \vartheta\}, \mathbb{K}, \| \cdot \|)$, where $\| \vartheta \| = 0$.

Proposition 7.23 *Let $\mathcal{X} := (\mathcal{X}, \mathbb{K}, \| \cdot \|)$ be a normed linear space. Then, the vector addition $\oplus_{\mathcal{X}} : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ is uniformly continuous; and the scalar multiplication $\otimes_{\mathcal{X}} : \mathbb{K} \times \mathcal{X} \rightarrow \mathcal{X}$ is continuous.*

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall (x_1, x_2), (y_1, y_2) \in \mathcal{X} \times \mathcal{X}$ with

$$\| (x_1, x_2) - (y_1, y_2) \|_{\mathcal{X} \times \mathcal{X}} < \epsilon / \sqrt{2}$$

we have

$$\begin{aligned} \| (x_1 + x_2) - (y_1 + y_2) \| &\leq \| x_1 - y_1 \| + \| x_2 - y_2 \| \\ &\leq \sqrt{2} (\| x_1 - y_1 \|^2 + \| x_2 - y_2 \|^2)^{1/2} = \sqrt{2} \| (x_1, x_2) - (y_1, y_2) \|_{\mathcal{X} \times \mathcal{X}} < \epsilon \end{aligned}$$

Hence, the vector addition $\oplus_{\mathcal{X}}$ is uniformly continuous.

$\forall \alpha_0 \in \mathbb{K}, \forall x_0 \in \mathcal{X}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta = \epsilon / (1 + \epsilon + |\alpha_0| + \|x_0\|) \in (0, 1) \subset \mathbb{R}$. $\forall (\alpha, x) \in \mathcal{B}_{\mathbb{K} \times \mathcal{X}}((\alpha_0, x_0), \delta)$, we have

$$\begin{aligned} \| \alpha x - \alpha_0 x_0 \| &\leq \| \alpha x - \alpha x_0 \| + \| \alpha x_0 - \alpha_0 x_0 \| \\ &= |\alpha| \| x - x_0 \| + \| x_0 \| |\alpha - \alpha_0| \\ &\leq \left(|\alpha|^2 + \|x_0\|^2 \right)^{1/2} \left(\|x - x_0\|^2 + |\alpha - \alpha_0|^2 \right)^{1/2} \\ &\leq (|\alpha| + \|x_0\|) \delta \leq (1 + |\alpha_0| + \|x_0\|) \delta < \epsilon \end{aligned}$$

where we have made use of Cauchy-Schwarz Inequality in the second inequality. Hence, $\otimes_{\mathcal{X}}$ is continuous at (α_0, x_0) . By the arbitrariness of (α_0, x_0) , we have that $\otimes_{\mathcal{X}}$ is continuous.

This completes the proof of the proposition. $\square$

Definition 7.24 *Let $\mathcal{X} := (\mathcal{X}, \mathbb{K}, \| \cdot \|_{\mathcal{X}})$ and $\mathcal{Y} := (\mathcal{Y}, \mathbb{K}, \| \cdot \|_{\mathcal{Y}})$ be two normed linear spaces over the same field $\mathbb{K}$ and $A : \mathcal{X} \rightarrow \mathcal{Y}$ be a vector space isomorphism. A is said to be an isometrical isomorphism if $\| Ax \|_{\mathcal{Y}} = \| x \|_{\mathcal{X}}, \forall x \in \mathcal{X}$. The $\mathcal{X}$ and $\mathcal{Y}$ are said to be isometrically isomorphic.*

Let $A : \mathcal{X} \rightarrow \mathcal{Y}$ be an isometrical isomorphism between $\mathcal{X}$ and $\mathcal{Y}$. Then, A is an isometry between $\mathcal{X}$ and $\mathcal{Y}$. Both A and A_{inv} are uniformly continuous. $\mathcal{X}$ and $\mathcal{Y}$ are equal to each other up to a relabeling of their vectors.

Proposition 7.25 *Let $\mathcal{X}_\alpha$, $\alpha \in \Lambda$, be normed linear spaces over $\mathbb{K}$, where Λ is a finite index set. Let $\Lambda = \bigcup_{\beta \in \Gamma} \Lambda_\beta$, where Λ_β 's are pairwise disjoint and finite and Γ is also finite. $\forall \beta \in \Gamma$, let $\mathcal{X}^{(\beta)} := \prod_{\alpha \in \Lambda_\beta} \mathcal{X}_\alpha$ be the Cartesian product space. Let $\mathcal{X}^{(\Gamma)} := \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} \mathcal{X}_\alpha$ be the Cartesian product space of product spaces, and $\mathcal{X} := \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha$ be the Cartesian product space. Then, $\mathcal{X}$ and $\mathcal{X}^{(\Gamma)}$ are isometrically isomorphic.*

Proof Define $E : \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} \mathcal{X}_\alpha \rightarrow \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha$ by, $\forall x \in \mathcal{X}^{(\Gamma)}$, $\forall \alpha \in \Lambda$, $\exists! \beta_\alpha \in \Gamma \ni \alpha \in \Lambda_{\beta_\alpha}$, $\pi_\alpha(E(x)) = \pi_\alpha^{(\beta_\alpha)}(\pi_{\beta_\alpha}^{(\Gamma)}(x))$. By Proposition 4.32, E is a isometry. It is clear that E is a linear operator since the projection functions are linear for vector spaces. Hence, E is a vector space isomorphism. Then, E is an isometrical isomorphism. This completes the proof of the proposition. $\square$

7.4 Banach Spaces

Definition 7.26 *If a normed linear space is complete with respect to the natural metric, then it is called a Banach space.*

Clearly, a Cauchy sequence in a normed linear space is bounded.

Proposition 7.27 *A normed linear space $\mathcal{X}$ is complete if, and only if, every absolutely summable series is summable, that is $\forall (x_n)_{n=1}^\infty \subseteq \mathcal{X}$, $\sum_{n=1}^\infty \|x_n\| \in \mathbb{R} \Rightarrow \sum_{n=1}^\infty x_n \in \mathcal{X}$.*

Proof “Only if” Let $\mathcal{X}$ be complete and $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$ be absolutely summable. Then, $\sum_{n=1}^\infty \|x_n\| \in \mathbb{R}$ and $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that, $\forall n, m \in \mathbb{N}$ with $n \geq m$, we have $\sum_{i=m}^n \|x_i\| < \epsilon$. Then, $\|\sum_{i=m}^n x_i\| < \epsilon$. Let $s_n := \sum_{i=1}^n x_i$, $\forall n \in \mathbb{N}$, be the partial sum. Then, $(s_n)_{n=1}^\infty \subseteq \mathcal{X}$ is a Cauchy sequence. By the completeness of $\mathcal{X}$, $\sum_{n=1}^\infty x_n = \lim_{n \in \mathbb{N}} s_n \in \mathcal{X}$.

“If” Let $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$ be a Cauchy sequence. Let $n_0 = 0$. $\forall k \in \mathbb{N}$, $\exists n_k \in \mathbb{N}$ with $n_k > n_{k-1}$ such that $\|x_n - x_m\| < 2^{-k}$, $\forall n, m \geq n_k$. Then, $(x_{n_k})_{k=1}^\infty$ is a subsequence of $(x_n)_{n=1}^\infty$. Let $y_1 := x_{n_1}$ and $y_k := x_{n_k} - x_{n_{k-1}}$, $\forall k \geq 2$. Then, the sequence $(y_k)_{k=1}^\infty \subseteq \mathcal{X}$ and its k th partial sum is x_{n_k} . Note that $\|y_k\| < 2^{-k+1}$, $\forall k \geq 2$. Then, $\sum_{k=1}^\infty \|y_k\| < \|y_1\| + 1$, which implies that $\sum_{k=1}^\infty y_k = x_0 \in \mathcal{X}$. Hence, we have $\lim_{k \in \mathbb{N}} x_{n_k} = x_0 \in \mathcal{X}$.

We will now show that $\lim_{n \in \mathbb{N}} x_n = x_0$. $\forall k \in \mathbb{N}$, $\exists N \in \mathbb{N}$ with $N \geq k$ such that $\|x_{n_i} - x_0\| < 2^{-k}$, $\forall i \geq N$. $\forall n \geq n_N$, we have $\|x_n - x_0\| \leq \|x_n - x_{n_N}\| + \|x_{n_N} - x_0\| < 2^{-N} + 2^{-k} \leq 2^{-k+1}$. Hence, $\lim_{n \in \mathbb{N}} x_n = x_0$. Therefore, $\mathcal{X}$ is complete.

This completes the proof of the proposition. $\square$

We frequently take great care to formulate problems arising in applications as equivalent problems in Banach spaces rather than other incomplete spaces. The principal advantage of Banach spaces in optimization problems is that when seeking an optimal vector, we often construct a sequence (net) of vectors, each member of which is superior to preceding ones. The desired optimal vector is then the limit of the sequence (net). In order for this scheme to be effective, there must be available a test for convergence which can be applied when the limit is unknown. The Cauchy criterion for convergence meets this requirement provided the space is complete.

Example 7.28 Consider Example 7.11 with $a = 0$ and $b = 1$. We will show that the space $(\mathcal{M}, \mathbb{R}, \|\cdot\|)$ is incomplete. Take $(x_n)_{n=1}^\infty \subseteq \mathcal{M}$ to be, $\forall n \in \mathbb{N}$,

$$x_n(t) = \begin{cases} 0 & 0 \leq t \leq \frac{1}{2} - \frac{1}{n+1} \\ (n+1)t - \frac{n+1}{2} + 1 & \frac{1}{2} - \frac{1}{n+1} < t \leq \frac{1}{2} \\ 1 & t > \frac{1}{2} \end{cases}$$

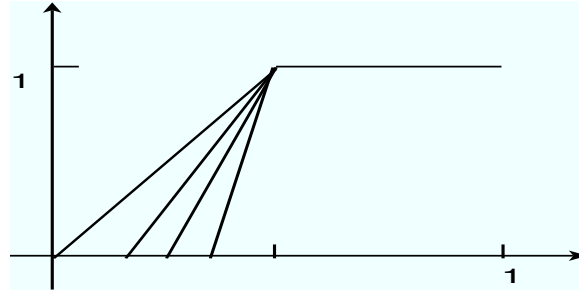


Figure 7.1: Sequence for Example 7.28.

Clearly, x_n is continuous and $\|x_n\| < 1$, $\forall n \in \mathbb{N}$. $\forall n, m \in \mathbb{N}$, we have

$$\begin{aligned} \|x_n - x_m\| &= \int_0^1 |x_n(t) - x_m(t)| dt = \left| \int_0^1 (x_n(t) - x_m(t)) dt \right| \\ &= \left| \frac{1}{2(n+1)} - \frac{1}{2(m+1)} \right| \end{aligned}$$

Then, $(x_n)_{n=1}^\infty$ is Cauchy. Yet, it is obvious that there is no continuous function $x \in \mathcal{M}$ such that $\lim_{n \in \mathbb{N}} x_n = x$, i. e., $\lim_{n \in \mathbb{N}} \int_0^1 |x_n(t) - x(t)| dt = 0$. Hence, the space is incomplete. $\diamond$

Example 7.29 $\mathbb{K}^n$ with norm defined as in Example 7.2 or Example 7.3 is a Banach space. $\diamond$

Example 7.30 Consider the real normed linear space $\mathcal{C}([a, b])$ defined in Example 7.4, where $a, b \in \mathbb{R}$ and $a \leq b$. We will show that it is complete.

Take a Cauchy sequence $(x_n)_{n=1}^\infty \subseteq \mathcal{C}([a, b])$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $\forall n, m \geq N$, $0 \leq |x_n(t) - x_m(t)| \leq \|x_n - x_m\| < \epsilon$, $\forall t \in [a, b]$. This shows that, $\forall t \in [a, b]$, $(x_n(t))_{n=1}^\infty \subseteq \mathbb{R}$ is a Cauchy sequence, which converges to $x(t) \in \mathbb{R}$ since $\mathbb{R}$ is complete. This defines a function $x : [a, b] \rightarrow \mathbb{R}$. It is easy to show that $(x_n)_{n=1}^\infty$, viewed as a sequence of functions of $[a, b]$ to $\mathbb{R}$, converges uniformly to x . By Proposition 4.26, x is continuous. Hence, $x \in \mathcal{C}([a, b])$. It is easy to see that $\lim_{n \in \mathbb{N}} \|x_n - x\| = 0$. Hence, $\lim_{n \in \mathbb{N}} x_n = x$. Hence, $\mathcal{C}([a, b])$ is a Banach space. $\diamond$

Example 7.31 Let $\mathcal{K}$ be a countably compact topological space, $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $Y := \{f : \mathcal{K} \rightarrow \mathcal{X}\}$. Define the usual vector addition $\oplus$ and scalar multiplication $\otimes$ and null vector ϑ on Y as in Example 6.20. Then, $\mathcal{Y} := (Y, \oplus, \otimes, \vartheta)$ is a vector space over $\mathbb{K}$. Let $M := \{f \in Y \mid f \text{ is continuous}\}$. Then, by Propositions 3.32, 6.25, and 7.23, $\mathcal{M} := (M, \oplus, \otimes, \vartheta)$ is a subspace of $(\mathcal{Y}, \mathbb{K})$. Define a function $\|\cdot\| : \mathcal{M} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\|f\| = \max\{\sup_{k \in \mathcal{K}} \|f(k)\|_{\mathcal{X}}, 0\}$, $\forall f \in \mathcal{M}$. This function is well defined by Propositions 7.21, 3.12, and 5.29. We will show that $\|\cdot\|$ defines a norm on $\mathcal{M}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{K} = \emptyset$; Case 2: $\mathcal{K} \neq \emptyset$. Case 1: $\mathcal{K} = \emptyset$. Then, $\mathcal{M}$ is a singleton set $\{\emptyset\}$ and $\|\emptyset\| = 0$. Clearly, $(\mathcal{M}, \mathbb{K}, \|\cdot\|)$ is a normed linear space. Case 2: $\mathcal{K} \neq \emptyset$. $\forall f, g \in \mathcal{M}$, $\forall \alpha \in \mathbb{K}$, $\|f\| = \max_{k \in \mathcal{K}} \|f(k)\|_{\mathcal{X}}$ by Propositions 7.21, 3.12, and 5.29. $\|f\| = 0 \Leftrightarrow \|f(k)\|_{\mathcal{X}} = 0$, $\forall k \in \mathcal{K} \Leftrightarrow f(k) = \vartheta_{\mathcal{X}}$, $\forall k \in \mathcal{K} \Leftrightarrow f = \vartheta$. $\|f + g\| = \max_{k \in \mathcal{K}} \|f(k) + g(k)\|_{\mathcal{X}} \leq \max_{k \in \mathcal{K}} (\|f(k)\|_{\mathcal{X}} + \|g(k)\|_{\mathcal{X}}) \leq \max_{k \in \mathcal{K}} \|f(k)\|_{\mathcal{X}} + \max_{k \in \mathcal{K}} \|g(k)\|_{\mathcal{X}} = \|f\| + \|g\|$. $\|\alpha f\| = \max_{k \in \mathcal{K}} \|\alpha f(k)\|_{\mathcal{X}} = \max_{k \in \mathcal{K}} |\alpha| \|f(k)\|_{\mathcal{X}} = |\alpha| \|f\|$. Hence, $(\mathcal{M}, \mathbb{K}, \|\cdot\|)$ is a normed linear space. In both case, we have shown that $\mathcal{C}(\mathcal{K}, \mathcal{X}) := (\mathcal{M}, \mathbb{K}, \|\cdot\|)$ is a normed linear space. $\diamond$

Example 7.32 Let $\mathcal{K}$ be a countably compact topological space, $\mathcal{X}$ be a Banach space over the field $\mathbb{K}$ (with norm $\|\cdot\|_{\mathcal{X}}$). Consider the normed linear space $\mathcal{C}(\mathcal{K}, \mathcal{X})$ (with norm $\|\cdot\|$) defined in Example 7.31. We will show that this space is a Banach space. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{K} = \emptyset$; Case 2: $\mathcal{K} \neq \emptyset$. Case 1: $\mathcal{K} = \emptyset$. Then, $\mathcal{C}(\mathcal{K}, \mathcal{X})$ is a singleton set. Hence, any Cauchy sequence must converge. Thus, $\mathcal{C}(\mathcal{K}, \mathcal{X})$ is a Banach space. Case 2: $\mathcal{K} \neq \emptyset$. Take a Cauchy sequence $(x_n)_{n=1}^\infty \subseteq \mathcal{C}(\mathcal{K}, \mathcal{X})$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $\forall n, m \geq N$, $0 \leq \|x_n(k) - x_m(k)\|_{\mathcal{X}} \leq \|x_n - x_m\| < \epsilon$, $\forall k \in \mathcal{K}$. This shows that, $\forall k \in \mathcal{K}$, $(x_n(k))_{n=1}^\infty \subseteq \mathcal{X}$ is a Cauchy sequence, which converges to $x(k) \in \mathcal{X}$ since $\mathcal{X}$ is complete. This defines a function $x : \mathcal{K} \rightarrow \mathcal{X}$. It is easy to show that $(x_n)_{n=1}^\infty$, viewed as a sequence of functions of $\mathcal{K}$ to $\mathcal{X}$, converges uniformly to x . By Proposition 4.26, x is continuous. Hence, $x \in \mathcal{C}(\mathcal{K}, \mathcal{X})$. It is easy to see that $\lim_{n \in \mathbb{N}} \|x_n - x\| = 0$. Hence, $\lim_{n \in \mathbb{N}} x_n = x$. Hence, $\mathcal{C}(\mathcal{K}, \mathcal{X})$ is a Banach space. In both cases, we have shown that $\mathcal{C}(\mathcal{K}, \mathcal{X})$ is a Banach space when $\mathcal{X}$ is a Banach space. $\diamond$

Example 7.33 Let $\mathcal{X}$ be a Banach space over the field $\mathbb{K}$. Consider the normed linear space $l_p(\mathcal{X})$ (with norm $\|\cdot\|_p$) defined in Example 7.10,

where $p \in [1, \infty] \subset \mathbb{R}_e$. We will show that this space is a Banach space. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $p \in [1, \infty)$; Case 2: $p = \infty$.

Case 1: $p \in [1, \infty)$. Take a Cauchy sequence $(x_n)_{n=1}^\infty \subseteq l_p(\mathcal{X})$, where $x_n = (\xi_{n,k})_{k=1}^\infty \subseteq \mathcal{X}$, $\forall n \in \mathbb{N}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N_\epsilon \in \mathbb{N}$ such that $\forall n, m \geq N_\epsilon$, we have $\|x_n - x_m\|_p < \epsilon$. $\forall k \in \mathbb{N}$, $\|\xi_{n,k} - \xi_{m,k}\|_{\mathcal{X}} \leq (\sum_{i=1}^\infty \|\xi_{n,i} - \xi_{m,i}\|_{\mathcal{X}}^p)^{1/p} = \|x_n - x_m\|_p < \epsilon$. Hence, $\forall k \in \mathbb{N}$, $(\xi_{n,k})_{n=1}^\infty \subseteq \mathcal{X}$ is a Cauchy sequence, which converges to some $\xi_k \in \mathcal{X}$ since $\mathcal{X}$ is a Banach space. Let $x := (\xi_k)_{k=1}^\infty \subseteq \mathcal{X}$. We will show that $x \in l_p(\mathcal{X})$ and $\lim_{n \in \mathbb{N}} x_n = x$. Since $(x_n)_{n=1}^\infty$ is a Cauchy sequence, then it is bounded, that is $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|x_n\|_p \leq M$, $\forall n \in \mathbb{N}$. Then, we have $\|x_n\|_p^p = \sum_{i=1}^\infty \|\xi_{n,i}\|_{\mathcal{X}}^p \leq M^p$, $\forall n \in \mathbb{N}$. Then, $\forall n, k \in \mathbb{N}$, we have $\sum_{i=1}^k \|\xi_{n,i}\|_{\mathcal{X}}^p \leq M^p$. By Propositions 7.21 and 3.66, we have $\sum_{i=1}^k \|\xi_i\|_{\mathcal{X}}^p = \lim_{n \in \mathbb{N}} \sum_{i=1}^k \|\xi_{n,i}\|_{\mathcal{X}}^p \leq M^p$. Hence, we have $\sum_{i=1}^\infty \|\xi_i\|_{\mathcal{X}}^p \leq M^p$ and $\|x\|_p = (\sum_{i=1}^\infty \|\xi_i\|_{\mathcal{X}}^p)^{1/p} \leq M$. Hence, we have $x \in l_p(\mathcal{X})$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall n, m \in \mathbb{N}$ with $n, m \geq N_\epsilon$, we have $\|x_n - x_m\|_p < \epsilon$. Then, $\forall k \in \mathbb{N}$, $\sum_{i=1}^k \|\xi_{n,i} - \xi_{m,i}\|_{\mathcal{X}}^p < \epsilon^p$. Taking limit as $m \rightarrow \infty$, by Propositions 7.21, 3.66, and 7.23, we have $\sum_{i=1}^k \|\xi_{n,i} - \xi_i\|_{\mathcal{X}}^p = \lim_{m \in \mathbb{N}} \sum_{i=1}^k \|\xi_{n,i} - \xi_{m,i}\|_{\mathcal{X}}^p \leq \epsilon^p$, $\forall k \in \mathbb{N}$. Hence, we have $\|x_n - x\|_p = (\sum_{i=1}^\infty \|\xi_{n,i} - \xi_i\|_{\mathcal{X}}^p)^{1/p} \leq \epsilon$. This shows that $\lim_{n \in \mathbb{N}} x_n = x$. Hence, $l_p(\mathcal{X})$ is complete and therefore a Banach space.

Case 2: $p = \infty$. Take a Cauchy sequence $(x_n)_{n=1}^\infty \subseteq l_\infty(\mathcal{X})$, where $x_n = (\xi_{n,k})_{k=1}^\infty \subseteq \mathcal{X}$, $\forall n \in \mathbb{N}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N_\epsilon \in \mathbb{N}$ such that $\forall n, m \geq N_\epsilon$, we have $\|x_n - x_m\|_\infty < \epsilon$. $\forall k \in \mathbb{N}$, $\|\xi_{n,k} - \xi_{m,k}\|_{\mathcal{X}} \leq \sup_{i \in \mathbb{N}} \|\xi_{n,i} - \xi_{m,i}\|_{\mathcal{X}} = \|x_n - x_m\|_\infty < \epsilon$. Hence, $\forall k \in \mathbb{N}$, $(\xi_{n,k})_{n=1}^\infty \subseteq \mathcal{X}$ is a Cauchy sequence, which converges to some $\xi_k \in \mathcal{X}$ since $\mathcal{X}$ is a Banach space. Let $x := (\xi_k)_{k=1}^\infty \subseteq \mathcal{X}$. We will show that $x \in l_\infty(\mathcal{X})$ and $\lim_{n \in \mathbb{N}} x_n = x$. Since $(x_n)_{n=1}^\infty$ is a Cauchy sequence, then it is bounded, that is $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|x_n\|_\infty \leq M$, $\forall n \in \mathbb{N}$. Then, we have $\|\xi_{n,k}\|_{\mathcal{X}} \leq \sup_{i \in \mathbb{N}} \|\xi_{n,i}\|_{\mathcal{X}} = \|x_n\|_\infty \leq M$, $\forall n, k \in \mathbb{N}$. By Propositions 7.21 and 3.66, we have $\|\xi_k\|_{\mathcal{X}} = \lim_{n \in \mathbb{N}} \|\xi_{n,k}\|_{\mathcal{X}} \leq M$. Hence, $\|x\|_\infty = \sup_{i \in \mathbb{N}} \|\xi_i\|_{\mathcal{X}} \leq M$. Therefore, $x \in l_\infty(\mathcal{X})$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall n, m \in \mathbb{N}$ with $n, m \geq N_\epsilon$, we have $\|x_n - x_m\|_\infty < \epsilon$. Then, $\forall k \in \mathbb{N}$, $\|\xi_{n,k} - \xi_{m,k}\|_{\mathcal{X}} < \epsilon$. Taking limit as $m \rightarrow \infty$, by Propositions 7.21, 3.66, and 7.23, we have $\|\xi_{n,k} - \xi_k\|_{\mathcal{X}} = \lim_{m \in \mathbb{N}} \|\xi_{n,k} - \xi_{m,k}\|_{\mathcal{X}} \leq \epsilon$, $\forall k \in \mathbb{N}$. Hence, we have $\|x_n - x\|_\infty = \sup_{i \in \mathbb{N}} \|\xi_{n,i} - \xi_i\|_{\mathcal{X}} \leq \epsilon$. This shows that $\lim_{n \in \mathbb{N}} x_n = x$. Hence, $l_\infty(\mathcal{X})$ is complete and therefore a Banach space.

In summary, we have shown that $l_p(\mathcal{X})$ is a Banach space, $\forall p \in [1, \infty] \subset \mathbb{R}_e$, when $\mathcal{X}$ is a Banach space. $\diamond$

Definition 7.34 Let $\mathcal{X}$ be a normed linear space and $S \subseteq \mathcal{X}$. S is said to be complete if S with the natural metric forms a complete metric space.

By Proposition 4.39, a subset in a Banach space is complete if, and only if, it is closed.

Proposition 7.35 *Let $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $S_1, S_2 \subseteq \mathcal{Y}$ be separable subsets, and $\alpha \in \mathbb{K}$. Then, $\text{span}(S_1)$, $\overline{\text{span}(S_1)}$, $S_1 + S_2$, αS_1 , $S_1 \cap S_2$, and $S_1 \cup S_2$ are separable subsets of $\mathcal{Y}$.*

Proof Let $\mathbb{K}_{\mathbb{Q}} := \mathbb{Q}$ if $\mathbb{K} = \mathbb{R}$; and $\mathbb{K}_{\mathbb{Q}} := \{a + ib \in \mathbb{C} \mid a, b \in \mathbb{Q}\}$, if $\mathbb{K} = \mathbb{C}$. Clearly, $\mathbb{K}_{\mathbb{Q}}$ is a countable dense set in $\mathbb{K}$. Let $D \subseteq S_1$ be a countable dense subset. Let $\hat{D} := \{\sum_{i=1}^n \alpha_i y_i \in \mathcal{Y} \mid n \in \mathbb{Z}_+, \alpha_i \in \mathbb{K}_{\mathbb{Q}}, y_i \in D, i = 1, \dots, n\}$, which is a countable set. Clearly, $\hat{D} \subseteq \text{span}(S_1)$ is a dense subset. Hence, $\text{span}(S_1) \subseteq \mathcal{Y}$ is separable. By Proposition 4.38, $\overline{\text{span}(S_1)}$ is separable. It is straightforward to show that $S_1 + S_2$, αS_1 , $S_1 \cap S_2$, and $S_1 \cup S_2$ are separable. $\square$

Theorem 7.36 *In a normed linear space $\mathcal{X}$, any finite-dimensional subspace $M \subseteq \mathcal{X}$ is complete.*

Proof Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$. Let $\bar{n} \in \mathbb{Z}_+$ be the dimension of M , which is well defined by Theorem 6.51. We will prove the theorem by mathematical induction on $\bar{n}$.

1° $\bar{n} = 0$. Then, $M = \{\vartheta\}$. Clearly, any Cauchy sequence in M must converge to $\vartheta \in M$. Hence, M is complete. $\bar{n} = 1$. Let $\{e_1\} \subseteq M$ be a basis for M . Clearly, $e_1 \neq \vartheta$ and $\|e_1\| > 0$. Fix any Cauchy sequence $(x_n)_{n=1}^{\infty} \subseteq M$. Then, $x_n = \alpha_n e_1$ for some $\alpha_n \in \mathbb{K}$, $\forall n \in \mathbb{N}$. Since $(x_n)_{n=1}^{\infty}$ is Cauchy, then $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$, $\forall n, m \geq N$, we have $\|x_n - x_m\| < \|e_1\| \epsilon$. Then, we have $|\alpha_n - \alpha_m| = \|x_n - x_m\| / \|e_1\| < \epsilon$. Hence, $(\alpha_n)_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{K}$. Then, $\lim_{n \in \mathbb{N}} \alpha_n = \alpha \in \mathbb{K}$ since $\mathbb{K}$ is complete. It is easy to show that $\lim_{n \in \mathbb{N}} x_n = \alpha e_1 \in M$. Hence, M is complete.

2° Assume M is complete when $\bar{n} = k - 1 \in \mathbb{N}$.

3° Consider the case $\bar{n} = k \in \{2, 3, \dots\} \subset \mathbb{N}$. Let $\{e_1, \dots, e_k\} \subseteq M$ be a basis for M . Define $M_i := \text{span}(\{e_1, \dots, e_k\} \setminus \{e_i\})$ and $\delta_i := \text{dist}(e_i, M_i)$, $i = 1, \dots, k$. $\forall i = 1, \dots, k$, we have $\delta_i \in [0, \infty) \subset \mathbb{R}$. M_i , which is a $k - 1$ -dimensional subspace of $\mathcal{X}$. By inductive assumption M_i is complete. By Proposition 4.39, M_i is closed. Clearly, $e_i \notin M_i$. By Proposition 4.10, $\delta_i > 0$.

Let $\delta := \min\{\delta_1, \dots, \delta_k\} > 0$. Fix any Cauchy sequence $(x_n)_{n=1}^{\infty} \subseteq M$. $\forall n \in \mathbb{N}$, x_n admits a unique representation $x_n = \sum_{i=1}^k \lambda_{n,i} e_i$ where $\lambda_{n,i} \in \mathbb{K}$, $i = 1, \dots, k$, by Definition 6.50 and Corollary 6.47. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $\forall n, m \geq N$, we have $\|x_n - x_m\| < \delta \epsilon$. $\forall i \in \{1, \dots, k\}$, we have $\epsilon > \|x_n - x_m\| / \delta = \left\| \sum_{j=1}^k (\lambda_{n,j} - \lambda_{m,j}) e_j \right\| / \delta \geq |\lambda_{n,i} - \lambda_{m,i}| \delta_i / \delta \geq |\lambda_{n,i} - \lambda_{m,i}|$. Hence, $(\lambda_{n,i})_{n=1}^{\infty} \subseteq \mathbb{K}$ is a Cauchy sequence, $\forall i \in \{1, \dots, k\}$. Then, $\lim_{n \in \mathbb{N}} \lambda_{n,i} = \lambda_i \in \mathbb{K}$. Let $x := \sum_{i=1}^k \lambda_i e_i \in M$.

Now, it is straightforward to show that $\lim_{n \in \mathbb{N}} \|x_n - x\| = 0$. Then, $\lim_{n \in \mathbb{N}} x_n = x \in M$. Hence, M is complete.

This completes the induction process. This completes the proof of the theorem. $\square$

Definition 7.37 Let $\mathcal{X}$ be a vector space over the field $\mathbb{K}$ and $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms defined on $\mathcal{X}$. These two norms are said to be equivalent if $\exists K \in (0, \infty) \subset \mathbb{R}$ such that $\|x\|_1/K \leq \|x\|_2 \leq K\|x\|_1$, $\forall x \in \mathcal{X}$.

Clearly, two norms are equivalent implies that the natural metrics are uniformly equivalent.

Theorem 7.38 Let $\mathcal{X}$ be a finite-dimensional vector space over the field $\mathbb{K}$. Any two norms on $\mathcal{X}$ are equivalent.

Proof Let $n \in \mathbb{Z}_+$ be the dimension of $\mathcal{X}$, which is well defined by Theorem 6.51. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms defined on $\mathcal{X}$. We will distinguish three exhaustive and mutually exclusive cases: Case 1: $n = 0$; Case 2: $n = 1$; Case 3: $n \geq 2$. Case 1: $n = 0$. Then, $\mathcal{X} = \{\vartheta\}$ and $\|\vartheta\|_1 = 0 = \|\vartheta\|_2$. Hence, the two norms are equivalent. Case 2: $n = 1$. Let $\{e_1\}$ be a basis for $\mathcal{X}$. Clearly, $e_1 \neq \vartheta$. Let $\delta := \|e_1\|_1 / \|e_1\|_2 \in (0, \infty) \subset \mathbb{R}$ and $K = \max\{\delta, \delta^{-1}\} \in (0, \infty) \subset \mathbb{R}$. $\forall x \in \mathcal{X}$, $\exists! \alpha \in \mathbb{K}$ such that $x = \alpha e_1$. Then, $\|x\|_1 = |\alpha| \|e_1\|_1 = |\alpha| \|e_1\|_2 \delta = \delta \|x\|_2$ and $\|x\|_1/K \leq \|x\|_2 \leq K\|x\|_1$. Hence, the two norms are equivalent.

Case 3: $n \geq 2$. Let $\{e_1, \dots, e_n\} \subseteq \mathcal{X}$ be a basis for $\mathcal{X}$. $\forall x \in \mathcal{X}$, by Definition 6.50 and Corollary 6.47, $\exists! \alpha_1, \dots, \alpha_n \in \mathbb{K}$ such that $x = \sum_{i=1}^n \alpha_i e_i$. We will show that $\exists K_1 \in (0, \infty) \subset \mathbb{R}$ such that $(\sum_{i=1}^n |\alpha_i|)/K_1 \leq \|x\|_1 \leq K_1(\sum_{i=1}^n |\alpha_i|)$. By a similar argument, $\exists K_2 \in (0, \infty) \subset \mathbb{R}$ such that $(\sum_{i=1}^n |\alpha_i|)/K_2 \leq \|x\|_2 \leq K_2(\sum_{i=1}^n |\alpha_i|)$. Then, $\|x\|_1/K \leq \|x\|_2 \leq K\|x\|_1$, where $K := K_1 K_2$. Hence, the two norms are equivalent.

Clearly, $e_i \neq \vartheta$ and $\|e_i\|_1 > 0$, $i = 1, \dots, n$. Let $\delta := \max_{1 \leq i \leq n} \|e_i\|_1 \in (0, \infty) \subset \mathbb{R}$. Then, $\|x\|_1 \leq \sum_{i=1}^n |\alpha_i| \|e_i\|_1 \leq \delta(\sum_{i=1}^n |\alpha_i|)$. Define $M_i = \text{span}(\{e_1, \dots, e_n\} \setminus \{e_i\})$, $i = 1, \dots, n$, and $\delta_i = \text{dist}(e_i, M_i)$ (with respect to $\|\cdot\|_1$). Since $\{e_1, \dots, e_n\}$ is a basis for $\mathcal{X}$, then $e_i \notin M_i$. $\forall i \in \{1, \dots, n\}$, by Theorem 7.36, M_i is complete (with respect to $\|\cdot\|_1$). By Proposition 4.39, M_i is closed (with respect to $\|\cdot\|_1$). By Proposition 4.10, we have $\delta_i \in (0, \infty) \subset \mathbb{R}$. Let $\bar{\delta} := \min\{\delta_1, \dots, \delta_n\} \in (0, \infty) \subset \mathbb{R}$. Then, $\|x\|_1 = \|\sum_{i=1}^n \alpha_i e_i\|_1 \geq |\alpha_i| \delta_i \geq |\alpha_i| \bar{\delta}$, $\forall i \in \{1, \dots, n\}$. Hence, we have $\|x\|_1 \geq (\bar{\delta}/n)(\sum_{i=1}^n |\alpha_i|)$. Let $K_1 = \max\{\delta, n/\bar{\delta}\} > 0$. Then, we have $(\sum_{i=1}^n |\alpha_i|)/K_1 \leq \|x\|_1 \leq K_1(\sum_{i=1}^n |\alpha_i|)$.

This completes the proof of the theorem. $\square$

7.5 Compactness

Definition 7.39 Let $\mathcal{X}$ be a normed linear space and $S \subseteq \mathcal{X}$. S is said to be compact if S together with the natural metric forms a compact metric space.

Proposition 5.29 says that a continuous function achieves its minimum and maximum on nonempty countably compact spaces. This has immediate generalization to infinite-dimensional spaces. Yet, the compactness restriction is so severe in infinite-dimensional spaces that it is applicable in minority of problems.

Lemma 7.40 *Let $(\mathcal{X}, \mathbb{C})$ be a vector space. Then $\mathcal{X}$ is also a vector space over the field $\mathbb{R}$. Furthermore, if $\mathcal{X} := (\mathcal{X}, \mathbb{C}, \|\cdot\|)$ is a normed linear space, then $\mathcal{X}_{\mathbb{R}} := (\mathcal{X}, \mathbb{R}, \|\cdot\|)$ is also a normed linear space. $\mathcal{X}$ and $\mathcal{X}_{\mathbb{R}}$ are isometric and admit the same metric space properties. In particular, $\mathcal{X}$ is a Banach space if, and only if, $\mathcal{X}_{\mathbb{R}}$ is a Banach space.*

Proof Let $(\mathcal{X}, \mathbb{C})$ be a vector space. Note that $\mathbb{R} \subset \mathbb{C}$. Then, $\forall x, y, z \in \mathcal{X}, \forall \alpha, \beta \in \mathbb{R}$, we have (i) $x+y = y+x$; (ii) $(x+y)+z = x+(y+z)$; (iii) $\vartheta_{\mathcal{X}} + x = x$; (iv) $\alpha(x+y) = \alpha x + \alpha y$; (v) $(\alpha + \beta)x = \alpha x + \beta x$; (vi) $(\alpha\beta)x = \alpha(\beta x)$; (vii) $0 \cdot x = \vartheta_{\mathcal{X}}$ and $1 \cdot x = x$. Hence, $(\mathcal{X}, \mathbb{R})$ is a vector space.

Let $\mathcal{X} := (\mathcal{X}, \mathbb{C}, \|\cdot\|)$ be a normed linear space. Then, $(\mathcal{X}, \mathbb{R})$ is a vector space. $\forall x, y \in \mathcal{X}, \forall \alpha \in \mathbb{R}$, we have (i) $\|x\| \in [0, \infty) \subset \mathbb{R}$ and $\|x\| = 0 \Leftrightarrow x = \vartheta_{\mathcal{X}}$; (ii) $\|x+y\| \leq \|x\| + \|y\|$; $\|\alpha x\| = |\alpha| \|x\|$. Hence, $\mathcal{X}_{\mathbb{R}} := (\mathcal{X}, \mathbb{R}, \|\cdot\|)$ is a normed linear space.

Clearly, $\text{id}_{\mathcal{X}} : \mathcal{X} \rightarrow \mathcal{X}_{\mathbb{R}}$ is an isometry and the natural metrics induced by $\mathcal{X}$ and $\mathcal{X}_{\mathbb{R}}$ on $\mathcal{X}$ are identical. Hence, $\mathcal{X}$ and $\mathcal{X}_{\mathbb{R}}$ admits the same metric space properties. Then, $\mathcal{X}$ is a Banach space if, and only if, $\mathcal{X}$ is complete if, and only if, $\mathcal{X}_{\mathbb{R}}$ is complete if, and only if, $\mathcal{X}_{\mathbb{R}}$ is a Banach space.

This completes the proof of the lemma. $\square$

Proposition 7.41 *Let $K \subseteq \mathbb{C}^n$ with $n \in \mathbb{Z}_+$. Then, K is compact if, and only if, K is closed and bounded.*

Proof “Necessity” By Proposition 5.38, K is complete and totally bounded. Then, K is bounded. By Proposition 4.39, K is closed.

“Sufficiency” Let $|\cdot|$ be the norm on $\mathbb{C}^n$. By Lemma 7.40, $\mathcal{X} := (\mathbb{C}^n, \mathbb{R}, |\cdot|)$ admits the same metric space property as $\mathbb{C}^n$. Then, $K \subseteq \mathcal{X}$ is closed and bounded. Note that $\mathcal{X}$ is isometrically isomorphic to $\mathbb{R}^{2n}$. Hence, $K \subseteq \mathcal{X}$ is compact. Then, $K \subseteq \mathbb{C}^n$ is compact. This completes the proof of the proposition. $\square$

Proposition 7.42 *Let $\mathcal{X}$ be a finite-dimensional normed linear space over the field $\mathbb{K}$. $K \subseteq \mathcal{X}$ is compact if, and only if, K is closed and bounded.*

Proof “Necessity” By Proposition 5.38, K is complete and totally bounded. Then, K is bounded. By Proposition 4.39, K is closed.

“Sufficiency” Let $S \subseteq \mathcal{X}$ be a basis for $\mathcal{X}$, that is, S is linearly independent and $\text{span}(S) = \mathcal{X}$. Since $\mathcal{X}$ is finite dimensional, then let $n \in \mathbb{Z}_+$ be the dimension of $\mathcal{X}$, which is well-defined by Theorem 6.51. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $n = 0$; Case 2:

$n \in \mathbb{N}$. Case 1: $n = 0$. Then, $\mathcal{X}$ is a singleton set. Clearly K is compact. Case 2: $n \in \mathbb{N}$. Let $S = \{e_1, \dots, e_n\}$. $\forall x \in \mathcal{X}$, x can be uniquely expressed as $\sum_{i=1}^n \alpha_i e_i$ for some $\alpha_1, \dots, \alpha_n \in \mathbb{K}$. This allows us to define a bijective mapping $\psi : \mathcal{X} \rightarrow \mathbb{K}^n$ by $\psi(x) = (\alpha_1, \dots, \alpha_n)$, $\forall x = \sum_{i=1}^n \alpha_i e_i \in \mathcal{X}$. Define a alternative norm $\|\cdot\|_1$ on $\mathcal{X}$ by $\|x\|_1 = \|\sum_{i=1}^n \alpha_i e_i\|_1 = \sqrt{\sum_{i=1}^n |\alpha_i|^2}$, $\forall x \in \mathcal{X}$. It is easy to show that $\|\cdot\|_1$ is a norm. By Theorem 7.38, there exists $\xi \in [1, \infty) \subset \mathbb{R}$ such that $\|x\|_1 / \xi \leq \|x\| \leq \xi \|x\|_1$, $\forall x \in \mathcal{X}$. Hence, ψ is a homeomorphism. By Proposition 3.10, $\psi(K)$ is closed. By the equivalence of the two norms, $\psi(K)$ is bounded. By Proposition 5.40 or 7.41, $\psi(K)$ is compact. By Proposition 5.7, K is compact.

This completes the proof of the proposition. $\square$

7.6 Quotient Spaces

Proposition 7.43 *Let $M \subseteq \mathcal{X}$ be a subspace of a vector space $\mathcal{X}$ over a field $\mathcal{F} := (F, +, \cdot, \mathbf{0}, \mathbf{1})$. $x_1, x_2 \in \mathcal{X}$ are said to be equivalent modulo M if $x_1 - x_2 \in M$. This equivalence relationship (easy to show) partitions the space $\mathcal{X}$ into disjoint subsets, or classes, of equivalent elements: namely, the linear varieties that are distinct translations of the subspace M . These classes are often called the cosets of M . $\forall x \in \mathcal{X}$, there is a unique coset of M , $[x] (= x + M)$, such that $x \in [x]$. The quotient of $\mathcal{X}$ modulo M is defined to be the set N of all cosets of M . Define vector addition and scalar multiplication on N by, $\forall [x_1], [x_2] \in N$, $\forall \alpha \in \mathcal{F}$, $[x_1] + [x_2] := [x_1 + x_2]$ and $\alpha[x_1] := [\alpha x_1]$. Let the null vector on N be $[\vartheta_{\mathcal{X}}]$. Then, N together with the vector addition, scalar multiplication, and the null vector form a vector space over $\mathcal{F}$. This vector space will be called the quotient space of $\mathcal{X}$ modulo M and denoted by $\mathcal{X}/M$. Define a function $\phi : \mathcal{X} \rightarrow \mathcal{X}/M$ by $\phi(x) = [x]$, $\forall x \in \mathcal{X}$. Then, ϕ is a linear function and will be called the natural homomorphism.*

Proof Fix any $[x_1], [x_2], [x_3] \in \mathcal{X}/M$ and any $\alpha, \beta \in \mathcal{F}$. We will first show that vector addition and scalar multiplication are uniquely defined. $\forall y_1 \in [x_1]$, $\forall y_2 \in [x_2]$, we have $x_1 - y_1, x_2 - y_2 \in M$; and $(x_1 + x_2) - (y_1 + y_2) \in M$, since M is a subspace. Then, $[x_1] + [x_2] = [x_1 + x_2] = [y_1 + y_2] = [y_1] + [y_2]$. Hence, the vector addition is uniquely defined. $\alpha x_1 - \alpha y_1 = \alpha(x_1 - y_1) \in M$. Then, $\alpha[x_1] = [\alpha x_1] = [\alpha y_1] = \alpha[y_1]$. Hence, the scalar multiplication is uniquely defined.

(i) $[x_1] + [x_2] = [x_1 + x_2] = [x_2 + x_1] = [x_2] + [x_1]$; (ii) $([x_1] + [x_2]) + [x_3] = [(x_1 + x_2) + x_3] = [x_1 + (x_2 + x_3)] = [x_1] + ([x_2] + [x_3])$; (iii) $[x_1] + [\vartheta_{\mathcal{X}}] = [x_1 + \vartheta_{\mathcal{X}}] = [x_1]$; (iv) $\alpha([x_1] + [x_2]) = \alpha[x_1 + x_2] = [\alpha(x_1 + x_2)] = [\alpha x_1 + \alpha x_2] = [\alpha x_1] + [\alpha x_2] = \alpha[x_1] + \alpha[x_2]$; (v) $(\alpha + \beta)[x_1] = [(\alpha + \beta)x_1] = [\alpha x_1 + \beta x_1] = [\alpha x_1] + [\beta x_1] = \alpha[x_1] + \beta[x_1]$; (vi) $(\alpha\beta)[x_1] = [(\alpha\beta)x_1] = [\alpha(\beta x_1)] = \alpha[\beta x_1] = \alpha(\beta[x_1])$; (vii) $\mathbf{0}[x_1] = [\mathbf{0}x_1] = [\vartheta_{\mathcal{X}}]$; $\mathbf{1}[x_1] = [x_1]$. Hence, $\mathcal{X}/M$ is a vector space over $\mathcal{F}$.

Clearly, ϕ is a linear function. This completes the proof of the proposition. $\square$

Proposition 7.44 *Let $\mathcal{X} := (\mathcal{X}, \mathbb{K}, \|\cdot\|_{\mathcal{X}})$ be a normed linear space and $M \subseteq \mathcal{X}$ be a closed subspace, and $\mathcal{X}/M$ be the quotient space of $\mathcal{X}$ modulo M . Define a norm $\|\cdot\|$ on $\mathcal{X}/M$ by, $\forall [x] \in \mathcal{X}/M$, $\|[x]\| := \inf_{m \in M} \|x - m\|_{\mathcal{X}} = \text{dist}(x, M)$. Then, $\mathcal{X}/M := (\mathcal{X}/M, \mathbb{K}, \|\cdot\|)$ is a normed linear space, which will be called the quotient space of $\mathcal{X}$ modulo M .*

Proof By Proposition 7.43, $\mathcal{X}/M$ is a vector space over $\mathbb{K}$. Here, we only need to show that $\|\cdot\|$ defines a norm on $\mathcal{X}/M$. First, we show that $\|\cdot\| : \mathcal{X}/M \rightarrow [0, \infty) \subset \mathbb{R}$ is uniquely defined. $\forall [x] \in \mathcal{X}/M$, $\forall y \in [x]$, $y - x \in M$. Then, we have $\|[y]\| = \inf_{m \in M} \|y - m\|_{\mathcal{X}} = \inf_{m \in M} \|x - m\|_{\mathcal{X}} = \|[x]\|$. Hence, $\|\cdot\|$ is uniquely defined.

Next, we show that $\|\cdot\|$ defines a norm on $\mathcal{X}/M$. $\forall [x_1], [x_2] \in \mathcal{X}/M$, $\forall \alpha \in \mathbb{K}$, we have (i) $\|[x_1]\| \in [0, \infty) \subset \mathbb{R}$, since $\|[x_1]\| \leq \|x_1\|_{\mathcal{X}} < +\infty$. $[x_1] = [\vartheta_{\mathcal{X}}] \Rightarrow \|[x_1]\| = \inf_{m \in M} \|m\|_{\mathcal{X}} = 0$ and $\|[x_1]\| = 0 \Rightarrow \text{dist}(x_1, M) = 0$, by Proposition 4.10, we have $x_1 \in M \Rightarrow [x_1] = [\vartheta_{\mathcal{X}}]$. Hence $[x_1] = [\vartheta_{\mathcal{X}}] \Leftrightarrow \|[x_1]\| = 0$.

(ii) $\|[x_1] + [x_2]\| = \|[x_1 + x_2]\| = \inf_{m \in M} \|x_1 + x_2 - m\|_{\mathcal{X}}$. Note that $\|[x_i]\| = \inf_{m \in M} \|x_i - m\|_{\mathcal{X}}$, $i = 1, 2$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $i = 1, 2$, $\exists m_i \in M$ such that $\|x_i - m_i\|_{\mathcal{X}} < \|[x_i]\| + \epsilon$. Then, we have $\|x_1 + x_2 - m_1 - m_2\|_{\mathcal{X}} \leq \|x_1 - m_1\|_{\mathcal{X}} + \|x_2 - m_2\|_{\mathcal{X}} \leq \|[x_1]\| + \|[x_2]\| + 2\epsilon$. Hence, we have $\|[x_1] + [x_2]\| \leq \|[x_1]\| + \|[x_2]\| + 2\epsilon$. By the arbitrariness of ϵ , we have $\|[x_1] + [x_2]\| \leq \|[x_1]\| + \|[x_2]\|$.

(iii) $\|\alpha[x_1]\| = \|\alpha x_1\| = \inf_{m \in M} \|\alpha x_1 - m\|_{\mathcal{X}}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\alpha = 0$; Case 2: $\alpha \neq 0$. Case 1: $\alpha = 0$. $\|\alpha[x_1]\| = \inf_{m \in M} \|m\|_{\mathcal{X}} = 0 = |\alpha| \|[x_1]\|$. Case 2: $\alpha \neq 0$. $\|\alpha[x_1]\| = \inf_{m \in M} \|\alpha x_1 - \alpha m\|_{\mathcal{X}} = \inf_{m \in M} |\alpha| \|x_1 - m\|_{\mathcal{X}} = |\alpha| \|[x_1]\|$. Hence, in both cases, we have $\|\alpha[x_1]\| = |\alpha| \|[x_1]\|$.

This shows that $\|\cdot\|$ is a norm on $\mathcal{X}/M$. Hence, $\mathcal{X}/M$ is a normed linear space. This completes the proof of the proposition. $\square$

Proposition 7.45 *Let $\mathcal{X}$ be a Banach space and $M \subseteq \mathcal{X}$ be a closed subspace. Then, the quotient space $\mathcal{X}/M$ is a Banach space.*

Proof By Proposition 7.44, $\mathcal{X}/M$ is a normed linear space. Here, we need only to show that $\mathcal{X}/M$ is complete. Let $\|\cdot\|_{\mathcal{X}}$ be the norm on $\mathcal{X}$ and $\|\cdot\|$ be the norm on $\mathcal{X}/M$. We will prove this using Proposition 7.27. Fix any absolutely summable series $([x_n])_{n=1}^{\infty} \subseteq \mathcal{X}/M$. Then, $\sum_{n=1}^{\infty} \|[x_n]\| \in \mathbb{R}$. $\forall n \in \mathbb{N}$, $\exists y_n \in [x_n]$ such that $\|y_n\|_{\mathcal{X}} < \|[x_n]\| + 2^{-n}$. Then, $\sum_{i=1}^{\infty} \|y_i\|_{\mathcal{X}} < \sum_{i=1}^{\infty} (\|[x_i]\| + 2^{-i}) = \sum_{i=1}^{\infty} \|[x_i]\| + 1$. Then, $(y_n)_{n=1}^{\infty} \subseteq \mathcal{X}$ is absolutely summable. By Proposition 7.27, $\sum_{n=1}^{\infty} y_n = y \in \mathcal{X}$, since $\mathcal{X}$ is complete. Now, it is easy to show that $\sum_{n=1}^{\infty} [x_n] = \sum_{n=1}^{\infty} [y_n] = [y] \in \mathcal{X}/M$. Hence, $\mathcal{X}/M$ is complete by Proposition 7.27. This completes the proof of the proposition. $\square$

Definition 7.46 Let $(\mathcal{X}, \mathbb{K})$ be a vector space and $\|\cdot\| : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$. Assume that $\|\cdot\|$ satisfies (ii) and (iii) of Definition 7.1 and $\|\vartheta\| = 0$, but not necessarily (i). Then, $\|\cdot\|$ is called a pseudo-norm.

Proposition 7.47 Let $(\mathcal{X}, \mathbb{K})$ be a vector space and $\|\cdot\| : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$ be a pseudo-norm on $\mathcal{X}$. Then, the set $M := \{x \in \mathcal{X} \mid \|x\| = 0\}$ is a subspace of $(\mathcal{X}, \mathbb{K})$. On the quotient space $\mathcal{X}/M$, define a norm $\|\cdot\|_1 : \mathcal{X}/M \rightarrow [0, \infty) \subset \mathbb{R}$ by $\|[x]\|_1 = \|x\|$, $\forall [x] \in \mathcal{X}/M$. Then, the space $(\mathcal{X}/M, \mathbb{K}, \|\cdot\|_1)$ is a normed linear space, which will be called the quotient space of $(\mathcal{X}, \mathbb{K})$ modulo $\|\cdot\|$.

Proof We need the following claim.

Claim 7.47.1 $\forall x \in \mathcal{X}$, $\forall m \in M$, we have $\|x\| = \|x + m\|$.

Proof of claim: Note that $\|x\| \leq \|x + m\| + \|-m\| = \|x + m\| \leq \|x\| + \|m\| = \|x\|$. This completes the proof of the claim. $\square$

Clearly, $\vartheta_{\mathcal{X}} \in M \neq \emptyset$. $\forall m_1, m_2 \in M$, $\forall \alpha \in \mathbb{K}$, we have $\alpha m_1 \in M$, by the properties of the pseudo-norm, and $m_1 + m_2 \in M$ by Claim 7.47.1. Hence, M is a subspace of $(\mathcal{X}, \mathbb{K})$. By Proposition 7.43, $\mathcal{X}/M$ is a vector space over $\mathbb{K}$.

By Claim 7.47.1, $\|\cdot\|_1 : \mathcal{X}/M \rightarrow [0, \infty) \subset \mathbb{R}$ is uniquely defined.

Next, we show that $\|\cdot\|_1$ is a norm on $\mathcal{X}/M$. $\forall [x_1], [x_2] \in \mathcal{X}/M$, $\forall \alpha \in \mathbb{K}$, $\|[x_1]\|_1 = \|x_1\| \in [0, \infty) \subset \mathbb{R}$. $[x_1] = [\vartheta_{\mathcal{X}}]$ implies that $\|[x_1]\|_1 = \|\vartheta_{\mathcal{X}}\|_1 = \|\vartheta_{\mathcal{X}}\| = 0$. $\|[x_1]\|_1 = 0$ implies that $\|x_1\| = 0$ and $x_1 \in M$, which further implies that $x_1 \in [\vartheta_{\mathcal{X}}]$ and $[x_1] = [\vartheta_{\mathcal{X}}]$. $\|[x_1] + [x_2]\|_1 = \|[x_1 + x_2]\|_1 = \|x_1 + x_2\| \leq \|x_1\| + \|x_2\| = \|[x_1]\|_1 + \|[x_2]\|_1$. $\|\alpha[x_1]\|_1 = \|\alpha x_1\|_1 = \|\alpha x_1\| = |\alpha| \|x_1\| = |\alpha| \|[x_1]\|_1$. Hence, $\|\cdot\|_1$ is a norm on $\mathcal{X}/M$. Therefore, $(\mathcal{X}/M, \mathbb{K}, \|\cdot\|_1)$ is a normed linear space. This completes the proof of the proposition. $\square$

7.7 The Stone-Weierstrass Theorem

Definition 7.48 Let X be a set, $f : X \rightarrow \mathbb{R}$, and $g : X \rightarrow \mathbb{R}$. $f \vee g : X \rightarrow \mathbb{R}$ and $f \wedge g : X \rightarrow \mathbb{R}$ are defined by $(f \vee g)(x) = \max\{f(x), g(x)\}$ and $(f \wedge g)(x) = \min\{f(x), g(x)\}$, $\forall x \in X$.

Proposition 7.49 Let $\mathcal{X}$ be a topological space and $f : \mathcal{X} \rightarrow \mathbb{R}$ and $g : \mathcal{X} \rightarrow \mathbb{R}$ be continuous at $x_0 \in \mathcal{X}$. Then, $f \vee g$ and $f \wedge g$ are continuous at x_0 . If furthermore f and g are continuous, then $f \vee g$ and $f \wedge g$ are continuous.

Proof This is straightforward. $\square$

Example 7.50 Let $\mathcal{X}$ be a topological space and $\mathcal{Y}$ be a normed linear space over the field $\mathbb{K}$. Let $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})$ be the vector space defined in

Example 6.20. Let $V := \{f \in \mathcal{M}(\mathcal{X}, \mathcal{Y}) \mid f \text{ is continuous}\}$. Then, by Propositions 3.12, 3.32, 6.25, and 7.23, V is a subspace of $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})$. This subspace will be denoted by $\mathcal{C}_v(\mathcal{X}, \mathcal{Y})$. $\diamond$

Definition 7.51 Let $\mathcal{X}$ be a topological space, $\mathcal{C}_v(\mathcal{X}, \mathbb{R})$ be the vector space defined in Example 7.50. $L \subseteq \mathcal{C}_v(\mathcal{X}, \mathbb{R})$ is said to be a lattice if $L \neq \emptyset$ and $\forall f, g \in L, f \vee g \in L$ and $f \wedge g \in L$. A subspace $M \subseteq \mathcal{C}_v(\mathcal{X}, \mathbb{R})$ is said to be an algebra if $\forall f, g \in M, fg \in M$, where fg is the product of functions f and g .

Proposition 7.52 Let $\mathcal{X} := (X, \mathcal{O})$ be a compact space, $\mathcal{C}(\mathcal{X}, \mathbb{R})$ be the Banach space as defined in Example 7.32, and $L \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ be a lattice. Assume $h : \mathcal{X} \rightarrow \mathbb{R}$ defined by $h(x) = \inf_{f \in L} f(x)$, $\forall x \in \mathcal{X}$ is continuous. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists g \in L$ such that $0 \leq g(x) - h(x) < \epsilon, \forall x \in \mathcal{X}$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X} \neq \emptyset$. Case 1: $\mathcal{X} = \emptyset$. $\mathcal{C}(\mathcal{X}, \mathbb{R})$ is a singleton set and $L = \mathcal{C}(\mathcal{X}, \mathbb{R})$ since L is a lattice and therefore nonempty. Then, the result holds. Case 2: $\mathcal{X} \neq \emptyset$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall x \in \mathcal{X}$, since $h(x) \in \mathbb{R}$, then $\exists f_x \in L$ such that $0 \leq f_x(x) - h(x) < \epsilon/3$. Since f_x and h are continuous, then $\exists O_x \in \mathcal{O}$ with $x \in O_x$ such that $\forall y \in O_x$, we have $|f_x(y) - f_x(x)| < \epsilon/3$ and $|h(y) - h(x)| < \epsilon/3$. Then, $|f_x(y) - h(y)| \leq |f_x(y) - f_x(x)| + |f_x(x) - h(x)| + |h(x) - h(y)| < \epsilon, \forall y \in O_x$. Then, $\mathcal{X} \subseteq \bigcup_{x \in \mathcal{X}} O_x$. Since $\mathcal{X}$ is compact, then there exists a finite set $X_N \subseteq \mathcal{X}$ such that $\mathcal{X} \subseteq \bigcup_{x \in X_N} O_x$. Since $X \neq \emptyset$ then X_N must be nonempty. Let $g := \bigwedge_{x \in X_N} f_x \in L$. $\forall x \in \mathcal{X}, \exists x_0 \in X_N$ such that $x \in O_{x_0}$ and $0 \leq g(x) - h(x) \leq f_{x_0}(x) - h(x) < \epsilon$. This completes the proof of the proposition. $\square$

Proposition 7.53 Let $\mathcal{X} := (X, \mathcal{O})$ be a compact space, $\mathcal{C}(\mathcal{X}, \mathbb{R})$ be the Banach space as defined in Example 7.32, and $L \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ be a lattice satisfying the following conditions:

- (i) L separates points, that is $\forall x, y \in \mathcal{X}$ with $x \neq y, \exists f \in L$ such that $f(x) \neq f(y)$;
- (ii) $\forall f \in L, \forall c \in \mathbb{R}, c + f, cf \in L$.

Then, $\forall h \in \mathcal{C}(\mathcal{X}, \mathbb{R}), \forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists $g \in L$ such that $0 \leq g(x) - h(x) < \epsilon, \forall x \in \mathcal{X}$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X} \neq \emptyset$. Case 1: $\mathcal{X} = \emptyset$. $\mathcal{C}(\mathcal{X}, \mathbb{R})$ is a singleton set and $L = \mathcal{C}(\mathcal{X}, \mathbb{R})$ since L is a lattice and therefore nonempty. Then, the result holds.

Case 2: $\mathcal{X} \neq \emptyset$. We need the following two results.

Claim 7.53.1 $\forall a, b \in \mathbb{R}, \forall x_1, x_2 \in \mathcal{X}$ with $x_1 \neq x_2, \exists f \in L$ such that $f(x_1) = a$ and $f(x_2) = b$.

Proof of claim: Let $g \in L$ be such that $g(x_1) \neq g(x_2)$, which exists by (i). Let

$$f = \frac{a - b}{g(x_1) - g(x_2)}g + \frac{bg(x_1) - ag(x_2)}{g(x_1) - g(x_2)} \in L$$

Then, f is the desired function. $\square$

Claim 7.53.2 $\forall a, b \in \mathbb{R}$ with $a \leq b, \forall$ closed set $F \subseteq \mathcal{X}, \forall x_0 \in \mathcal{X}$ with $x_0 \notin F, \exists f \in L$ such that $f(x_0) = a, f(x) \geq a, \forall x \in \mathcal{X}$, and $f(x) > b, \forall x \in F$.

Proof of claim: $\forall x \in F$, we have $x \neq x_0$. By Claim 7.53.1, $\exists f_x \in L$ such that $f_x(x_0) = a$ and $f_x(x) = b + 1$. Let $O_x := \{\bar{x} \in \mathcal{X} \mid f_x(\bar{x}) > b\}$. Then, $O_x \in \mathcal{O}$ since f_x is continuous. Clearly, $x \in O_x$. $F \subseteq \bigcup_{x \in F} O_x$. By Proposition 5.5, F is compact. Then, there exists a finite set $F_N \subseteq F$ such that $F \subseteq \bigcup_{x \in F_N} O_x$. Take $g \in L \neq \emptyset$. Then, $a = 0g + a \in L$ by (ii). Let $f := a \vee (\bigvee_{x \in F_N} f_x) \in L$. Clearly, $f(x_0) = a$ and $f(x) \geq a, \forall x \in \mathcal{X}$. $\forall x \in F, \exists x_0 \in F_N$ such that $x \in O_{x_0}$. Then, $f(x) \geq f_{x_0}(x) > b$. Hence, f is the desired function. This completes the proof of the claim. $\square$

$\forall h \in \mathcal{C}(\mathcal{X}, \mathbb{R})$, let $\bar{L} := \{f \in L \mid f(x) \geq h(x), \forall x \in \mathcal{X}\}$. By the continuity of h , the compactness of $\mathcal{X}$, and Proposition 5.29, $\exists b \in \mathbb{R}$ such that $h(x) \leq b, \forall x \in \mathcal{X}$. Then, the constant function $b \in \bar{L} \neq \emptyset$ ($b \in L$ by (ii)). $\forall f_1, f_2 \in \bar{L}$, it is easy to show that $f_1 \vee f_2 \in \bar{L}$ and $f_1 \wedge f_2 \in \bar{L}$. Hence, $\bar{L}$ is a lattice. We will show that $h(x) = \inf_{f \in \bar{L}} f(x), \forall x \in \mathcal{X}$. Then, the result follows from Proposition 7.52. $\forall x_0 \in \mathcal{X}, \forall \eta \in (0, \infty) \subset \mathbb{R}$, let $F_{x_0, \eta} := \{x \in \mathcal{X} \mid h(x) \geq h(x_0) + \eta\}$. By the continuity of h and Proposition 3.10, $F_{x_0, \eta}$ is closed. Clearly, $x_0 \notin F_{x_0, \eta}$. Now, by Claim 7.53.2, $\exists f_{x_0, \eta} \in L$ such that $f_{x_0, \eta}(x_0) = h(x_0) + \eta, f_{x_0, \eta}(x) \geq h(x_0) + \eta, \forall x \in \mathcal{X}$, and $f_{x_0, \eta}(x) > b, \forall x \in F_{x_0, \eta}$. It is clear that $f_{x_0, \eta}(x) > h(x), \forall x \in \mathcal{X}$. Hence, $f_{x_0, \eta} \in \bar{L}$. Then, $\inf_{f \in \bar{L}} f(x_0) \leq f_{x_0, \eta}(x_0) = h(x_0) + \eta$. By the definition of $\bar{L}$ and the arbitrariness of η , we have $h(x_0) \leq \inf_{f \in \bar{L}} f(x_0) \leq h(x_0)$. Hence, we have $h(x_0) = \inf_{f \in \bar{L}} f(x_0)$.

This completes the proof of the proposition. $\square$

Lemma 7.54 $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists a polynomial $P(s)$ in one variable such that $|P(s) - |s|| < \epsilon, \forall s \in [-1, 1] \subset \mathbb{R}$.

Proof This result is a special case of Bernstein Approximation Theorem (Bartle, 1976, pg. 171). For an generalization of the Bernstein Approximation Theorem to multi-variable case, see Page 527. $\square$

Lemma 7.55 Let $\mathcal{X}$ be a nonempty compact space and $\mathcal{A} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ be an algebra. Then, $\overline{\mathcal{A}}$ is an algebra.

Proof Note that $\mathcal{A}$ is a subspace of $\mathcal{C}(\mathcal{X}, \mathbb{R})$ by Definition 7.51. Then, by Proposition 7.17, $\overline{\mathcal{A}}$ is a subspace of $\mathcal{C}(\mathcal{X}, \mathbb{R})$. $\forall f, g \in \overline{\mathcal{A}}, \forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\exists \bar{f}, \bar{g} \in \mathcal{A}$ such that $\|f - \bar{f}\| < \frac{\epsilon}{2(1+\|g\|)}$ and $\|g - \bar{g}\| < \frac{\epsilon}{2(1+\|f\|)}$. Then, $\bar{f}\bar{g} \in \mathcal{A}$ since $\mathcal{A}$ is an algebra, and $\|fg - \bar{f}\bar{g}\| = \max_{x \in \mathcal{X}} |f(x)g(x) - \bar{f}(x)\bar{g}(x)| \leq \max_{x \in \mathcal{X}} (|f(x)g(x) - f(x)\bar{g}(x)| + |f(x)\bar{g}(x) - \bar{f}(x)\bar{g}(x)|) \leq \max_{x \in \mathcal{X}} |f(x)| |g(x) - \bar{g}(x)| + \max_{x \in \mathcal{X}} |\bar{g}(x)| \cdot |f(x) - \bar{f}(x)| \leq \|f\| \|g - \bar{g}\| + \|g + \bar{g} - g\| \|f - \bar{f}\| < \frac{\epsilon \|f\|}{2(1+\|f\|)} + \frac{\epsilon(\|g\| + \|g - \bar{g}\|)}{2(1+\|g\|)} < \epsilon$. Hence, $fg \in \overline{\mathcal{A}}$ by the arbitrariness of ϵ . This shows that $\overline{\mathcal{A}}$ is an algebra. This completes the proof of the lemma. $\square$

Theorem 7.56 (Stone-Weierstrass Theorem) *Let $\mathcal{X} := (X, \mathcal{O})$ be a compact space and $\mathcal{C}(\mathcal{X}, \mathbb{R})$ be the Banach space defined in Example 7.32. Assume that $\mathcal{A} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ is an algebra and satisfies*

- (i) $\mathcal{A}$ separates points, that is, $\forall x_1, x_2 \in \mathcal{X}$ with $x_1 \neq x_2$, then $\exists f \in \mathcal{A}$ such that $f(x_1) \neq f(x_2)$;
- (ii) $\mathcal{A}$ contains all constant functions.

Then, $\overline{\mathcal{A}} = \mathcal{C}(\mathcal{X}, \mathbb{R})$, that is $\mathcal{A}$ is dense in $\mathcal{C}(\mathcal{X}, \mathbb{R})$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X} \neq \emptyset$. Case 1: $\mathcal{X} = \emptyset$. $\mathcal{C}(\mathcal{X}, \mathbb{R})$ is a singleton set and $\mathcal{A} = \mathcal{C}(\mathcal{X}, \mathbb{R})$ since $\mathcal{A}$ is a subspace and therefore nonempty. Then, the result holds.

Case 2: $\mathcal{X} \neq \emptyset$. By Lemma 7.55, $\overline{\mathcal{A}}$ is an algebra. We need the following claim.

Claim 7.56.1 $\forall f \in \overline{\mathcal{A}}$, then $|f| \in \overline{\mathcal{A}}$.

Proof of claim: Fix $f \in \overline{\mathcal{A}}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists g \in \mathcal{A}$ such that $\|f - g\| < \epsilon/2$. By Proposition 5.29 and Definitions 5.1 and 5.21, $\exists N \in (0, \infty) \subset \mathbb{R}$ such that $\|g\| < N$. Then, $g/N \in \mathcal{A}$ since $\mathcal{A}$ is an algebra. By Lemma 7.54, $\exists$ a polynomial $P(s)$ such that $|P(s) - |s|| < \frac{\epsilon}{2N}$, $\forall s \in [-1, 1] \subset \mathbb{R}$. Hence, we have $P \circ (g/N) \in \mathcal{A}$ since $\mathcal{A}$ is an algebra. Furthermore, $|P(g(x)/N) - |g(x)|/N| < \frac{\epsilon}{2N}$, $\forall x \in \mathcal{X}$. Let $h = NP \circ (g/N) \in \mathcal{A}$. Then, we have $\|h - |g|\| < \epsilon/2$. Note that $\||f| - h\| \leq \||f| - |g|\| + \||g| - h\| \leq \|f - g\| + \||g| - h\| < \epsilon$. Hence, $|f| \in \overline{\mathcal{A}}$, by the arbitrariness of ϵ . This completes the proof of the claim. $\square$

$\forall f, g \in \overline{\mathcal{A}}$, we have

$$\begin{aligned} f \vee g &= \frac{1}{2}(f + g) + \frac{1}{2}|f - g| \in \overline{\mathcal{A}} \\ f \wedge g &= \frac{1}{2}(f + g) - \frac{1}{2}|f - g| \in \overline{\mathcal{A}} \end{aligned}$$

Hence, $\overline{\mathcal{A}}$ is a lattice. Clearly, $\overline{\mathcal{A}}$ separates points since $\mathcal{A}$ separates points. By Proposition 7.53, $\forall h \in \mathcal{C}(\mathcal{X}, \mathbb{R})$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists $g \in \overline{\mathcal{A}}$ such that $0 \leq g(x) - h(x) < \epsilon$, $\forall x \in \mathcal{X}$. Then, $\|g - h\| < \epsilon$, and $h \in \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$, where the last equality follows from Proposition 3.3. Hence, we have $\overline{\mathcal{A}} = \mathcal{C}(\mathcal{X}, \mathbb{R})$. This completes the proof of the theorem. $\square$

Corollary 7.57 *Let $X \subseteq \mathbb{R}^n$ be a closed and bounded set with the subset topology $\mathcal{O}$, where $n \in \mathbb{Z}_+$, $\mathcal{X} := (X, \mathcal{O})$, and $f : \mathcal{X} \rightarrow \mathbb{R}$ be a continuous function. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists a polynomial $P : \mathcal{X} \rightarrow \mathbb{R}$ in n variables such that $|f(x) - P(x)| < \epsilon$, $\forall x \in \mathcal{X}$.*

Proof By Proposition 5.40, $\mathcal{X}$ is a compact space. Let $\mathcal{C}(\mathcal{X}, \mathbb{R})$ be the Banach space defined in Example 7.32. Then, $f \in \mathcal{C}(\mathcal{X}, \mathbb{R})$. Let $\mathcal{A}$ be all polynomial in n variables on $\mathcal{X}$. Clearly, $\mathcal{A} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ and is a linear subspace of $\mathcal{C}(\mathcal{X}, \mathbb{R})$. Then, it is easy to show that $\mathcal{A}$ is an algebra. $\forall x_1, x_2 \in \mathcal{X}$ with $x_1 \neq x_2$, there exists a coordinate $i_0 \in \{1, \dots, n\}$ such that $\pi_{i_0}(x_1) \neq \pi_{i_0}(x_2)$. Then, the polynomial $f \in \mathcal{A}$ given by $f(x) = \pi_{i_0}(x)$, $\forall x \in \mathcal{X}$, separates x_1 and x_2 . Hence, $\mathcal{A}$ separates points. Clearly, $\mathcal{A}$ contains all the constant functions. By the Stone-Weierstrass Theorem, $\overline{\mathcal{A}} = \mathcal{C}(\mathcal{X}, \mathbb{R})$. This completes the proof of the corollary. $\square$

Corollary 7.58 *Let $S_1 = [0, 2\pi] \subset \mathbb{R}$ and $f \in \mathcal{C}(S_1, \mathbb{R})$. Assume that $f(0) = f(2\pi)$. Let $\mathcal{M} = \{g \in \mathcal{C}(S_1, \mathbb{R}) \mid g(x) = \cos(nx), \forall x \in S_1 \text{ or } g(x) = \sin(nx), \forall x \in S_1, \text{ where } n \in \mathbb{Z}_+\}$. Let $\mathcal{A} = \text{span}(\mathcal{M}) \subseteq \mathcal{C}(S_1, \mathbb{R})$. Then, $f \in \overline{\mathcal{A}}$.*

Proof Let $S_2 \subset \mathbb{R}^2$ be the unit circle: $S_2 := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1^2 + y_2^2 = 1\}$. Define a mapping $\Psi : S_1 \rightarrow S_2$ by $\Psi(x) = (\cos(x), \sin(x))$, $\forall x \in S_1$. Clearly, Ψ is surjective and continuous. Note that S_1 is compact and S_2 is Hausdorff. It is obvious that we may define a function $\Phi : S_2 \rightarrow \mathbb{R}$ such that $\Phi \circ \Psi = f$, which is continuous. By Proposition 5.18, we have Φ is continuous. Note that S_2 is closed and bounded in $\mathbb{R}^2$, then S_2 is compact by Proposition 5.40. Hence, $\Phi \in \mathcal{C}(S_2, \mathbb{R})$. By Corollary 7.57, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists a polynomial $P : S_2 \rightarrow \mathbb{R}$ such that $|\Phi(y_1, y_2) - P(y_1, y_2)| < \epsilon$, $\forall y_1, y_2 \in S_2$. Then, $|f(x) - P \circ \Psi(x)| = |\Phi \circ \Psi(x) - P \circ \Psi(x)| < \epsilon$, $\forall x \in S_1$. Note that $P \circ \Psi \in \mathcal{A}$, since, $\forall \gamma, \theta \in \mathbb{R}$,

$$\begin{aligned} (\sin(\theta))^2 &= \frac{1}{2}(1 - \cos(2\theta)); & (\cos(\theta))^2 &= \frac{1}{2}(1 + \cos(2\theta)); \\ \sin(\gamma)\cos(\theta) &= \frac{1}{2}(\sin(\gamma + \theta) + \sin(\gamma - \theta)); \\ \sin(\gamma)\sin(\theta) &= \frac{1}{2}(\cos(\gamma - \theta) - \cos(\gamma + \theta)); \\ \cos(\gamma)\cos(\theta) &= \frac{1}{2}(\cos(\gamma - \theta) + \cos(\gamma + \theta)) \end{aligned}$$

Hence, $f \in \overline{\mathcal{A}}$, by the arbitrariness of ϵ . This completes the proof of the corollary. $\square$

For ease of presentation below, we will define two functions: $\text{sqr} : \mathbb{R} \rightarrow \mathbb{R}$ by $\text{sqr}(x) = x^2$, $\forall x \in \mathbb{R}$, and $\text{sqrt} : [0, \infty) \rightarrow [0, \infty)$ by $\text{sqrt}(x) = \sqrt{x}$, $\forall x \in [0, \infty)$.

Corollary 7.59 *Let $\mathcal{X} := (X, \mathcal{O})$ be a compact space and $\mathcal{A} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ be an algebra that satisfies*

- (i) $\mathcal{A}$ separates points;
- (ii) $\exists f_0 \in \mathcal{A}$ such that $f_0(x) \neq 0$, $\forall x \in \mathcal{X}$.

Then, the constant function $1 \in \overline{\mathcal{A}} = \mathcal{C}(\mathcal{X}, \mathbb{R})$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X} \neq \emptyset$. Case 1: $\mathcal{X} = \emptyset$. $\mathcal{C}(\mathcal{X}, \mathbb{R})$ is a singleton set and $\mathcal{A} = \mathcal{C}(\mathcal{X}, \mathbb{R})$ since $\mathcal{A}$ is a subspace and therefore nonempty. Then, the result holds.

Case 2: $\mathcal{X} \neq \emptyset$. By Lemma 7.55, $\overline{\mathcal{A}}$ is an algebra.

Since $\mathcal{A}$ is an algebra, then, $\text{sqr} \circ f_0 \in \mathcal{A}$, which satisfies that $\text{sqr} \circ f_0(x) > 0$, $\forall x \in \mathcal{X}$. Furthermore, $g_0 := \text{sqr} \circ f_0 / \|\text{sqr} \circ f_0\| \in \mathcal{A}$, which satisfies $g_0 : \mathcal{X} \rightarrow (0, 1] \subset \mathbb{R}$.

By Corollary 7.57, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists a polynomial Q_ϵ in one variable such that $|\sqrt{s} - Q_\epsilon(s)| < \epsilon$, $\forall s \in [0, 1] \subset \mathbb{R}$, and $Q_\epsilon(0) = 0$.

$\forall f : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R}$ with $f \in \overline{\mathcal{A}}$, $Q_\epsilon \circ f \in \overline{\mathcal{A}}$ since $\overline{\mathcal{A}}$ is an algebra and $Q_\epsilon(0) = 0$. Then, $\text{sqrt} \circ f \in \overline{\mathcal{A}} = \overline{\overline{\mathcal{A}}}$. Recursively, we may conclude that $\text{sqrt}^n \circ g_0 \in \overline{\mathcal{A}}$, $\forall n \in \mathbb{N}$.

By Proposition 5.29, $\exists x_m \in \mathcal{X}$ such that $g_0(x) \geq g_0(x_m) =: \gamma > 0$, $\forall x \in \mathcal{X}$. $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $|1 - \gamma^{2^{-n_0}}| < \epsilon$. Then, $\|1 - \text{sqrt}^{n_0} \circ g_0\| < \epsilon$. Hence, the constant function $1 \in \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$.

Clearly, $\overline{\mathcal{A}}$ separates points since $\mathcal{A}$ does. Hence, by Theorem 7.56, $\mathcal{C}(\mathcal{X}, \mathbb{R}) = \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$. This completes the proof of the corollary. $\square$

Proposition 7.60 *Let $\mathcal{X}$ be a compact space and $\mathcal{A} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ be an algebra that separates points on $\mathcal{X}$. Then, either $\overline{\mathcal{A}} = \mathcal{C}(\mathcal{X}, \mathbb{R})$ or $\exists x_0 \in \mathcal{X}$ such that $\overline{\mathcal{A}} = \{f \in \mathcal{C}(\mathcal{X}, \mathbb{R}) \mid f(x_0) = 0\}$.*

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X} \neq \emptyset$. Case 1: $\mathcal{X} = \emptyset$. $\mathcal{C}(\mathcal{X}, \mathbb{R})$ is a singleton set and $\mathcal{A} = \mathcal{C}(\mathcal{X}, \mathbb{R})$ since $\mathcal{A}$ is a subspace and therefore nonempty. Then, the result holds.

Case 2: $\mathcal{X} \neq \emptyset$. By Lemma 7.55, $\overline{\mathcal{A}}$ is an algebra. We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: $1 \in \overline{\mathcal{A}}$; Case 2b: $1 \notin \overline{\mathcal{A}}$. Case 2a: $1 \in \overline{\mathcal{A}}$. Clearly, $\overline{\mathcal{A}}$ is an algebra that separates points. By Stone-Weierstrass Theorem, $\mathcal{C}(\mathcal{X}, \mathbb{R}) = \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$. Then, the result holds.

Case 2b: $1 \notin \overline{\mathcal{A}}$. Then, $\overline{\mathcal{A}} \subset \mathcal{C}(\mathcal{X}, \mathbb{R})$. By Corollary 7.59, $\forall f \in \overline{\mathcal{A}}$, there exists $x \in \mathcal{X}$ such that $f(x) = 0$. Then, we have the following claims.

Claim 7.60.1 $\forall f \in \overline{\mathcal{A}}, |f| \in \overline{\mathcal{A}}$. Hence, $\overline{\mathcal{A}}$ is a lattice.

Proof of claim: $\forall f \in \overline{\mathcal{A}}$, by the compactness of $\mathcal{X}$ and Proposition 5.29, $\exists M \in (0, \infty) \subset \mathbb{R}$ such that $\|f\| \leq M$. Let $g := f/M \in \overline{\mathcal{A}}$. Then, $g : \mathcal{X} \rightarrow [-1, 1] \subset \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Lemma 7.54, there exists a polynomial $P : [-1, 1] \rightarrow \mathbb{R}$ with $P(0) = 0$ such that $|P(s) - |s|| < \epsilon/M$, $\forall s \in [-1, 1] \subset \mathbb{R}$. Since $\overline{\mathcal{A}}$ is an algebra, then $M \cdot P \circ g \in \overline{\mathcal{A}}$ and $\| |f| - M \cdot P \circ g \| = M \| |g| - P \circ g \| < \epsilon$. Hence, $|f| \in \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$.
 $\forall f, g \in \overline{\mathcal{A}}$, we have

$$\begin{aligned} f \vee g &= \frac{1}{2}(f + g) + \frac{1}{2}|f - g| \in \overline{\mathcal{A}} \\ f \wedge g &= \frac{1}{2}(f + g) - \frac{1}{2}|f - g| \in \overline{\mathcal{A}} \end{aligned}$$

Hence, $\overline{\mathcal{A}}$ is a lattice. This completes the proof of the claim. $\square$

Claim 7.60.2 $\exists x_0 \in \mathcal{X}$ such that $f(x_0) = 0, \forall f \in \overline{\mathcal{A}}$.

Proof of claim: We will prove this using an argument of contradiction. Suppose the claim is false. $\forall x \in \mathcal{X}, \exists f_x \in \overline{\mathcal{A}}$ such that $f_x(x) \neq 0$. By Claim 7.60.1, $|f_x| \in \overline{\mathcal{A}}$ and $|f_x(x)| > 0$. Let $O_x := \{\bar{x} \in \mathcal{X} \mid |f_x(\bar{x})| > 0\}$. Since $f_x \in \overline{\mathcal{A}} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$, then $x \in O_x \in \mathcal{O}_{\mathcal{X}}$. Hence, we have $\mathcal{X} \subseteq \bigcup_{x \in \mathcal{X}} O_x$. By the compactness of $\mathcal{X}$, there exists a finite set $X_N \subseteq \mathcal{X}$ such that $\mathcal{X} \subseteq \bigcup_{x \in X_N} O_x$. Clearly, $X_N \neq \emptyset$ since $\mathcal{X} \neq \emptyset$. Let $f := \sum_{x \in X_N} |f_x| \in \overline{\mathcal{A}}$. Hence, $f(x) > 0, \forall x \in \mathcal{X}$. By Corollary 7.59, we have $1 \in \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$, which is a contradiction. Hence, the claim is true. This completes the proof of the claim. $\square$

Claim 7.60.3 $\forall$ closed set $F \subseteq \mathcal{X}$ with $x_0 \notin F$, we have

(i) let $\mathcal{A}_F := \{h \in \mathcal{C}(F, \mathbb{R}) \mid \exists f \in \overline{\mathcal{A}} \text{ such that } h = f|_F\}$, then $\overline{\mathcal{A}_F} = \mathcal{C}(F, \mathbb{R})$;

(ii) $\forall \epsilon \in (0, 1) \subset \mathbb{R}, \exists g \in \overline{\mathcal{A}}$ such that $g : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R}$ and $g(x) \geq 1 - \epsilon, \forall x \in F$.

Proof of claim: Clearly, $\mathcal{A}_F$ is an algebra since $\overline{\mathcal{A}}$ is an algebra. Furthermore, $\mathcal{A}_F$ separate points on F since $\overline{\mathcal{A}}$ separates points on $\mathcal{X}$.

By Proposition 5.5, F with the subset topology is compact.

$\forall x \in F$, we have $x \neq x_0$. $\exists f_x \in \mathcal{A}$ such that $f_x(x) \neq f_x(x_0) = 0$. By Claim 7.60.1, $|f_x| \in \overline{\mathcal{A}}$ and $|f_x(x)| > 0$. Let $O_x := \{\bar{x} \in \mathcal{X} \mid |f_x(\bar{x})| > 0\}$. Then, we have $x \in O_x \in \mathcal{O}_{\mathcal{X}}$. Hence, $F \subseteq \bigcup_{x \in F} O_x$. By the compactness of F , there exists a finite set $F_N \subseteq F$ such that $F \subseteq \bigcup_{x \in F_N} O_x$. Let $f := \sum_{x \in F_N} |f_x|$. Then, $f \in \overline{\mathcal{A}}$ and $f(x) > 0, \forall x \in F$.

Then, $h := f|_F \in \mathcal{A}_F$ and $h(x) > 0, \forall x \in F$. By Corollary 7.59, we have $\overline{\mathcal{A}_F} = \mathcal{C}(F, \mathbb{R})$. Hence, (i) is true.

$\exists M \in (0, \infty) \subset \mathbb{R}$ such that $\|f\| \leq M$. Then, $g_0 := f/M \in \overline{\mathcal{A}}$ and $g_0 : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R}$.

By Corollary 7.57, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists a polynomial Q_ϵ in one variable such that $|\sqrt{s} - Q_\epsilon(s)| < \epsilon$, $\forall s \in [0, 1] \subset \mathbb{R}$, and $Q_\epsilon(0) = 0$.

$\forall \bar{f} : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R}$ with $\bar{f} \in \overline{\mathcal{A}}$, $Q_\epsilon \circ \bar{f} \in \overline{\mathcal{A}}$ since $\overline{\mathcal{A}}$ is an algebra and $Q_\epsilon(0) = 0$. Then, $\text{sqrt} \circ \bar{f} \in \overline{\mathcal{A}} = \overline{\mathcal{A}}$. Recursively, we may conclude that $\text{sqrt}^n \circ g_0 \in \overline{\mathcal{A}}$, $\forall n \in \mathbb{N}$.

By the compactness of F and Proposition 5.29, $\exists \gamma \in (0, 1] \subset \mathbb{R}$ such that $g_0(x) \geq \gamma$, $\forall x \in F$. $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\gamma^{2^{-n_0}} \geq 1 - \epsilon$. Then, $g := \text{sqrt}^{n_0} \circ g_0 \in \overline{\mathcal{A}}$, $g : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R}$ and $g(x) \geq 1 - \epsilon$, $\forall x \in F$. Hence, (ii) is true.

This completes the proof of the claim. $\square$

Claim 7.60.4 $\forall g \in \mathcal{C}(\mathcal{X}, \mathbb{R})$ with $g(x_0) = 0$ and $g(x) \geq 0$, $\forall x \in \mathcal{X}$, we have $g \in \overline{\mathcal{A}}$.

Proof of claim: $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $O := \{x \in \mathcal{X} \mid g(x) < \epsilon/2\}$. Then, $x_0 \in O \in \mathcal{O}_{\mathcal{X}}$ since g is continuous. Let $M := \|g\| \in [0, \infty) \subset \mathbb{R}$.

By Claim 7.60.3, $\exists h_1 \in \overline{\mathcal{A}}$ such that $h_1 : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R}$ and $h_1(x) \geq 2/3$, $\forall x \in \tilde{O}$. Note that $h_1(x_0) = 0$ by Claim 7.60.2. Let $h_2 := 3Mh_1/2 \in \overline{\mathcal{A}}$. Then, we have $0 \leq h_2(x) \leq 3M/2$, $\forall x \in \mathcal{X}$, and $h_2(x) \geq M \geq g(x)$, $\forall x \in \tilde{O}$.

Let $U := \{x \in O \mid h_2(x) < \epsilon/2\}$. Then, $x_0 \in U \in \mathcal{O}_{\mathcal{X}}$ since $O \in \mathcal{O}_{\mathcal{X}}$, h_2 is continuous, and $h_2(x_0) = 0$. Furthermore, $g|_{\tilde{U}} \in \mathcal{C}(\tilde{U}, \mathbb{R})$. By Claim 7.60.3, $\exists h_3 \in \overline{\mathcal{A}}$ such that $|h_3|_{\tilde{U}}(x) - g|_{\tilde{U}}(x)| < \epsilon$, $\forall x \in \tilde{U}$. Define $h_4 := (h_3 \vee 0) \wedge h_2$. Then, $h_4 \in \overline{\mathcal{A}}$ by Claim 7.60.1. $\forall x \in \mathcal{X}$, we will show that $|g(x) - h_4(x)| < \epsilon$ by distinguishing three exhaustive and mutually exclusive cases: Case A: $x \in U$; Case B: $x \in O \setminus U$; Case C: $x \in \tilde{O}$.

Case A: $x \in U$. Then, we have $0 \leq h_2(x) < \epsilon/2$, $0 \leq g(x) < \epsilon/2$. Then, $0 \leq h_4(x) = (h_3(x) \vee 0) \wedge h_2(x) \leq h_2(x) < \epsilon/2$. Therefore, $|g(x) - h_4(x)| < \epsilon$.

Case B: $x \in O \setminus U$. Then, we have $|h_3(x) - g(x)| < \epsilon$, $h_2(x) \geq 0$, and $0 \leq g(x) < \epsilon/2$. Hence, $0 \leq h_3(x) \vee 0 < g(x) + \epsilon$. Then, $0 \leq h_4(x) < g(x) + \epsilon$ and $-\epsilon/2 < -g(x) \leq h_4(x) - g(x) < \epsilon$. Therefore, we have $|g(x) - h_4(x)| < \epsilon$.

Case C: $x \in \tilde{O}$. Then, we have $|h_3(x) - g(x)| < \epsilon$ and $h_2(x) \geq M \geq g(x) \geq 0$. Hence, $g(x) - \epsilon < h_3(x) < g(x) + \epsilon$, $g(x) - \epsilon < h_3(x) \vee 0 < g(x) + \epsilon$, and $g(x) - \epsilon < h_4(x) < g(x) + \epsilon$. Therefore, $|g(x) - h_4(x)| < \epsilon$.

Hence, in all three cases, we have $|g(x) - h_4(x)| < \epsilon$. Hence $\|g - h_4\| < \epsilon$.

By the arbitrariness of ϵ , we have $g \in \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$. $\square$

$\forall f \in \mathcal{C}(\mathcal{X}, \mathbb{R})$ with $f(x_0) = 0$, let $f_+ := f \vee 0$ and $f_- := (-f) \vee 0$. Clearly, $f_+, f_- \in \mathcal{C}(\mathcal{X}, \mathbb{R})$, $f_+(x_0) = 0$, $f_-(x_0) = 0$, and $f = f_+ - f_-$. By Claim 7.60.4, $f_+, f_- \in \overline{\mathcal{A}}$. Then, $f \in \overline{\mathcal{A}}$ since $\overline{\mathcal{A}}$ is an algebra. This coupled with Claim 7.60.2, we have that $\overline{\mathcal{A}} = \{f \in \mathcal{C}(\mathcal{X}, \mathbb{R}) \mid f(x_0) = 0\}$. Hence, the result holds in this case.

This completes the proof of the proposition. $\square$

7.8 Linear Operators

Definition 7.61 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$. A linear operator $A : \mathcal{X} \rightarrow \mathcal{Y}$ is said to be bounded if $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|Ax\|_{\mathcal{Y}} \leq M \|x\|_{\mathcal{X}}, \forall x \in \mathcal{X}$.

Proposition 7.62 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$ and $A : \mathcal{X} \rightarrow \mathcal{Y}$ be a linear operator. Then,

(i) if A is bounded then A is uniformly continuous;

(ii) if A is continuous at some $x_0 \in \mathcal{X}$ then A is bounded.

Proof (i). $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|Ax\|_{\mathcal{Y}} \leq M \|x\|_{\mathcal{X}}, \forall x \in \mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta = \epsilon/(1 + M) \in (0, \infty) \subset \mathbb{R}$. $\forall x_1, x_2 \in \mathcal{X}$ with $\|x_1 - x_2\|_{\mathcal{X}} < \delta$, we have $\|Ax_1 - Ax_2\|_{\mathcal{Y}} = \|A(x_1 - x_2)\|_{\mathcal{Y}} \leq M \|x_1 - x_2\|_{\mathcal{X}} < \epsilon$. Hence A is uniformly continuous.

(ii). $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, we have $\|Ax - Ax_0\|_{\mathcal{Y}} < \epsilon$. $\forall \bar{x} \in \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, \delta)$, we have $\|A\bar{x}\|_{\mathcal{Y}} = \|A(x_0 + \bar{x}) - Ax_0\|_{\mathcal{Y}} < \epsilon$. Hence, A is continuous at $\vartheta_{\mathcal{X}}$. Then, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $\|Ax\|_{\mathcal{Y}} < 1, \forall x \in \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, \delta_0)$. Then, $\forall x \in \mathcal{X}$, we have two possibilities: (a) $x = \vartheta_{\mathcal{X}}$, then $\|Ax\|_{\mathcal{Y}} = 0 \leq \frac{2}{\delta_0} \|x\|_{\mathcal{X}}$; (b) $x \neq \vartheta_{\mathcal{X}}$, then, $\frac{\delta_0}{2\|x\|_{\mathcal{X}}}x \in \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, \delta_0)$ and we have $\|Ax\|_{\mathcal{Y}} = \left\| \frac{2\|x\|_{\mathcal{X}}}{\delta_0} A\left(\frac{\delta_0}{2\|x\|_{\mathcal{X}}}x\right) \right\|_{\mathcal{Y}} = \frac{2\|x\|_{\mathcal{X}}}{\delta_0} \left\| A\left(\frac{\delta_0}{2\|x\|_{\mathcal{X}}}x\right) \right\|_{\mathcal{Y}} < \frac{2}{\delta_0} \|x\|_{\mathcal{X}}$. Hence, A is bounded.

This completes the proof of the proposition. $\square$

Proposition 7.63 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$. Let $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})$ be the vector space defined in Example 6.20 with the null vector ϑ . Let $P := \{A \in \mathcal{M}(\mathcal{X}, \mathcal{Y}) \mid A \text{ is linear.}\}$. Then, P is a subspace of $\mathcal{M}(\mathcal{X}, \mathcal{Y})$. Define a functional $\|\cdot\|$ on P by $\|A\| := \inf \{M \in [0, \infty) \subset \mathbb{R} \mid \|Ax\|_{\mathcal{Y}} \leq M \|x\|_{\mathcal{X}}, \forall x \in \mathcal{X}\}, \forall A \in P$. Then, $\forall A \in P, \|A\|$ may be equivalently expressed as

$$\begin{aligned} \|A\| &= \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \|Ax\|_{\mathcal{Y}} \\ &= \max \left\{ \sup_{\substack{x \in \mathcal{X} \\ x \neq \vartheta_{\mathcal{X}}}} \frac{\|Ax\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}}, 0 \right\} = \max \left\{ \sup_{\substack{x \in \mathcal{X} \\ \|x\|_{\mathcal{X}} = 1}} \|Ax\|_{\mathcal{Y}}, 0 \right\} \end{aligned}$$

Let $N := \{A \in P \mid \|A\| < +\infty\}$. Then, $(N, \mathbb{K}, \|\cdot\|) =: \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is a normed linear space.

Proof It is easy to check that P is closed under vector addition and scalar multiplication in $\mathcal{M}(\mathcal{X}, \mathcal{Y})$. Clearly, $\vartheta \in P \neq \emptyset$. Hence, P is a subspace of $\mathcal{M}(\mathcal{X}, \mathcal{Y})$.

Next, we prove the four equivalent definitions of $\|A\|$. $\forall A \in P$, define the set $B_A := \{M \in [0, \infty) \subset \mathbb{R} \mid \|Ax\|_{\mathcal{Y}} \leq M\|x\|_{\mathcal{X}}, \forall x \in \mathcal{X}\}$. Let $\delta_{A1} := \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \|Ax\|_{\mathcal{Y}}$. Since $\|\vartheta_{\mathcal{X}}\|_{\mathcal{X}} = 0$ and $\|A\vartheta_{\mathcal{X}}\|_{\mathcal{Y}} = \|\vartheta_{\mathcal{Y}}\|_{\mathcal{Y}} = 0$, then $\delta_{A1} \geq 0$. $\forall M \in B_A$, we have $\|Ax\|_{\mathcal{Y}} \leq M\|x\|_{\mathcal{X}} \leq M$, $\forall x \in \mathcal{X}$ with $\|x\|_{\mathcal{X}} \leq 1$. Hence, we have $\delta_{A1} \leq M$ and $\delta_{A1} \leq \|A\|$.

Let $\delta_{A2} := \max \left\{ \sup_{x \in \mathcal{X}, x \neq \vartheta_{\mathcal{X}}} \frac{\|Ax\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}}, 0 \right\}$. $\forall x \in \mathcal{X}$ with $x \neq \vartheta_{\mathcal{X}}$, we have $\delta_{A1} \geq \left\| A \left(\frac{x}{\|x\|_{\mathcal{X}}} \right) \right\|_{\mathcal{Y}} = \frac{\|Ax\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}}$. Hence, we have $\delta_{A2} \leq \delta_{A1}$. Let $\delta_{A3} := \max \left\{ \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}}=1} \|Ax\|_{\mathcal{Y}}, 0 \right\}$. Clearly, we have $0 \leq \delta_{A3} \leq \delta_{A2}$.

On the other hand, $\forall x \in \mathcal{X}$, we have either $x = \vartheta_{\mathcal{X}}$, then $\|Ax\|_{\mathcal{Y}} = 0 \leq \delta_{A3}\|x\|_{\mathcal{X}}$; or $x \neq \vartheta_{\mathcal{X}}$, then $\|x\|_{\mathcal{X}} > 0$ and $\delta_{A3} \geq \left\| A \left(\frac{x}{\|x\|_{\mathcal{X}}} \right) \right\|_{\mathcal{Y}} = \frac{\|Ax\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}}$, and hence, $\|Ax\|_{\mathcal{Y}} \leq \delta_{A3}\|x\|_{\mathcal{X}}$. Thus, $\forall x \in \mathcal{X}$, we have $\|Ax\|_{\mathcal{Y}} \leq \delta_{A3}\|x\|_{\mathcal{X}}$. If $\delta_{A3} = +\infty$, then $\|A\| \leq \delta_{A3}$. If $\delta_{A3} < +\infty$, then $\delta_{A3} \in B_A$ and $\|A\| \leq \delta_{A3}$. Therefore, $\|A\| \leq \delta_{A3}$.

Hence, $\|A\| = \delta_{A1} = \delta_{A2} = \delta_{A3}$.

It is easy to check that N is closed under vector addition and scalar multiplication in P . Clearly, $\vartheta \in N \neq \emptyset$. Hence, N is a subspace of P .

Finally, we will show that $\|\cdot\|$ defines a norm on N . $\forall A_1, A_2 \in N$, $\forall \alpha \in \mathbb{K}$. (i) $\|A_1\| \in [0, \infty) \subset \mathbb{R}$. If $A_1 = \vartheta$, then $\|A_1\| = \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \|A_1x\|_{\mathcal{Y}} = \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} 0 = 0$. On the other hand, if $\|A_1\| = 0$, then, $\forall x \in \mathcal{X}$, $0 \leq \|A_1x\|_{\mathcal{Y}} \leq \|A_1\|\|x\|_{\mathcal{X}} = 0$, which implies that $A_1x = \vartheta_{\mathcal{Y}}$ and $A_1 = \vartheta$. Hence, $\|A_1\| = 0 \Leftrightarrow A_1 = \vartheta$.

(ii) $\|A_1 + A_2\| = \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \|(A_1 + A_2)x\|_{\mathcal{Y}} = \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \|A_1x\|_{\mathcal{Y}} + \|A_2x\|_{\mathcal{Y}} \leq \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} (\|A_1x\|_{\mathcal{Y}} + \|A_2x\|_{\mathcal{Y}}) \leq \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \|A_1x\|_{\mathcal{Y}} + \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \|A_2x\|_{\mathcal{Y}} = \|A_1\| + \|A_2\|$.

(iii) $\|\alpha A_1\| = \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \|(\alpha A_1)x\|_{\mathcal{Y}} = \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \|\alpha(A_1x)\|_{\mathcal{Y}} = \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} |\alpha| \|A_1x\|_{\mathcal{Y}} = |\alpha| \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} \|A_1x\|_{\mathcal{Y}} = |\alpha| \|A_1\|$, where we have made use of Proposition 3.81 in the fourth equality.

Hence, $B(\mathcal{X}, \mathcal{Y})$ is a normed linear space. $\square$

Proposition 7.64 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over the field $\mathbb{K}$, $A \in B(\mathcal{X}, \mathcal{Y})$, and $B \in B(\mathcal{Y}, \mathcal{Z})$. Then, $\forall x \in \mathcal{X}$, we have $\|Ax\| \leq \|A\|\|x\|$, where the three norms are over three different normed linear spaces. Furthermore, $\|BA\| \leq \|B\|\|A\|$, where the three norms are over three different normed linear spaces.*

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 7.65 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$ and $B(\mathcal{X}, \mathcal{Y})$ be the normed linear space of bounded linear operators of $\mathcal{X}$*

to $\mathcal{Y}$. Define an operator $\psi : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \rightarrow \mathcal{Y}$ by $\psi(A, x) = Ax$, $\forall A \in B(\mathcal{X}, \mathcal{Y})$, $\forall x \in \mathcal{X}$. Then, ψ is continuous.

Proof $\forall (A_0, x_0) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta = \frac{\epsilon}{\epsilon + \|x_0\|} \wedge \frac{\epsilon}{\epsilon + \|A_0\|} \wedge \epsilon \in (0, \infty) \subset \mathbb{R}$. $\forall (A, x) \in \mathcal{B}_{B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}}((A_0, x_0), \delta)$, we have $\|\psi(A, x) - \psi(A_0, x_0)\| = \|Ax - A_0x_0\| = \|Ax - Ax_0 + Ax_0 - A_0x_0\| \leq \|A(x - x_0)\| + \|(A - A_0)x_0\| \leq \|A\|\|x - x_0\| + \|A - A_0\|\|x_0\| \leq (\|A_0\| + \|A - A_0\|)\delta + \|x_0\|\delta \leq (\delta + \|A_0\|)\delta + \|x_0\|\delta \leq \epsilon + \epsilon = 2\epsilon$. Hence, ψ is continuous. This completes the proof of the proposition. $\square$

Proposition 7.66 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$. Let $B(\mathcal{X}, \mathcal{Y})$ be the normed linear space defined in Proposition 7.63. If $\mathcal{Y}$ is a Banach space, then $B(\mathcal{X}, \mathcal{Y})$ is also a Banach space.

Proof All we need to show is that $B(\mathcal{X}, \mathcal{Y})$ is complete. Take a Cauchy sequence $(A_n)_{n=1}^\infty \subseteq B(\mathcal{X}, \mathcal{Y})$. $\forall x \in \mathcal{X}$, we have $\|A_nx - A_mx\| \leq \|A_n - A_m\|\|x\|$, $\forall n, m \in \mathbb{N}$, by Proposition 7.64. Hence, $(A_nx)_{n=1}^\infty \subseteq \mathcal{Y}$ is a Cauchy sequence. Since $\mathcal{Y}$ is a Banach space, then $\exists! y_x \in \mathcal{Y}$ such that $\lim_{n \in \mathbb{N}} A_nx = y_x$. Hence, we may define a function $f : \mathcal{X} \rightarrow \mathcal{Y}$ by $f(x) = y_x$, $\forall x \in \mathcal{X}$.

$\forall x_1, x_2 \in \mathcal{X}$, $\forall \alpha, \beta \in \mathbb{K}$, by Propositions 7.23, 4.15, and 3.67, we have $f(\alpha x_1 + \beta x_2) = \lim_{n \in \mathbb{N}} A_n(\alpha x_1 + \beta x_2) = \lim_{n \in \mathbb{N}} (\alpha A_nx_1 + \beta A_nx_2) = \alpha \lim_{n \in \mathbb{N}} A_nx_1 + \beta \lim_{n \in \mathbb{N}} A_nx_2 = \alpha f(x_1) + \beta f(x_2)$. Hence, f is linear.

Since, $(A_n)_{n=1}^\infty$ is Cauchy, then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$, $\forall n, m \geq N$, $\|A_n - A_m\| < \epsilon$. $\forall x \in \mathcal{X}$, $\|A_nx - A_mx\| \leq \|A_n - A_m\|\|x\| \leq \epsilon\|x\|$. Then, by Propositions 3.66, 7.21, and 7.23, $\|f(x) - A_mx\| = \lim_{n \in \mathbb{N}} \|A_nx - A_mx\| \leq \lim_{n \in \mathbb{N}} \|A_n - A_m\|\|x\| \leq \epsilon\|x\|$. Therefore, $\|f(x)\| \leq \|f(x) - A_Nx\| + \|A_Nx\| \leq \epsilon\|x\| + \|A_N\|\|x\| = (\epsilon + \|A_N\|)\|x\|$. This implies that f is bounded. Hence, $f \in B(\mathcal{X}, \mathcal{Y})$. Note that $0 \leq \lim_{n \in \mathbb{N}} \|A_n - f\| = \lim_{n \in \mathbb{N}} \sup_{x \in \mathcal{X}, \|x\| \leq 1} \|(A_n - f)(x)\| = \lim_{n \in \mathbb{N}} \sup_{x \in \mathcal{X}, \|x\| \leq 1} \|A_nx - f(x)\| \leq \epsilon$. By the arbitrariness of ϵ , we have $\lim_{n \in \mathbb{N}} \|A_n - f\| = 0$ and hence $\lim_{n \in \mathbb{N}} A_n = f \in B(\mathcal{X}, \mathcal{Y})$. Hence, $B(\mathcal{X}, \mathcal{Y})$ is complete.

This completes the proof of the proposition. $\square$

Proposition 7.67 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$, $\mathcal{X}$ be finite dimensional with dimension $n \in \mathbb{Z}_+$, and $A : \mathcal{X} \rightarrow \mathcal{Y}$ be a linear operator. Then, $A \in B(\mathcal{X}, \mathcal{Y})$.

Proof Let $X_N \subseteq \mathcal{X}$ be a basis for $\mathcal{X}$, which then contains exactly n vectors. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $n = 0$; Case 2: $n \in \mathbb{N}$. Case 1: $n = 0$. Then, $\mathcal{X} = \{0\}$. Then, by Proposition 7.63, we have $\|A\| = \sup_{x \in \mathcal{X}, \|x\| \leq 1} \|Ax\| = 0$. Hence, $A \in B(\mathcal{X}, \mathcal{Y})$.

Case 2: $n \in \mathbb{N}$. Let $X_N = \{x_1, \dots, x_n\}$. $\forall x \in \mathcal{X}$, by Corollary 6.47 and Definition 6.50, $\exists! \alpha_1, \dots, \alpha_n \in \mathbb{K}$ such that $x = \sum_{i=1}^n \alpha_i x_i$. Then, we may define an alternative norm $\|\cdot\|_1$ on $\mathcal{X}$ by $\|x\|_1 = \sqrt{\sum_{i=1}^n |\alpha_i|^2}$.

It is easy to check that $\|\cdot\|_1$ defines a norm on $\mathcal{X}$. By Theorem 7.38, $\exists M \in (0, \infty) \subset \mathbb{R}$ such that $\|x\|/M \leq \|x\|_1 \leq M\|x\|$, $\forall x \in \mathcal{X}$. Define $r_i := \|Ax_i\|$, $i = 1, \dots, n$. Then, by Proposition 7.63,

$$\begin{aligned}
\|A\| &= \sup_{x \in \mathcal{X}, \|x\| \leq 1} \|Ax\| = \sup_{\alpha_1, \dots, \alpha_n \in \mathbb{R}, \|\sum_{i=1}^n \alpha_i x_i\| \leq 1} \left\| A \left(\sum_{i=1}^n \alpha_i x_i \right) \right\| \\
&\leq \sup_{\alpha_1, \dots, \alpha_n \in \mathbb{R}, \|\sum_{i=1}^n \alpha_i x_i\| \leq 1} \sum_{i=1}^n |\alpha_i| r_i \\
&\leq \sup_{\alpha_1, \dots, \alpha_n \in \mathbb{R}, \|\sum_{i=1}^n \alpha_i x_i\| \leq 1} \left(\sum_{i=1}^n r_i^2 \right)^{1/2} \left(\sum_{i=1}^n |\alpha_i|^2 \right)^{1/2} \\
&= \sup_{x \in \mathcal{X}, \|x\| \leq 1} \left(\sum_{i=1}^n r_i^2 \right)^{1/2} \|x\|_1 \leq \sup_{x \in \mathcal{X}, \|x\| \leq 1} \left(\sum_{i=1}^n r_i^2 \right)^{1/2} M \|x\| \\
&\leq M \left(\sum_{i=1}^n r_i^2 \right)^{1/2}
\end{aligned}$$

where we have applied the Cauchy-Schwarz Inequality in the second inequality. Hence, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.

This completes the proof of the proposition. $\square$

Proposition 7.68 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$ and $A : \mathcal{X} \rightarrow \mathcal{Y}$ be a linear operator. If A is bounded, then $\mathcal{N}(A) \subseteq \mathcal{X}$ is closed. On the other hand, if $\mathcal{N}(A)$ is closed and $\mathcal{R}(A) \subseteq \mathcal{Y}$ is finite dimensional, then A is bounded.*

Proof Let A be bounded. By Proposition 7.62, A is continuous. Note that the set $\{\vartheta_{\mathcal{Y}}\} \subseteq \mathcal{Y}$ is closed by Proposition 3.34. By Proposition 3.10, $\mathcal{N}(A) = A_{\text{inv}}(\{\vartheta_{\mathcal{Y}}\})$ is closed.

Let $\mathcal{N}(A)$ be closed and $\mathcal{R}(A)$ be finite dimensional. Let $n \in \mathbb{Z}_+$ be the dimension of $\mathcal{R}(A)$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $n = 0$; Case 2: $n \in \mathbb{N}$. Case 1: $n = 0$. Then, $\mathcal{R}(A) = \{\vartheta_{\mathcal{Y}}\}$. Hence, $A = \vartheta_{\mathcal{B}(\mathcal{X}, \mathcal{Y})}$, which is clearly bounded.

Case 2: $n \in \mathbb{N}$. Then, $\exists X_N = \{x_1, \dots, x_n\} \subseteq \mathcal{X}$ with exactly n elements such that $A(X_N) \subseteq \mathcal{Y}$ is a basis of $\mathcal{R}(A)$. Let $\mathcal{X}/\mathcal{N}(A)$ be the quotient normed linear space as defined in Proposition 7.44. $\forall x \in \mathcal{X}$, $Ax \in \mathcal{R}(A) = \text{span}(A(X_N))$. Then, by Definition 6.50 and Corollary 6.47, $\exists! \alpha_1, \dots, \alpha_n \in \mathbb{K}$ such that $Ax = \sum_{i=1}^n \alpha_i Ax_i$. Then, we have $A(x - \sum_{i=1}^n \alpha_i x_i) = \vartheta_{\mathcal{Y}}$ and $x - \sum_{i=1}^n \alpha_i x_i \in \mathcal{N}(A)$. Hence, $[x] = [\sum_{i=1}^n \alpha_i x_i] = \sum_{i=1}^n \alpha_i [x_i]$. This implies that $\mathcal{X}/\mathcal{N}(A) \subseteq \text{span}(\{[x_1], \dots, [x_n]\})$. Hence, $\mathcal{X}/\mathcal{N}(A)$ is finite dimensional.

Define a mapping $\bar{A} : \mathcal{X}/\mathcal{N}(A) \rightarrow \mathcal{Y}$ by $\bar{A}[x] = Ax$, $\forall [x] \in \mathcal{X}/\mathcal{N}(A)$. Note that $\forall \bar{x} \in [x]$, we have $x - \bar{x} \in \mathcal{N}(A)$ and $Ax = A\bar{x}$. Hence, $\bar{A}$ is uniquely defined. $\forall [x_1], [x_2] \in \mathcal{X}/\mathcal{N}(A)$, $\forall \alpha, \beta \in \mathbb{K}$, we have $\bar{A}(\alpha[x_1] +$

$\beta[x_2]) = \bar{A}([\alpha x_1 + \beta x_2]) = A(\alpha x_1 + \beta x_2) = \alpha Ax_1 + \beta Ax_2 = \alpha \bar{A}([x_1]) + \beta \bar{A}([x_2])$. Hence, $\bar{A}$ is a linear operator. By Proposition 7.67, $\bar{A}$ is bounded. $\forall x \in \mathcal{X}$, $\|Ax\| = \|\bar{A}[x]\| \leq \|\bar{A}\| \|x\| \leq \|\bar{A}\| \|x\|$. Hence, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.

This completes the proof of the proposition. $\square$

Proposition 7.69 *Let $\mathcal{X}$ be a normed linear space and $M \subseteq \mathcal{X}$ be a closed subspace. Let $\mathcal{X}/M$ be the quotient normed linear space and $\phi : \mathcal{X} \rightarrow \mathcal{X}/M$ be the natural homomorphism. Then, $\|\phi\| \leq 1$.*

Proof By Proposition 7.43, ϕ is linear. $\forall x \in \mathcal{X}$, $\|\phi(x)\| = \inf_{m \in M} \|x - m\| \leq \|x\|$. Then, $\phi \in \mathcal{B}(\mathcal{X}, \mathcal{X}/M)$ and $\|\phi\| \leq 1$. This completes the proof of the proposition. $\square$

Proposition 7.70 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, and $\phi : \mathcal{X} \rightarrow \mathcal{X}/\mathcal{N}(A)$ be the natural homomorphism. Then, $\exists B \in \mathcal{B}(\mathcal{X}/\mathcal{N}(A), \mathcal{Y})$ such that $A = B \circ \phi$, B is injective, and $\|A\| = \|B\|$.*

Proof By Proposition 7.68, $\mathcal{N}(A)$ is closed. Then, by Proposition 7.44, $\mathcal{X}/\mathcal{N}(A)$ is a normed linear space. Define $B : \mathcal{X}/\mathcal{N}(A) \rightarrow \mathcal{Y}$ by $B([x]) = Ax$, $\forall [x] \in \mathcal{X}/\mathcal{N}(A)$. Note that $\forall \bar{x} \in [x]$, we have $x - \bar{x} \in \mathcal{N}(A)$ and $Ax = A\bar{x}$. Hence, B is uniquely defined. $\forall [x_1], [x_2] \in \mathcal{X}/\mathcal{N}(A)$, $\forall \alpha, \beta \in \mathbb{K}$, we have $B(\alpha[x_1] + \beta[x_2]) = B([\alpha x_1 + \beta x_2]) = A(\alpha x_1 + \beta x_2) = \alpha Ax_1 + \beta Ax_2 = \alpha B([x_1]) + \beta B([x_2])$. Hence, B is linear. $\forall [x_1], [x_2] \in \mathcal{X}/\mathcal{N}(A)$, $B[x_1] = B[x_2]$ implies that $Ax_1 = Ax_2$, $x_1 - x_2 \in \mathcal{N}(A)$, and hence, $[x_1] = [x_2]$. This shows that B is injective. $\forall x \in \mathcal{X}$, $B \circ \phi(x) = B[x] = Ax$. Hence, $A = B \circ \phi$.

$\forall [x] \in \mathcal{X}/\mathcal{N}(A)$, $\forall m \in \mathcal{N}(A)$, $\|B[x]\| = \|Ax\| = \|A(x - m)\| \leq \|A\| \|x - m\|$. This implies that $\|B[x]\| \leq \|A\| \|x\|$. Hence, $B \in \mathcal{B}(\mathcal{X}/\mathcal{N}(A), \mathcal{Y})$ and $\|B\| \leq \|A\|$. By Propositions 7.64 and 7.69, $\|A\| = \|B \circ \phi\| \leq \|B\| \|\phi\| \leq \|B\|$. Then, $\|A\| = \|B\|$. This completes the proof of the proposition. $\square$

7.9 Dual Spaces

7.9.1 Basic concepts

Definition 7.71 *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$. Then, the space $\mathcal{B}(\mathcal{X}, \mathbb{K})$, which consists of all bounded linear functional on $\mathcal{X}$, is called the dual of $\mathcal{X}$ and denoted by $\mathcal{X}^*$. We will denote the vectors in $\mathcal{X}^*$ by x_* and denote $x_*(x)$ by $\langle x_*, x \rangle$.*

Proposition 7.72 *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$ and $\mathcal{X}^*$ be its dual. Then, the following statements hold.*

- (i) *If a linear functional $f : \mathcal{X} \rightarrow \mathbb{K}$ is continuous at some $x_0 \in \mathcal{X}$, then $f \in \mathcal{X}^*$.*

- (ii) $\forall x_* \in \mathcal{X}^*$, x_* is uniformly continuous.
- (iii) For any linear functional $f : \mathcal{X} \rightarrow \mathbb{K}$, we have $\|f\| = \inf \{ M \in [0, \infty) \subset \mathbb{R} \mid |f(x)| \leq M \|x\|_{\mathcal{X}}, \forall x \in \mathcal{X} \} = \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1} |f(x)| = \max \left\{ \sup_{x \in \mathcal{X}, x \neq \vartheta_{\mathcal{X}}} \frac{|f(x)|}{\|x\|_{\mathcal{X}}}, 0 \right\} = \max \left\{ \sup_{x \in \mathcal{X}, \|x\|_{\mathcal{X}}=1} |f(x)|, 0 \right\}.$
- (iv) $|\langle x_*, x \rangle| \leq \|x_*\| \|x\|, \forall x \in \mathcal{X}, \forall x_* \in \mathcal{X}^*.$
- (v) $\langle \cdot, \cdot \rangle$ is a continuous function on $\mathcal{X}^* \times \mathcal{X}.$
- (vi) $\mathcal{X}^*$ is a Banach space.
- (vii) A linear functional $f : \mathcal{X} \rightarrow \mathbb{K}$ is bounded if, and only if, $\mathcal{N}(f)$ is closed.

Proof (i) and (ii) are direct consequences of Proposition 7.62. (iii) is a direct consequence of Proposition 7.63. (iv) follows from Proposition 7.64. (v) follows from Proposition 7.65. (vi) follows from Proposition 7.66 since $\mathbb{K}$ is a Banach space. Finally, (vii) follows from Proposition 7.68. This completes the proof of the proposition. $\square$

7.9.2 Duals of some common Banach spaces

Example 7.73 Let us consider the dual of $\mathcal{X} := (\{\vartheta_{\mathcal{X}}\}, \mathbb{K}, \|\cdot\|)$. Clearly, there is a single linear functional on $\mathcal{X}$ given by $f : \mathcal{X} \rightarrow \mathbb{K}$ and $f(\vartheta_{\mathcal{X}}) = 0$. Clearly, this linear functional is bounded with norm 0. Hence, $\mathcal{X}^* = (\{f\}, \mathbb{K}, \|\cdot\|_1)$, which is isometrically isomorphic to $\mathcal{X}$. $\diamond$

Example 7.74 Let us consider the dual of $\mathcal{X} := \mathbb{K}^n, n \in \mathbb{N}$. $\forall x := (\xi_1, \dots, \xi_n) \in \mathbb{K}^n$, any functional of the form $f(x) = \sum_{i=1}^n \eta_i \xi_i$ with $\eta_1, \dots, \eta_n \in \mathbb{K}$ is clearly linear. By Cauchy-Schwarz Inequality, we have $|f(x)| = |\sum_{i=1}^n \eta_i \xi_i| \leq \sum_{i=1}^n |\eta_i| |\xi_i| \leq \sqrt{\sum_{i=1}^n |\eta_i|^2} \sqrt{\sum_{i=1}^n |\xi_i|^2} = \sqrt{\sum_{i=1}^n |\eta_i|^2} |x|$. Hence, f is bounded and $\|f\| \leq \sqrt{\sum_{i=1}^n |\eta_i|^2}$. Since the equality is achieved at $x = (\overline{\eta_1}, \dots, \overline{\eta_n})$ in the above inequality, where $\overline{\eta_i}$ denotes the complex conjugate of η_i , then $\|f\| = \sqrt{\sum_{i=1}^n |\eta_i|^2}$. Clearly, for different n -tuple $(\eta_1, \dots, \eta_n)$, the linear functional f is distinct. Now, let f be a bounded linear functional on $\mathcal{X}$. Let $e_i \in \mathcal{X}$ be the i th unit vector (all components of e_i are zero except a 1 at the i th component), $i = 1, \dots, n$. Let $\eta_i = f(e_i) \in \mathbb{K}, i = 1, \dots, n$. Then, $\forall x = (\xi_1, \dots, \xi_n) \in \mathcal{X}$, we have $x = \sum_{i=1}^n \xi_i e_i$ and $f(x) = \sum_{i=1}^n \eta_i \xi_i$. Define a function $\Psi : \mathcal{X}^* \rightarrow \mathbb{K}^n$ by $\Psi(f) = (\eta_1, \dots, \eta_n), \forall f \in \mathcal{X}^*$. Clearly, Ψ is linear, bijective and norm preserving. Hence, the dual of $\mathbb{K}^n$ is (isometrically isomorphic to) $\mathbb{K}^n$. $\diamond$

Lemma 7.75 *Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $\mathcal{X}^*$ be its dual, and $x_* \in \mathcal{X}^*$. Then, $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\exists x \in \mathcal{X}$ with $\|x\| \leq 1$ such that $\langle x_*, x \rangle \in \mathbb{R}$ and $\langle x_*, x \rangle \geq (1 - \epsilon) \|x_*\|$.*

Proof $\forall \epsilon \in (0, 1) \subset \mathbb{R}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\|x_*\| = 0$; Case 2: $\|x_*\| > 0$. Case 1: $\|x_*\| = 0$. Then, $x_* = \vartheta_*$. Take $x = \vartheta$ and the result holds. Case 2: $\|x_*\| > 0$. Note that $\|x_*\| = \sup_{x \in \mathcal{X}, \|x\| \leq 1} |\langle x_*, x \rangle|$. Then, $\exists \bar{x} \in \mathcal{X}$ with $\|\bar{x}\| \leq 1$ such that $|\langle x_*, \bar{x} \rangle| \geq (1 - \epsilon) \|x_*\| > 0$. Take $x := \frac{|\langle x_*, \bar{x} \rangle|}{\langle x_*, \bar{x} \rangle} \bar{x} \in \mathcal{X}$. Then, $\|x\| = \|\bar{x}\| \leq 1$ and $\langle x_*, x \rangle = \left\langle x_*, \frac{|\langle x_*, \bar{x} \rangle|}{\langle x_*, \bar{x} \rangle} \bar{x} \right\rangle = \frac{|\langle x_*, \bar{x} \rangle|}{\langle x_*, \bar{x} \rangle} \langle x_*, \bar{x} \rangle = |\langle x_*, \bar{x} \rangle| \in \mathbb{R}$ and $\langle x_*, x \rangle \geq (1 - \epsilon) \|x_*\|$. This completes the proof of the lemma. $\square$

Example 7.76 Let $\mathcal{Y}_i$ be a normed linear space over the field $\mathbb{K}$, $\mathcal{Y}_i^*$ be its dual, $i = 1, \dots, n$, $n \in \mathbb{N}$. We will study the dual of $\mathcal{X} := \prod_{i=1}^n \mathcal{Y}_i$. We will show that $\mathcal{X}^* = \prod_{i=1}^n \mathcal{Y}_i^*$ isometrically isomorphically.

Let $f \in \mathcal{X}^*$. $\forall i = 1, \dots, n$, define $A_i : \mathcal{Y}_i \rightarrow \mathcal{X}$ by $x = (y_1, \dots, y_n) = A_i y$, $\forall y \in \mathcal{Y}_i$, and $y_j = \vartheta_{\mathcal{Y}_j}$, $\forall j = 1, \dots, n$ with $j \neq i$, and $y_i = y$. Clearly, A_i is well-defined and linear and bounded with $\|A_i\| \leq 1$. Define $f_i : \mathcal{Y}_i \rightarrow \mathbb{K}$ by $f_i = f \circ A_i$. Then, f_i is a bounded linear functional on $\mathcal{Y}_i$ with $\|f_i\| \leq \|f\| \|A_i\| \leq \|f\|$. Hence, $f_i \in \mathcal{Y}_i^*$. Denote $f_i =: y_{*i}$. $\forall x = (y_1, \dots, y_n) \in \mathcal{X}$, we have $y_i \in \mathcal{Y}_i$, $i = 1, \dots, n$. By the linearity of f , we have

$$f(x) = f\left(\sum_{i=1}^n A_i y_i\right) = \sum_{i=1}^n f(A_i y_i) = \sum_{i=1}^n \langle y_{*i}, y_i \rangle \quad (7.1)$$

Let $x_* := (y_{*1}, \dots, y_{*n}) \in \prod_{i=1}^n \mathcal{Y}_i^*$. $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\forall i = 1, \dots, n$, by Lemma 7.75, $\exists \bar{y}_i \in \mathcal{Y}_i$ with $\|\bar{y}_i\| \leq 1$ such that $\langle y_{*i}, \bar{y}_i \rangle \in \mathbb{R}$ and $\langle y_{*i}, \bar{y}_i \rangle \geq (1 - \epsilon) \|y_{*i}\|$. Let $\hat{y}_i = \|y_{*i}\| \bar{y}_i$. Then, $\langle y_{*i}, \hat{y}_i \rangle \geq (1 - \epsilon) \|y_{*i}\|^2$. Then, we have $|f(\sum_{i=1}^n A_i \hat{y}_i)| = |\sum_{i=1}^n \langle y_{*i}, \hat{y}_i \rangle| = \sum_{i=1}^n \langle y_{*i}, \hat{y}_i \rangle \geq (1 - \epsilon) \sum_{i=1}^n \|y_{*i}\|^2 = (1 - \epsilon) \|x_*\|^2$. Since $f \in \mathcal{X}^*$, then, by Proposition 7.72, we have $(1 - \epsilon) \|x_*\|^2 \leq \|f\| \|\sum_{i=1}^n A_i \hat{y}_i\| = \|f\| \left(\sum_{i=1}^n \|\hat{y}_i\|^2\right)^{1/2} \leq \|f\| \left(\sum_{i=1}^n \|y_{*i}\|^2\right)^{1/2} = \|f\| \|x_*\|$. By the arbitrariness of ϵ , we have $\|x_*\| \leq \|f\|$.

Based on the preceding analysis, we may define a function $\psi : \mathcal{X}^* \rightarrow \prod_{i=1}^n \mathcal{Y}_i^*$ by $\psi(f) = (f \circ A_1, \dots, f \circ A_n)$, $\forall f \in \mathcal{X}^*$. Then, $\|\psi(f)\| \leq \|f\|$. Clearly, ψ is linear. $\forall f_1, f_2 \in \mathcal{X}^*$ with $\psi(f_1) = \psi(f_2)$. Then, we have $f_1 \circ A_i = f_2 \circ A_i$, $\forall i \in \mathbb{N}$. $\forall x := (y_1, \dots, y_n) \in \mathcal{X}$, by (7.1), we have $f_1(x) = \sum_{i=1}^n f_1(A_i y_i) = \sum_{i=1}^n f_2(A_i y_i) = f_2(x)$. Then, we have $f_1 = f_2$. Hence, ψ is injective. $\forall x_* := (y_{*1}, \dots, y_{*n}) \in \prod_{i=1}^n \mathcal{Y}_i^*$, define $\phi(x_*) : \mathcal{X} \rightarrow \mathbb{K}$ by $\phi(x_*)(x) = \sum_{i=1}^n \langle y_{*i}, y_i \rangle$, $\forall x := (y_1, \dots, y_n) \in \mathcal{X}$. Note that, $\forall x := (y_1, \dots, y_n) \in \mathcal{X}$, $|\sum_{i=1}^n \langle y_{*i}, y_i \rangle| \leq \sum_{i=1}^n \|y_{*i}\| \|y_i\| \leq \|x_*\| \|x\|$, where we have made use of Proposition 7.72 in the first inequality and

Cauchy-Schwarz Inequality in the second inequality. Clearly, $\phi(x_*)$ is linear and bounded with

$$\begin{aligned} \|\phi(x_*)\| &= \sup_{x \in \mathcal{X}, \|x\| \leq 1} |\phi(x_*)(x)| \leq \sup_{x=(y_1, \dots, y_n) \in \mathcal{X}, \|x\| \leq 1} \sum_{i=1}^n |\langle y_{*i}, y_i \rangle| \\ &\leq \sup_{x \in \mathcal{X}, \|x\| \leq 1} \|x_*\| \|x\| \leq \|x_*\| \end{aligned} \quad (7.2)$$

Hence, $\phi(x_*) \in \mathcal{X}^*$ and $\psi(\phi(x_*)) = (\phi(x_*) \circ A_1, \dots, \phi(x_*) \circ A_n) = (y_{*1}, \dots, y_{*n}) = x_*$. Then, $\psi \circ \phi = \text{id}_{\prod_{i=1}^n \mathcal{Y}_i^*}$. Hence, ψ is surjective. Therefore, ψ is bijective and admits inverse ψ_{inv} . By Proposition 2.4, $\phi = \psi_{\text{inv}}$. $\forall f \in \mathcal{X}^*$, by (7.2), we have $\|f\| = \|\phi(\psi(f))\| \leq \|\psi(f)\| \leq \|f\|$. Then, $\|\psi(f)\| = \|f\|$ and ψ is an isometry. Hence, ψ is an isometrical isomorphism. $\diamond$

Example 7.77 Let us consider the dual of $\mathcal{X} := l_p(\mathcal{Y})$, $p \in [1, \infty) \subset \mathbb{R}$, $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$. Let $\mathcal{Y}^*$ be the dual of $\mathcal{Y}$ and $q \in (1, \infty] \subset \mathbb{R}_e$ with $1/p + 1/q = 1$. We will show that $\mathcal{X}^*$ is isometrically isomorphic to $l_q(\mathcal{Y}^*)$.

Let $f \in \mathcal{X}^*$. $\forall i \in \mathbb{N}$, define $A_i : \mathcal{Y} \rightarrow \mathcal{X}$ by $x = (y_1, y_2, \dots) = A_i y$, $\forall y \in \mathcal{Y}$, and $y_j = \vartheta y$, $\forall j \in \mathbb{N}$ with $j \neq i$, and $y_i = y$. Clearly, A_i is well-defined, linear, and bounded with $\|A_i\| \leq 1$. Define $f_i : \mathcal{Y} \rightarrow \mathbb{K}$ by $f_i = f \circ A_i$. Then, f_i is a bounded linear functional on $\mathcal{Y}$ with $\|f_i\| \leq \|f\| \|A_i\| \leq \|f\|$. Hence, $f_i \in \mathcal{Y}^*$. Denote $f_i =: y_{*i}$. $\forall x = (y_1, y_2, \dots) \in \mathcal{X}$, we have $y_i \in \mathcal{Y}$, $\forall i \in \mathbb{N}$, and $\sum_{i=1}^{\infty} \|y_i\|^p < +\infty$. Then, $\lim_{n \in \mathbb{N}} \sum_{i=n+1}^{\infty} \|y_i\|^p = 0$. This implies that $\lim_{n \in \mathbb{N}} \|x - \sum_{i=1}^n A_i y_i\| = \lim_{n \in \mathbb{N}} (\sum_{i=n+1}^{\infty} \|y_i\|^p)^{1/p} = 0$. Hence, $\lim_{n \in \mathbb{N}} \sum_{i=1}^n A_i y_i = x$. By the continuity of f and Proposition 3.66, we have

$$f(x) = \lim_{n \in \mathbb{N}} f\left(\sum_{i=1}^n A_i y_i\right) = \lim_{n \in \mathbb{N}} \sum_{i=1}^n f(A_i y_i) = \lim_{n \in \mathbb{N}} \sum_{i=1}^n \langle y_{*i}, y_i \rangle \quad (7.3)$$

Claim 7.77.1 $x_* := (y_{*1}, y_{*2}, \dots) \in l_q(\mathcal{Y}^*)$ and $\|x_*\| \leq \|f\|$.

Proof of claim: We will distinguish two exhaustive and mutually exclusive cases: Case 1: $1 < p < +\infty$; Case 2: $p = 1$.

Case 1: $1 < p < +\infty$. Then, $1 < q < +\infty$. $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, by Lemma 7.75, $\exists \bar{y}_i \in \mathcal{Y}$ with $\|\bar{y}_i\| \leq 1$ such that $\langle y_{*i}, \bar{y}_i \rangle \in \mathbb{R}$ and $\langle y_{*i}, \bar{y}_i \rangle \geq (1 - \epsilon) \|y_{*i}\|$. Let $\hat{y}_i = \|y_{*i}\|^{q/p} \bar{y}_i$, then $\|\hat{y}_i\| \leq \|y_{*i}\|^{q/p}$, $\langle y_{*i}, \hat{y}_i \rangle \in \mathbb{R}$, and $\langle y_{*i}, \hat{y}_i \rangle \geq (1 - \epsilon) \|y_{*i}\|^{q/p+1} = (1 - \epsilon) \|y_{*i}\|^q$. Then, $\forall n \in \mathbb{N}$, $|f(\sum_{i=1}^n A_i \hat{y}_i)| = |\sum_{i=1}^n \langle y_{*i}, \hat{y}_i \rangle| = \sum_{i=1}^n \langle y_{*i}, \hat{y}_i \rangle \geq (1 - \epsilon) \sum_{i=1}^n \|y_{*i}\|^q$. Since $f \in \mathcal{X}^*$, then, by Proposition 7.72, we have $(1 - \epsilon) \sum_{i=1}^n \|y_{*i}\|^q \leq \|f\| \|\sum_{i=1}^n A_i \hat{y}_i\| \leq \|f\| (\sum_{i=1}^n \|\hat{y}_i\|^p)^{1/p} \leq \|f\| \cdot (\sum_{i=1}^n \|y_{*i}\|^q)^{1/p}$. Then, we have $(\sum_{i=1}^n \|y_{*i}\|^q)^{1/q} \leq \|f\| / (1 - \epsilon)$. By the arbitrariness of n , we have $(\sum_{i=1}^{\infty} \|y_{*i}\|^q)^{1/q} \leq \|f\| / (1 - \epsilon)$. By the

arbitrariness of ϵ , we have $(\sum_{i=1}^{\infty} \|y_{*i}\|^q)^{1/q} \leq \|f\|$. Hence, the result holds in this case.

Case 2: $p = 1$. Then, $q = +\infty$. $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, by Lemma 7.75, $\exists \hat{y}_i \in \mathcal{Y}$ with $\|\hat{y}_i\| \leq 1$ such that $\langle\langle y_{*i}, \hat{y}_i \rangle\rangle \in \mathbb{R}$ and $\langle\langle y_{*i}, \hat{y}_i \rangle\rangle \geq (1 - \epsilon) \|y_{*i}\|$. Then, $(1 - \epsilon) \|y_{*i}\| \leq |\langle\langle y_{*i}, \hat{y}_i \rangle\rangle| = |f(A_i \hat{y}_i)| \leq \|f\| \|A_i\| \|\hat{y}_i\| \leq \|f\|$. By the arbitrariness of n , we have $\sup_{i \geq 1} \|y_{*i}\| \leq \|f\| / (1 - \epsilon)$. By the arbitrariness of ϵ , we have $\sup_{i \geq 1} \|y_{*i}\| \leq \|f\|$. Hence, the result holds in this case.

This completes the proof of the claim. $\square$

The preceding analysis shows that we may define a function $\psi : \mathcal{X}^* \rightarrow l_q(\mathcal{Y}^*)$ by $\psi(f) = (f \circ A_1, f \circ A_2, \dots)$, $\forall f \in \mathcal{X}^*$. Clearly, ψ is linear. $\forall f_1, f_2 \in \mathcal{X}^*$ with $\psi(f_1) = \psi(f_2)$. Then, we have $f_1 \circ A_i = f_2 \circ A_i$, $\forall i \in \mathbb{N}$. $\forall x := (y_1, y_2, \dots) \in \mathcal{X}$, by (7.3), we have $f_1(x) = \lim_{n \in \mathbb{N}} \sum_{i=1}^n f_1(A_i y_i) = \lim_{n \in \mathbb{N}} \sum_{i=1}^n f_2(A_i y_i) = f_2(x)$. Then, we have $f_1 = f_2$. Hence, ψ is injective. $\forall x_* := (y_{*1}, y_{*2}, \dots) \in l_q(\mathcal{Y}^*)$, define $\phi(x_*) : \mathcal{X} \rightarrow \mathbb{K}$ by $\phi(x_*)(x) = \sum_{i=1}^{\infty} \langle\langle y_{*i}, y_i \rangle\rangle$, $\forall x := (y_1, y_2, \dots) \in \mathcal{X}$. Note that, $\forall x := (y_1, y_2, \dots) \in \mathcal{X}$, $\sum_{i=1}^{\infty} |\langle\langle y_{*i}, y_i \rangle\rangle| \leq \sum_{i=1}^{\infty} \|y_{*i}\| \|y_i\| \leq \|x_*\| \|x\|$, where we have made use of Proposition 7.72 in the first inequality and Hölder's Inequality in the second inequality. Hence, $\phi(x_*)(x)$ is well-defined. Hence, $\phi(x_*)$ is well-defined. Clearly, $\phi(x_*)$ is linear and bounded with

$$\begin{aligned} \|\phi(x_*)\| &= \sup_{x \in \mathcal{X}, \|x\| \leq 1} |\phi(x_*)(x)| \leq \sup_{x = (y_1, y_2, \dots) \in \mathcal{X}, \|x\| \leq 1} \sum_{i=1}^{\infty} |\langle\langle y_{*i}, y_i \rangle\rangle| \\ &\leq \sup_{x \in \mathcal{X}, \|x\| \leq 1} \|x_*\| \|x\| \leq \|x_*\| \end{aligned} \quad (7.4)$$

Hence, $\phi(x_*) \in \mathcal{X}^*$ and $\psi(\phi(x_*)) = (\phi(x_*) \circ A_1, \phi(x_*) \circ A_2, \dots) = (y_{*1}, y_{*2}, \dots) = x_*$. Then, $\psi \circ \phi = \text{id}_{l_q(\mathcal{Y}^*)}$. Hence, ψ is surjective. Therefore, ψ is bijective and admits inverse ψ_{inv} . By Proposition 2.4, $\phi = \psi_{\text{inv}}$. $\forall f \in \mathcal{X}^*$, by Claim 7.77.1 and (7.4), we have $\|f\| = \|\phi(\psi(f))\| \leq \|\psi(f)\| \leq \|f\|$. Then, $\|\psi(f)\| = \|f\|$ and ψ is an isometry. Hence, ψ is an isometrical isomorphism. $\diamond$

Example 7.78 Let $\mathcal{Y}$ be a normed linear space over the field $\mathbb{K}$, $\mathcal{Y}^*$ be its dual, and $M = \{x = (y_1, y_2, \dots) \in l_{\infty}(\mathcal{Y}) \mid \lim_{n \in \mathbb{N}} y_n = \vartheta_{\mathcal{Y}}\}$. Clearly, M is a subspace of $l_{\infty}(\mathcal{Y})$. By Proposition 7.13, M is a normed linear space over $\mathbb{K}$, which will be denoted by $c_0(\mathcal{Y})$. Next, we will study the dual of $\mathcal{X} := c_0(\mathcal{Y})$. We will show that $\mathcal{X}^* = l_1(\mathcal{Y}^*)$ isometrically isomorphically.

Let $f \in \mathcal{X}^*$. $\forall i \in \mathbb{N}$, define $A_i : \mathcal{Y} \rightarrow \mathcal{X}$ by $x = (y_1, y_2, \dots) = A_i y$, $\forall y \in \mathcal{Y}$, and $y_j = \vartheta_{\mathcal{Y}}$, $\forall j \in \mathbb{N}$ with $j \neq i$, and $y_i = y$. Clearly, A_i is well-defined and linear and bounded with $\|A_i\| \leq 1$. Define $f_i : \mathcal{Y} \rightarrow \mathbb{K}$ by $f_i = f \circ A_i$. Then, f_i is a bounded linear functional on $\mathcal{Y}$ with $\|f_i\| \leq \|f\| \|A_i\| \leq \|f\|$. Hence, $f_i \in \mathcal{Y}^*$. Denote $f_i =: y_{*i}$. $\forall x = (y_1, y_2, \dots) \in \mathcal{X}$, we have $y_i \in \mathcal{Y}$, $\forall i \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} y_i = \vartheta_{\mathcal{Y}}$, and $\lim_{n \in \mathbb{N}} \|y_i\| = 0$ (by Proposition 7.21). Then, $\lim_{n \in \mathbb{N}} \sup_{i \geq n+1} \|y_i\| = 0$. This implies that $\lim_{n \in \mathbb{N}} \|x - \sum_{i=1}^n A_i y_i\| = \lim_{n \in \mathbb{N}} \sup_{k \geq n+1} \|y_k\| = 0$.

$\lim_{n \in \mathbb{N}} \max_{k \geq n+1} \|y_k\| = 0$. Hence, $\lim_{n \in \mathbb{N}} \sum_{i=1}^n A_i y_i = x$. By the continuity of f and Proposition 3.66, we have

$$f(x) = \lim_{n \in \mathbb{N}} f\left(\sum_{i=1}^n A_i y_i\right) = \lim_{n \in \mathbb{N}} \sum_{i=1}^n f(A_i y_i) = \lim_{n \in \mathbb{N}} \sum_{i=1}^n \langle y_{*i}, y_i \rangle \quad (7.5)$$

Let $x_* := (y_{*1}, y_{*2}, \dots)$. $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, by Lemma 7.75, $\exists \hat{y}_i \in \mathcal{Y}$ with $\|\hat{y}_i\| \leq 1$ such that $\langle y_{*i}, \hat{y}_i \rangle \in \mathbb{R}$ and $\langle y_{*i}, \hat{y}_i \rangle \geq (1 - \epsilon) \|y_{*i}\|$. Then, $\forall n \in \mathbb{N}$, $(1 - \epsilon) \sum_{i=1}^n \|y_{*i}\| \leq |\sum_{i=1}^n \langle y_{*i}, \hat{y}_i \rangle| = |\sum_{i=1}^n f(A_i \hat{y}_i)| = |f(\sum_{i=1}^n A_i \hat{y}_i)| \leq \|f\| \|\sum_{i=1}^n A_i \hat{y}_i\| \leq \|f\|$. By the arbitrariness of n , we have $\sum_{i=1}^{\infty} \|y_{*i}\| \leq \|f\| / (1 - \epsilon)$. By the arbitrariness of ϵ , we have

$$\sum_{i=1}^{\infty} \|y_{*i}\| \leq \|f\| \quad (7.6)$$

Hence, $x_* \in l_1(\mathcal{Y}^*)$.

Based on the preceding analysis, we may define a function $\psi : \mathcal{X}^* \rightarrow l_1(\mathcal{Y}^*)$ by $\psi(f) = (f \circ A_1, f \circ A_2, \dots)$, $\forall f \in \mathcal{X}^*$. Clearly, ψ is linear. $\forall f_1, f_2 \in \mathcal{X}^*$ with $\psi(f_1) = \psi(f_2)$. Then, we have $f_1 \circ A_i = f_2 \circ A_i$, $\forall i \in \mathbb{N}$. $\forall x := (y_1, y_2, \dots) \in \mathcal{X}$, by (7.5), we have $f_1(x) = \lim_{n \in \mathbb{N}} \sum_{i=1}^n f_1(A_i y_i) = \lim_{n \in \mathbb{N}} \sum_{i=1}^n f_2(A_i y_i) = f_2(x)$. Then, we have $f_1 = f_2$. Hence, ψ is injective. $\forall x_* := (y_{*1}, y_{*2}, \dots) \in l_1(\mathcal{Y}^*)$, define $\phi(x_*) : \mathcal{X} \rightarrow \mathbb{K}$ by $\phi(x_*)(x) = \sum_{i=1}^{\infty} \langle y_{*i}, y_i \rangle$, $\forall x := (y_1, y_2, \dots) \in \mathcal{X}$. Note that, $\forall x := (y_1, y_2, \dots) \in \mathcal{X}$, $|\sum_{i=1}^{\infty} \langle y_{*i}, y_i \rangle| \leq \sum_{i=1}^{\infty} \|y_{*i}\| \|y_i\| \leq \|x_*\| \|x\|$, where we have made use of Proposition 7.72 in the first inequality and Hölder's Inequality in the second inequality. Hence, $\phi(x_*)(x)$ is well-defined. Hence, $\phi(x_*)$ is well-defined. Clearly, $\phi(x_*)$ is linear and bounded with

$$\begin{aligned} \|\phi(x_*)\| &= \sup_{x \in \mathcal{X}, \|x\| \leq 1} |\phi(x_*)(x)| \leq \sup_{x=(y_1, y_2, \dots) \in \mathcal{X}, \|x\| \leq 1} \sum_{i=1}^{\infty} |\langle y_{*i}, y_i \rangle| \\ &\leq \sup_{x \in \mathcal{X}, \|x\| \leq 1} \|x_*\| \|x\| \leq \|x_*\| \end{aligned} \quad (7.7)$$

Hence, $\phi(x_*) \in \mathcal{X}^*$ and $\psi(\phi(x_*)) = (\phi(x_*) \circ A_1, \phi(x_*) \circ A_2, \dots) = (y_{*1}, y_{*2}, \dots) = x_*$. Then, $\psi \circ \phi = \text{id}_{l_1(\mathcal{Y}^*)}$. Hence, ψ is surjective. Therefore, ψ is bijective and admits inverse ψ_{inv} . By Proposition 2.4, $\phi = \psi_{\text{inv}}$. $\forall f \in \mathcal{X}^*$, by (7.6) and (7.7), we have $\|f\| = \|\phi(\psi(f))\| \leq \|\psi(f)\| \leq \|f\|$. Then, $\|\psi(f)\| = \|f\|$ and ψ is an isometry. Hence, ψ is an isometrical isomorphism. $\diamond$

7.9.3 Extension form of Hahn-Banach Theorem

Definition 7.79 Let $\mathcal{X}$ be a vector space over $\mathbb{K}$. A sublinear functional is $p : \mathcal{X} \rightarrow \mathbb{R}$ satisfying, $\forall x_1, x_2 \in \mathcal{X}$, $\forall \alpha \in \mathbb{R}$ with $\alpha \geq 0$, (i) $p(x_1 + x_2) \leq p(x_1) + p(x_2)$; and (ii) $p(\alpha x_1) = \alpha p(x_1)$.

Note that any $\|\cdot\|$ on $\mathcal{X}$ is a sublinear functional. We introduce the above definition to illustrate the full generality of the Hahn-Banach Theorem.

Theorem 7.80 (Extension Form of Hahn-Banach Theorem) *Let $\mathcal{X}$ be a vector space over the field $\mathbb{R}$, $p : \mathcal{X} \rightarrow \mathbb{R}$ be a sublinear functional, $M \subseteq \mathcal{X}$ be a subspace, and $f : M \rightarrow \mathbb{R}$ be a linear functional on M satisfying $f(m) \leq p(m)$, $\forall m \in M$. Then, $\exists$ a linear functional $F : \mathcal{X} \rightarrow \mathbb{R}$ such that $F|_M = f$ and $F(x) \leq p(x)$, $\forall x \in \mathcal{X}$. Furthermore, if $\mathcal{X}$ is normed with $\|\cdot\|$ and p is continuous at $\vartheta_{\mathcal{X}}$, then F is continuous.*

Proof We will prove the theorem using Zorn's Lemma. Define a collection of extensions of f , $\mathcal{E}$, by

$$\mathcal{E} := \{ (g, N) \mid N \subseteq \mathcal{X} \text{ is a subspace, } M \subseteq N, \ g : N \rightarrow \mathbb{R} \text{ is a linear functional, such that } g|_M = f \text{ and } g(n) \leq p(n), \forall n \in N \}$$

Clearly, $(f, M) \in \mathcal{E} \neq \emptyset$. Define a relation $\preceq$ on $\mathcal{E}$ by $\forall (g_1, N_1), (g_2, N_2) \in \mathcal{E}$, we say $(g_1, N_1) \preceq (g_2, N_2)$ if $N_1 \subseteq N_2$ and $g_2|_{N_1} = g_1$. Clearly, $\preceq$ is reflexive, antisymmetric, and transitive. Hence, $\preceq$ is an antisymmetric partial ordering on $\mathcal{E}$.

For any nonempty subcollection $\mathcal{C} \subseteq \mathcal{E}$ such that $\preceq$ is a total ordering on $\mathcal{C}$, let $N_c = \bigcup_{(g, N) \in \mathcal{C}} N$. Then, $\vartheta_{\mathcal{X}} \in M \subseteq N_c \subseteq \mathcal{X}$. $\forall x_1, x_2 \in N_c$, $\forall \alpha, \beta \in \mathbb{R}$, $\exists (g_1, N_1), (g_2, N_2) \in \mathcal{C}$ such that $x_1 \in N_1$ and $x_2 \in N_2$. Since $\mathcal{C}$ is totally ordered, by Proposition 2.12, then, without loss of generality, we may assume that $(g_1, N_1) \preceq (g_2, N_2)$. Then, $N_1 \subseteq N_2$ and $x_1, x_2 \in N_2$. This implies that $\alpha x_1 + \beta x_2 \in N_2 \subseteq N_c$, since N_2 is a subspace. The above shows that N_c is a subspace of $\mathcal{X}$.

Define a functional $g_c : N_c \rightarrow \mathbb{R}$ by $\forall x \in N_c$, $\exists (g, N) \in \mathcal{C}$ such that $x \in N$, we assign $g_c(x) := g(x)$. Such a functional is uniquely defined because of the following reasoning. $\forall x \in N_c$, $\forall (g_1, N_1), (g_2, N_2) \in \mathcal{C}$ such that $x \in N_1 \cap N_2$. By the total ordering of $\mathcal{C}$ and Proposition 2.12, we may assume that, without loss of generality, $(g_1, N_1) \preceq (g_2, N_2)$. Then, $x \in N_1 \subseteq N_2$ and $g_2|_{N_1} = g_1$. Hence, $g_1(x) = g_2(x)$. Therefore, g_c is well-defined.

$\forall x_1, x_2 \in N_c$, $\forall \alpha, \beta \in \mathbb{R}$, $\exists (g_1, N_1), (g_2, N_2) \in \mathcal{C}$ such that $x_1 \in N_1$ and $x_2 \in N_2$. By the total ordering of $\mathcal{C}$ and Proposition 2.12, we may assume that, without loss of generality, $(g_1, N_1) \preceq (g_2, N_2)$. Then, $x_1, x_2 \in N_2$ and $g_c(\alpha x_1 + \beta x_2) = g_2(\alpha x_1 + \beta x_2) = \alpha g_2(x_1) + \beta g_2(x_2) = \alpha g_c(x_1) + \beta g_c(x_2)$. Hence, g_c is a linear functional. $g_c(x_1) = g_2(x_1) \leq p(x_1)$. $\forall m \in M$, $g_c(m) = g_2(m) = f(m)$. Therefore, $(g_c, N_c) \in \mathcal{E}$.

$\forall (g, N) \in \mathcal{C}$, we have $N \subseteq N_c$ and $\forall n \in N$, $g_c(n) = g(n)$. Then, $(g, N) \preceq (g_c, N_c)$. Hence, (g_c, N_c) is an upper bound of $\mathcal{C}$.

By Zorn's Lemma, $\exists (g_M, N_M) \in \mathcal{E}$, which is maximal with respect to $\preceq$. Now, we are going to show that $N_M = \mathcal{X}$. Suppose $N_M \subset \mathcal{X}$. Then, $\exists x_0 \in \mathcal{X} \setminus N_M$. Let $N_e = \{x \in \mathcal{X} \mid \exists \alpha \in \mathbb{R}, \exists n \in N_M \ni x = \alpha x_0 + n\}$. Clearly, $N_M \subset N_e$, $x_0 \in N_e$, and N_e is a subspace of $\mathcal{X}$. $\forall x \in N_e$, $\exists! \alpha \in \mathbb{R}$

and $\exists! n \in N_M$ such that $x = \alpha x_0 + n$ (otherwise, we may deduce that $x_0 \in N_M$). Define $g_e : N_e \rightarrow \mathbb{R}$ by $g_e(x) = g_e(\alpha x_0 + n) = g_M(n) + \alpha g_e(x_0)$, $\forall x = \alpha x_0 + n \in N_e$, $\alpha \in \mathbb{R}$, and $n \in N_M$, where $g_e(x_0) \in \mathbb{R}$ is a constant to be determined. Clearly, g_e is well-defined and is a linear functional on N_e . $\forall n \in N_M$, we have $g_e(n) = g_M(n)$, and hence, $g_e|_{N_M} = g_M$. Then, $g_e|_M = f$.

We will show that $(g_e, N_e) \in \mathcal{E}$ by finding an admissible constant $g_e(x_0)$ such that $g_e(x) \leq p(x)$, $\forall x \in N_e$. $\forall n_1, n_2 \in N_M$, we have $g_M(n_1) + g_M(n_2) = g_M(n_1 + n_2) \leq p(n_1 + n_2) \leq p(n_1 + x_0) + p(n_2 - x_0)$. Then, $g_M(n_2) - p(n_2 - x_0) \leq p(n_1 + x_0) - g_M(n_1)$. By the arbitrariness of n_1 and n_2 , we have $\sup_{n_2 \in N_M} (g_M(n_2) - p(n_2 - x_0)) \leq \inf_{n_1 \in N_M} (p(n_1 + x_0) - g_M(n_1))$. Since $\vartheta_{\mathcal{X}} \in N_M$, then $\exists c \in \mathbb{R}$ such that

$$\begin{aligned} -\infty &< \sup_{n_2 \in N_M} (g_M(n_2) - p(n_2 - x_0)) \leq c \\ &\leq \inf_{n_1 \in N_M} (p(n_1 + x_0) - g_M(n_1)) < +\infty \end{aligned}$$

Let $g_e(x_0) := c$.

$\forall x \in N_e$, $x = \alpha x_0 + n$, where $\alpha \in \mathbb{R}$ and $n \in N_M$. We will distinguish three exhaustive and mutually exclusive cases: Case 1: $\alpha > 0$; Case 2: $\alpha = 0$; Case 3: $\alpha < 0$. Case 1: $\alpha > 0$. $g_e(x) = \alpha c + g_M(n) = \alpha(c + g_M(n/\alpha)) \leq \alpha(p(n/\alpha + x_0) - g_M(n/\alpha) + g_M(n/\alpha)) = \alpha p(x/\alpha) = p(x)$. Case 2: $\alpha = 0$. Then, $x \in N_M$ and $g_e(x) = g_M(n) \leq p(n) = p(x)$. Case 3: $\alpha < 0$. $g_e(x) = \alpha c + g_M(n) = \alpha(c + g_M(n/\alpha)) \leq \alpha(g_M(-n/\alpha) - p(-n/\alpha - x_0) + g_M(n/\alpha)) = -\alpha p(-x/\alpha) = p(x)$. Hence, we have $g_e(x) \leq p(x)$ in all three cases. Therefore, $(g_e, N_e) \in \mathcal{E}$. We have shown that $(g_M, N_M) \preceq (g_e, N_e)$ and $(g_M, N_M) \neq (g_e, N_e)$. This contradicts the fact that (g_M, N_M) is maximal in $\mathcal{E}$ by Proposition 2.12.

Hence, $N_M = \mathcal{X}$ and $F = g_M$.

If, in addition, $\mathcal{X}$ is normed with $\|\cdot\|$ and p is continuous at $\vartheta_{\mathcal{X}}$, then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $|p(x)| < \epsilon$, $\forall x \in \mathcal{B}(\vartheta_{\mathcal{X}}, \delta)$. $\forall x \in \mathcal{B}(\vartheta_{\mathcal{X}}, \delta)$, we will distinguish two exhaustive and mutually exclusive cases: Case 1: $F(x) \geq 0$; Case 2: $F(x) < 0$. Case 1: $F(x) \geq 0$. Then, $0 \leq F(x) \leq p(x) < \epsilon$ and $|F(x)| < \epsilon$. Case 2: $F(x) < 0$. Then, $|F(x)| = -F(x) = F(-x) \leq p(-x) < \epsilon$. Hence, in both cases, we have $|F(x)| < \epsilon$. Then, F is continuous at ϑ . Thus, by Proposition 7.72, F is a bounded linear functional, and therefore continuous.

This completes the proof of the theorem. $\square$

To obtain the counterpart result for complex vector spaces, we need the following result.

Lemma 7.81 *Let $(\mathcal{X}, \mathbb{C})$ be a vector space, $f : \mathcal{X} \rightarrow \mathbb{C}$ be a functional on $(\mathcal{X}, \mathbb{C})$, and $g : \mathcal{X} \rightarrow \mathbb{R}$ be given by $g(x) = \operatorname{Re}(f(x))$, $\forall x \in \mathcal{X}$. Then, f is a linear functional on $(\mathcal{X}, \mathbb{C})$ if, and only if, g is a linear functional on $(\mathcal{X}, \mathbb{R})$ and $f(x) = g(x) - ig(ix)$, $\forall x \in \mathcal{X}$.*

Proof By Lemma 7.40, $(\mathcal{X}, \mathbb{R})$ is a vector space. “Necessity” $\forall x_1, x_2 \in \mathcal{X}$, $\forall \alpha, \beta \in \mathbb{R}$, we have $g(\alpha x_1 + \beta x_2) = \operatorname{Re}(f(\alpha x_1 + \beta x_2)) = \operatorname{Re}(\alpha f(x_1) + \beta f(x_2)) = \alpha \operatorname{Re}(f(x_1)) + \beta \operatorname{Re}(f(x_2)) = \alpha g(x_1) + \beta g(x_2)$. Hence, g is a linear functional on $(\mathcal{X}, \mathbb{R})$. $\forall x \in \mathcal{X}$, let $f(x) = g(x) + i \operatorname{Im}(f(x))$. Since f is a linear functional on $(\mathcal{X}, \mathbb{C})$, then $f(ix) = g(ix) + i \operatorname{Im}(f(ix)) = if(x) = -\operatorname{Im}(f(x)) + ig(x)$. Then, $\operatorname{Im}(f(x)) = -g(ix)$ and $f(x) = g(x) - ig(ix)$.

“Sufficiency” $\forall x_1, x_2 \in \mathcal{X}$, $\forall \alpha := \alpha_r + i\alpha_i, \beta := \beta_r + i\beta_i \in \mathbb{C}$, where $\alpha_r, \alpha_i, \beta_r, \beta_i \in \mathbb{R}$, we have

$$\begin{aligned} f(\alpha x_1 + \beta x_2) &= g(\alpha x_1 + \beta x_2) - ig(ix_1 + ix_2) \\ &= g(\alpha_r x_1 + i\alpha_i x_1 + \beta_r x_2 + i\beta_i x_2) - ig(\alpha_r ix_1 - \alpha_i x_1 + \beta_r ix_2 - \beta_i x_2) \\ &= \alpha_r g(x_1) + \alpha_i g(ix_1) + \beta_r g(x_2) + \beta_i g(ix_2) - i\alpha_r g(ix_1) + i\alpha_i g(x_1) \\ &\quad - i\beta_r g(ix_2) + i\beta_i g(x_2) \\ &= \alpha_r f(x_1) + i\alpha_i f(x_1) + \beta_r f(x_2) + i\beta_i f(x_2) = \alpha f(x_1) + \beta f(x_2) \end{aligned}$$

Hence, f is a linear functional on $(\mathcal{X}, \mathbb{C})$.

This completes the proof of the lemma. $\square$

Theorem 7.82 (Extension Form of Hahn-Banach Theorem) *Let $\mathcal{X}$ be a vector space over the field $\mathbb{C}$, $p : \mathcal{X} \rightarrow \mathbb{R}$ be a sublinear functional, $M \subseteq \mathcal{X}$ be a subspace (of $(\mathcal{X}, \mathbb{C})$), and $f : M \rightarrow \mathbb{C}$ be a linear functional satisfying $\operatorname{Re}(f(m)) \leq p(m)$, $\forall m \in M$. Then, there exists a linear functional $F : \mathcal{X} \rightarrow \mathbb{C}$, such that $F|_M = f$ and $\operatorname{Re}(F(x)) \leq p(x)$, $\forall x \in \mathcal{X}$. Furthermore, if $\mathcal{X}$ is normed with $\|\cdot\|$ and p is continuous at $\vartheta_{\mathcal{X}}$, then F is continuous.*

Proof By Lemma 7.40, $(\mathcal{X}, \mathbb{R})$ is a vector space and $(M, \mathbb{R})$ is also a vector space. It is easy to show that $(M, \mathbb{R})$ is a subspace of $(\mathcal{X}, \mathbb{R})$. Define $g : M \rightarrow \mathbb{R}$ by $g(m) = \operatorname{Re}(f(m))$, $\forall m \in M$. By Lemma 7.81, g is a linear functional on $(M, \mathbb{R})$ and $f(m) = g(m) - ig(im)$, $\forall m \in M$. Note that p is a sublinear functional on $(\mathcal{X}, \mathbb{C})$, then it is a sublinear functional on $(\mathcal{X}, \mathbb{R})$. Furthermore, $g(m) = \operatorname{Re}(f(m)) \leq p(m)$, $\forall m \in M$. Then, by Hahn-Banach Theorem, Theorem 7.80, $\exists$ a linear functional G on $(\mathcal{X}, \mathbb{R})$ such that $G|_M = g$ and $G(x) \leq p(x)$, $\forall x \in \mathcal{X}$.

Define a functional F on $(\mathcal{X}, \mathbb{C})$ by $F(x) = G(x) - iG(ix)$, $\forall x \in \mathcal{X}$. By Lemma 7.81, F is a linear functional on $(\mathcal{X}, \mathbb{C})$. $\forall m \in M$, $im \in M$, since M is a subspace of $(\mathcal{X}, \mathbb{C})$. Then, $F(m) = G(m) - iG(im) = g(m) - ig(im) = f(m)$. Hence, we have $F|_M = f$. $\forall x \in \mathcal{X}$, $\operatorname{Re}(F(x)) = G(x) \leq p(x)$. Hence, F is the functional we seek.

Furthermore, if $(\mathcal{X}, \mathbb{C})$ is normed with $\|\cdot\|$ and p is continuous at $\vartheta_{\mathcal{X}}$, then, by Lemma 7.40, $\mathcal{X}_{\mathbb{R}} := (\mathcal{X}, \mathbb{R}, \|\cdot\|)$ is a normed linear space. By Theorem 7.80, G is continuous. Then, $G : \mathcal{X}_{\mathbb{R}} \rightarrow \mathbb{C}$ is continuous. By Propositions 3.12, 3.32, and 7.23, $F : \mathcal{X}_{\mathbb{R}} \rightarrow \mathbb{C}$ is continuous. By Lemma 7.40, F is continuous on $\mathcal{X} := (\mathcal{X}, \mathbb{C}, \|\cdot\|)$.

This completes the proof of the theorem. $\square$

Theorem 7.83 (Simple Version of Hahn-Banach Theorem) *Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $M \subseteq \mathcal{X}$ be a subspace, $f : M \rightarrow \mathbb{K}$ be a linear functional which is bounded (on M), that is $\|f\|_M := \sup_{m \in M, \|m\| \leq 1} |f(m)| \in [0, \infty) \subset \mathbb{R}$. Then, there exists a $F \in \mathcal{X}^*$ such that $F|_M = f$ and $\|F\| = \|f\|_M$.*

Proof Define a functional $p : \mathcal{X} \rightarrow \mathbb{R}$ by $p(x) = \|f\|_M \|x\|$, $\forall x \in \mathcal{X}$. It is easy to check that p is a sublinear functional. We are going to distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathbb{K} = \mathbb{R}$; Case 2: $\mathbb{K} = \mathbb{C}$.

Case 1: $\mathbb{K} = \mathbb{R}$. Clearly, by Proposition 7.72, we have $f(m) \leq |f(m)| \leq p(m)$, $\forall m \in M$, and p is continuous. By Hahn-Banach Theorem, Theorem 7.80, $\exists$ a linear functional $F : \mathcal{X} \rightarrow \mathbb{R}$ such that $F|_M = f$ and $F(x) \leq p(x)$, $\forall x \in \mathcal{X}$. Furthermore, F is continuous. Hence, $F \in \mathcal{X}^*$, by Proposition 7.72. Note that $\|f\|_M = \sup_{m \in M, \|m\| \leq 1} |F(m)| \leq \sup_{x \in \mathcal{X}, \|x\| \leq 1} |F(x)| = \|F\|$. On the other hand, we have, $\forall x \in \mathcal{X}$, $-F(x) = F(-x) \leq p(-x) = \|f\|_M \|x\|$ and $F(x) \leq p(x) = \|f\|_M \|x\|$. Hence, $|F(x)| \leq \|f\|_M \|x\|$. Then, $\|F\| \leq \|f\|_M$. Hence, $\|F\| = \|f\|_M$.

Case 2: $\mathbb{K} = \mathbb{C}$. Note that, $\forall m \in M$, $\operatorname{Re}(f(m)) \leq |\operatorname{Re}(f(m))| \leq |f(m)| \leq p(m)$, by Proposition 7.72. By Hahn-Banach Theorem, Theorem 7.82, there exists a linear functional $F : \mathcal{X} \rightarrow \mathbb{C}$ such that $F|_M = f$ and $\operatorname{Re}(F(x)) \leq p(x)$, $\forall x \in \mathcal{X}$. Furthermore, since p is continuous, then F is continuous. By Proposition 7.72, $F \in \mathcal{X}^*$. Note that $\|f\|_M = \sup_{m \in M, \|m\| \leq 1} |F(m)| \leq \sup_{x \in \mathcal{X}, \|x\| \leq 1} |F(x)| = \|F\|$. On the other hand, $\forall x \in \mathcal{X}$, we have either $F(x) = 0$, then $|F(x)| \leq \|f\|_M \|x\|$; or $F(x) \neq 0$, then $|F(x)| = \frac{|F(x)|}{F(x)} F(x) = F\left(\frac{|F(x)|}{F(x)} x\right) = \operatorname{Re}\left(F\left(\frac{|F(x)|}{F(x)} x\right)\right) \leq p\left(\frac{|F(x)|}{F(x)} x\right) = \|f\|_M \left\|\frac{|F(x)|}{F(x)} x\right\| = \|f\|_M \left|\frac{|F(x)|}{F(x)}\right| \|x\| = \|f\|_M \|x\|$. Thus, we have $|F(x)| \leq \|f\|_M \|x\|$, $\forall x \in \mathcal{X}$. Therefore, $\|F\| \leq \|f\|_M$. Hence, $\|F\| = \|f\|_M$.

This completes the proof of the theorem. $\square$

Corollary 7.84 *Let $\mathcal{X}$ be a vector space over the field $\mathbb{K}$, $M \subseteq \mathcal{X}$ be a subspace, $p : \mathcal{X} \rightarrow \mathbb{R}$ be a sublinear functional, $f : M \rightarrow \mathbb{K}$ be a linear functional on M , and $(\mathcal{G}, \circ, E)$ be an abelian semigroup of linear operators of $\mathcal{X}$ to $\mathcal{X}$. Assume that the following conditions hold.*

- (i) $Am \in M$, $\forall A \in \mathcal{G}$ and $\forall m \in M$.
- (ii) $p(Ax) \leq p(x)$, $\forall x \in \mathcal{X}$ and $\forall A \in \mathcal{G}$.
- (iii) $f(Am) = f(m)$, $\forall A \in \mathcal{G}$ and $\forall m \in M$.
- (iv) $\operatorname{Re}(f(m)) \leq p(m)$, $\forall m \in M$.

Then, there exists a linear functional $F : \mathcal{X} \rightarrow \mathbb{K}$ such that $F|_M = f$, $F(Ax) = F(x)$, $\forall A \in \mathcal{G}$ and $\forall x \in \mathcal{X}$, and $\operatorname{Re}(F(x)) \leq p(x)$, $\forall x \in \mathcal{X}$.

Furthermore, if, in addition, $\mathcal{X}$ is normed with $\|\cdot\|$ and p is continuous at $\vartheta_{\mathcal{X}}$, then F is continuous.

Proof Define a functional $q : \mathcal{X} \rightarrow \mathbb{R}_e$ by

$$q(x) = \inf_{\substack{n \in \mathbb{N} \\ A_1, \dots, A_n \in \mathcal{G}}} \frac{1}{n} p \left(\sum_{i=1}^n A_i x \right); \quad \forall x \in \mathcal{X}$$

Since $p(\vartheta_{\mathcal{X}}) = 0$, then $q(\vartheta_{\mathcal{X}}) = 0$. Since $E \in \mathcal{G}$, then, $\forall x \in \mathcal{X}$, $q(x) \leq p(Ex) \leq p(x) < +\infty$. $\forall x_1, x_2 \in \mathcal{X}$, $\forall n, k \in \mathbb{N}$, $\forall A_1, \dots, A_n \in \mathcal{G}$, $\forall B_1, \dots, B_k \in \mathcal{G}$, we have

$$\begin{aligned} q(x_1 + x_2) &\leq \frac{1}{nk} p \left(\sum_{i=1}^n \sum_{j=1}^k A_i \circ B_j (x_1 + x_2) \right) \\ &= \frac{1}{nk} p \left(\sum_{i=1}^n \sum_{j=1}^k A_i \circ B_j (x_1) + \sum_{i=1}^n \sum_{j=1}^k A_i \circ B_j (x_2) \right) \\ &= \frac{1}{nk} p \left(\sum_{i=1}^n A_i \left(\sum_{j=1}^k B_j (x_1) \right) + \sum_{j=1}^k B_j \left(\sum_{i=1}^n A_i (x_2) \right) \right) \\ &\leq \frac{1}{nk} p \left(\sum_{i=1}^n A_i \left(\sum_{j=1}^k B_j (x_1) \right) \right) + \frac{1}{nk} p \left(\sum_{j=1}^k B_j \left(\sum_{i=1}^n A_i (x_2) \right) \right) \\ &\leq \frac{1}{nk} \sum_{i=1}^n p \left(A_i \left(\sum_{j=1}^k B_j (x_1) \right) \right) + \frac{1}{nk} \sum_{j=1}^k p \left(B_j \left(\sum_{i=1}^n A_i (x_2) \right) \right) \\ &\leq \frac{1}{nk} \sum_{i=1}^n p \left(\sum_{j=1}^k B_j (x_1) \right) + \frac{1}{nk} \sum_{j=1}^k p \left(\sum_{i=1}^n A_i (x_2) \right) \\ &= \frac{1}{k} p \left(\sum_{j=1}^k B_j (x_1) \right) + \frac{1}{n} p \left(\sum_{i=1}^n A_i (x_2) \right) \end{aligned}$$

where the first two equalities follows from the fact that $\mathcal{G}$ is an albelian semigroup of linear operators on $\mathcal{X}$, the second and third inequalities follows from the fact that p is sublinear, the fourth inequality follows from (ii) in the assumption. Then, by the definition of q , we have

$$q(x_1 + x_2) \leq q(x_1) + q(x_2)$$

and the right-hand-side makes sense since $q(x_1) < +\infty$ and $q(x_2) < +\infty$. Note that $0 = q(\vartheta_{\mathcal{X}}) \leq q(x_1) + q(-x_1)$. Then, $q(x_1) > -\infty$. Hence, $q : \mathcal{X} \rightarrow \mathbb{R}$.

$\forall \alpha \in [0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{X}$, we have

$$\begin{aligned} q(\alpha x) &= \inf_{\substack{n \in \mathbb{N} \\ A_1, \dots, A_n \in \mathcal{G}}} \frac{1}{n} p \left(\sum_{i=1}^n A_i(\alpha x) \right) = \inf_{\substack{n \in \mathbb{N} \\ A_1, \dots, A_n \in \mathcal{G}}} \frac{1}{n} p \left(\alpha \sum_{i=1}^n A_i x \right) \\ &= \inf_{\substack{n \in \mathbb{N} \\ A_1, \dots, A_n \in \mathcal{G}}} \left(\alpha \frac{1}{n} p \left(\sum_{i=1}^n A_i x \right) \right) = \alpha q(x) \end{aligned}$$

where the fourth equality follows from Proposition 3.81 and the fact that $q(x) \in \mathbb{R}$. Hence, we have shown that q is a sublinear functional.

$\forall m \in M$, $\forall n \in \mathbb{N}$, $\forall A_1, \dots, A_n \in \mathcal{G}$, we have

$$\begin{aligned} \operatorname{Re}(f(m)) &= \frac{1}{n} \sum_{i=1}^n \operatorname{Re}(f(A_i m)) = \frac{1}{n} \operatorname{Re} \left(f \left(\sum_{i=1}^n A_i m \right) \right) \\ &\leq \frac{1}{n} p \left(\sum_{i=1}^n A_i m \right) \end{aligned}$$

where the first equality follows from (i) and (iii) in the assumptions, the second equality follows from the linearity of f , and the first inequality follows from (iv) in the assumptions. Hence, we have $\operatorname{Re}(f(m)) \leq q(m)$, $\forall m \in M$.

By the extension forms of Hahn-Banach Theorem (Theorem 7.80 for $\mathbb{K} = \mathbb{R}$ and Theorem 7.82 for $\mathbb{K} = \mathbb{C}$), there exists a linear functional $F : \mathcal{X} \rightarrow \mathbb{K}$ such that $F|_M = f$ and $\operatorname{Re}(F(x)) \leq q(x)$, $\forall x \in \mathcal{X}$.

$\forall x \in \mathcal{X}$, $\forall A \in \mathcal{G}$, $\forall n \in \mathbb{N}$, we have

$$\begin{aligned} q(x - Ax) &\leq \frac{1}{n} p(E(x - Ax) + A(x - Ax) + \dots + A^{n-1}(x - Ax)) \\ &= \frac{1}{n} p(Ex - A^n x) \leq \frac{1}{n} (p(Ex) + p(A^n(-x))) \leq \frac{1}{n} (p(x) + p(-x)) \end{aligned}$$

where the first inequality follows from the definition of q , the second inequality follows from the fact that p is sublinear, and the third inequality follows from (ii) in the assumption. Hence, by the arbitrariness of n , we have $q(x - Ax) \leq 0$. We will show that $F(Ax) = F(x)$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\mathbb{K} = \mathbb{R}$; Case 2: $\mathbb{K} = \mathbb{C}$.

Case 1: $\mathbb{K} = \mathbb{R}$. We have $F(x) - F(Ax) = F(x - Ax) \leq q(x - Ax) \leq 0$. On the other hand, we have $F(Ax) - F(x) = F(-x) - F(A(-x)) = F((-x) - A(-x)) \leq q((-x) - A(-x)) \leq 0$. Hence, $F(x) = F(Ax)$.

Case 2: $\mathbb{K} = \mathbb{C}$. Define $H : \mathcal{X} \rightarrow \mathbb{R}$ by $H(x) = \operatorname{Re}(F(x))$, $\forall x \in \mathcal{X}$. By Lemma 7.40, $(\mathcal{X}, \mathbb{R})$ is a vector space. By Lemma 7.81, H is a linear functional on $(\mathcal{X}, \mathbb{R})$ and $F(x) = H(x) - iH(ix)$, $\forall x \in \mathcal{X}$. Note that $H(x) - H(Ax) = H(x - Ax) = \operatorname{Re}(F(x - Ax)) \leq q(x - Ax) \leq 0$. On the other hand, we have $H(Ax) - H(x) = H(-x) - H(A(-x)) =$

$H((-x) - A(-x)) = \operatorname{Re}(F((-x) - A(-x))) \leq q((-x) - A(-x)) \leq 0$. Hence, $H(x) = H(Ax)$. Then, $F(Ax) = H(Ax) - iH(iAx) = H(Ax) - iH(A(ix)) = H(x) - iH(ix) = F(x)$.

Furthermore, if, in addition, $\mathcal{X}$ is normed with $\|\cdot\|$ and p is continuous at $\vartheta_{\mathcal{X}}$, then $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}, \forall x \in \mathcal{B}(\vartheta_{\mathcal{X}}, \delta)$, we have $|p(x) - p(\vartheta_{\mathcal{X}})| = |p(x)| < \epsilon$. Note that $q(x) \leq p(x) < \epsilon$. Then, $q(-x) \leq p(-x) < \epsilon$. Since q is a sublinear functional, we have $0 = q(\vartheta_{\mathcal{X}}) \leq q(x) + q(-x)$. Hence, we have $q(x) \geq -q(-x) > -\epsilon$. Therefore, $|q(x) - q(\vartheta_{\mathcal{X}})| = |q(x)| < \epsilon$. Hence, q is continuous at $\vartheta_{\mathcal{X}}$. Therefore, by the extension forms of Hahn-Banach Theorem, F is continuous.

This completes the proof of the corollary. $\square$

Proposition 7.85 *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$ and $x_0 \in \mathcal{X}$ with $x_0 \neq \vartheta_{\mathcal{X}}$. Then, $\exists x_* \in \mathcal{X}^*$ with $\|x_*\| = 1$ such that $\langle x_*, x_0 \rangle = \|x_*\| \|x_0\|$.*

Proof Consider the subspace $M := \operatorname{span}(\{x_0\})$. Define $f : M \rightarrow \mathbb{K}$ by $f(\alpha x_0) = \alpha \|x_0\|$, $\forall \alpha \in \mathbb{K}$. Clearly, f is uniquely defined by Proposition 6.17. Clearly f is a linear functional on M and $\|f\|_M := \sup_{m \in M, \|m\| \leq 1} |f(m)| = \sup_{\alpha \in \mathbb{K}, \|\alpha x_0\| \leq 1} |\alpha \|x_0\|| = 1$. By the simple version of Hahn-Banach Theorem, there exists $x_* \in \mathcal{X}^*$ such that $x_*|_M = f$ and $\|x_*\| = \|f\|_M = 1$. Then, $\langle x_*, x_0 \rangle = f(x_0) = \|x_0\| = \|x_*\| \|x_0\|$. This completes the proof of the proposition. $\square$

Example 7.86 Let $\mathcal{Y}$ be a normed linear space over the field $\mathbb{K}$, $\mathcal{X} := l_{\infty}(\mathcal{Y})$ be the normed linear space as defined in Example 7.10, and $y_0 \in \mathcal{Y}$ with $y_0 \neq \vartheta_{\mathcal{Y}}$. We will show that $\mathcal{X}^* \neq l_1(\mathcal{Y}^*)$ by Hahn-Banach Theorem.

Let $M := \{x := (\xi_1, \xi_2, \dots) \in \mathcal{X} \mid \lim_{n \in \mathbb{N}} \xi_n \in \mathcal{Y}\}$. Clearly, M is a subspace of $\mathcal{X}$. By Proposition 7.85, there exists $y_* \in \mathcal{Y}^*$ such that $\|y_*\| = 1$ and $\langle y_*, y_0 \rangle = \|y_0\| > 0$. Define $f : M \rightarrow \mathbb{K}$ by $f(m) = \langle y_*, \lim_{n \in \mathbb{N}} \xi_n \rangle$, $\forall m := (\xi_1, \xi_2, \dots) \in M$. It is easy to show that f is a linear functional on M . $\forall m = (\xi_1, \xi_2, \dots) \in M$ with $\|m\| \leq 1$, we have

$$\begin{aligned} |f(m)| &= |\langle y_*, \lim_{n \in \mathbb{N}} \xi_n \rangle| \leq \|\lim_{n \in \mathbb{N}} \xi_n\| = \lim_{n \in \mathbb{N}} \|\xi_n\| \leq \sup_{n \geq 1} \|\xi_n\| \\ &= \|m\| \leq 1 \end{aligned}$$

where we have applied Propositions 7.72, 7.21, and 3.66. Hence, we have $\|f\|_M := \sup_{m \in M, \|m\| \leq 1} |f(m)| \leq 1$. Consider $m_0 = (y_0, y_0, \dots) \in M$, we have $|f(m_0)| = |\langle y_*, y_0 \rangle| = \|y_0\| = \|m_0\| > 0$. Hence, by Proposition 7.72, $\|f\|_M = 1$. By the simple version of Hahn-Banach Theorem, there exists $x_* \in \mathcal{X}^*$ such that $x_*|_M = f$ and $\|x_*\| = \|f\|_M = 1$. Clearly, there does not exist $(\eta_{*1}, \eta_{*2}, \dots) \in l_1(\mathcal{Y}^*)$ such that x_* is given by $\langle x_*, x \rangle = \sum_{n=1}^{\infty} \langle \eta_{*n}, \xi_n \rangle$, $\forall x = (\xi_1, \xi_2, \dots) \in \mathcal{X}$. Hence, $\mathcal{X}^* \neq l_1(\mathcal{Y}^*)$. $\diamond$

7.9.4 Second dual space

Definition 7.87 Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$ and $\mathcal{X}^*$ be its dual. The dual space of $\mathcal{X}^*$ is called the second dual of $\mathcal{X}$ and is denoted by $\mathcal{X}^{**}$.

Remark 7.88 Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $\mathcal{X}^*$ be its dual, and $\mathcal{X}^{**}$ be its second dual. Then, $\mathcal{X}$ is isometrically isomorphic to a dense subset of a Banach space, which can be taken as a subspace of $\mathcal{X}^{**}$. This can be proved as follows.

By Proposition 7.72, $\mathcal{X}^{**}$ is a Banach space over $\mathbb{K}$. $\forall x \in \mathcal{X}$, define a functional $f : \mathcal{X}^* \rightarrow \mathbb{K}$ by $f(x_*) = \langle x_*, x \rangle$, $\forall x_* \in \mathcal{X}^*$. Clearly, f is a linear functional on $\mathcal{X}^*$. Note that $|f(x_*)| = |\langle x_*, x \rangle| \leq \|x\| \|x_*\|$, $\forall x_* \in \mathcal{X}^*$, by Proposition 7.72. Then, $\|f\| \leq \|x\|$. When $x = \vartheta_{\mathcal{X}}$, then $\|f\| = 0 = \|x\|$. When $x \neq \vartheta_{\mathcal{X}}$, then, by Proposition 7.85, $\exists x_{*0} \in \mathcal{X}^*$ with $\|x_{*0}\| = 1$ such that $\langle x_{*0}, x \rangle = \|x\|$. Then, $\|f\| = \sup_{x_* \in \mathcal{X}^*, \|x_*\| \leq 1} |f(x_*)| \geq |f(x_{*0})| = \|x\|$. Hence, we have $\|f\| = \|x\|$. Hence, $f \in \mathcal{X}^{**}$. Thus, we may define a natural mapping $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ by $\phi(x) = f$, $\forall x \in \mathcal{X}$.

$\forall x_1, x_2 \in \mathcal{X}$, $\forall \alpha, \beta \in \mathbb{K}$, $\forall x_* \in \mathcal{X}^*$, we have $\phi(\alpha x_1 + \beta x_2)(x_*) = \langle x_*, \alpha x_1 + \beta x_2 \rangle = \alpha \langle x_*, x_1 \rangle + \beta \langle x_*, x_2 \rangle = \alpha \phi(x_1)(x_*) + \beta \phi(x_2)(x_*)$. Hence, ϕ is a linear function. $\forall x \in \mathcal{X}$, we have $\|\phi(x)\| = \|x\|$. Hence, ϕ is an isometry and ϕ is injective. Therefore, ϕ is an isometrical isomorphism between $\mathcal{X}$ and $\phi(\mathcal{X})$. Clearly, $\phi(\mathcal{X}) \subseteq \mathcal{X}^{**}$ is a subspace. Then, by Proposition 7.17, $\overline{\phi(\mathcal{X})} \subseteq \mathcal{X}^{**}$ is a closed subspace. By Proposition 4.39, $\overline{\phi(\mathcal{X})}$ is complete. Hence, $\overline{\phi(\mathcal{X})}$ is a Banach space. By Proposition 3.5, $\phi(\mathcal{X})$ is dense in $\overline{\phi(\mathcal{X})}$. This completes the proof of the remark. $\diamond$

Definition 7.89 Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $\mathcal{X}^{**}$ be its second dual, and $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ be the natural mapping as defined in Remark 7.88. $\mathcal{X}$ is said to be reflexive if $\phi(\mathcal{X}) = \mathcal{X}^{**}$, that is $\mathcal{X}$ and $\mathcal{X}^{**}$ are isometrically isomorphic. Then, we may label $\mathcal{X}^{**}$ such that $\mathcal{X}^{**} = \mathcal{X}$ and $\phi = \text{id}_{\mathcal{X}}$.

A reflexive normed linear space is clearly a Banach space.

Proposition 7.90 Let $\mathcal{X}$ be a Banach space over the field $\mathbb{K}$ and $\mathcal{X}^*$ be its dual. Then, $\mathcal{X}$ is reflexive if, and only if, $\mathcal{X}^*$ is reflexive.

Proof Let $\mathcal{X}^{**}$ be the second dual of $\mathcal{X}$ and $\mathcal{X}^{***}$ be the dual of $\mathcal{X}^{**}$.

“Necessity” Let $\mathcal{X}$ be reflexive. Then, $\mathcal{X}^{**} = \mathcal{X}$ isometrically isomorphically. Then, $\mathcal{X}^* = (\mathcal{X}^{**})^* = \mathcal{X}^{***}$ isometrically isomorphically. Hence, $\mathcal{X}^*$ is reflexive.

“Sufficiency” Let $\mathcal{X}^*$ be reflexive. Then, $\mathcal{X}^{***} = \phi_*(\mathcal{X}^*)$, where $\phi_* : \mathcal{X}^* \rightarrow \mathcal{X}^{***}$ is the natural mapping on $\mathcal{X}^*$. We will show that $\mathcal{X}$ is reflexive by an argument of contradiction. Suppose $\mathcal{X}$ is not reflexive. Let $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ be the natural mapping. ϕ is an isometrical isomorphism between

$\mathcal{X}$ and $\phi(\mathcal{X})$. Since $\mathcal{X}$ is complete, then $\phi(\mathcal{X})$ is also complete. Then, by Proposition 4.39, $\phi(\mathcal{X})$ is a closed subspace of $\mathcal{X}^{**}$. Since $\mathcal{X}$ is not reflexive, then there exists $y_{**} \in \mathcal{X}^{**} \setminus \phi(\mathcal{X})$. Then, $y_{**} \neq \vartheta_{**} = \phi(\vartheta)$. Let $\delta := \inf_{m_{**} \in \phi(\mathcal{X})} \|y_{**} - m_{**}\|$. By Proposition 4.10, $\delta \in (0, \infty) \subset \mathbb{R}$. Consider the subspace $N := \{m_{**} + \alpha y_{**} \in \mathcal{X}^{**} \mid \alpha \in \mathbb{K}, m_{**} \in \phi(\mathcal{X})\}$. Since $y_{**} \notin \phi(\mathcal{X})$, then $\forall n_{**} \in N, \exists! \alpha \in \mathbb{K}$ and $\exists! m_{**} \in \phi(\mathcal{X})$ such that $n_{**} = \alpha y_{**} + m_{**}$. Define a linear functional $f : N \rightarrow \mathbb{K}$ by $f(n_{**}) = \alpha$, $\forall n_{**} = \alpha y_{**} + m_{**} \in N$, where $m_{**} \in \phi(\mathcal{X})$ and $\alpha \in \mathbb{K}$. $\forall n_{**} \in N$, let $n_{**} = \alpha y_{**} + m_{**}$, where $m_{**} \in \phi(\mathcal{X})$ and $\alpha \in \mathbb{K}$. We have $|f(n_{**})| = |\alpha| = |\alpha\delta|/\delta \leq \frac{1}{\delta} \|\alpha y_{**} + m_{**}\| = \frac{1}{\delta} \|n_{**}\|$. Hence, $\|f\|_M \leq 1/\delta$. By the simple version of Hahn-Banach Theorem, there exists $F \in \mathcal{X}^{***}$ such that $F|_M = f$ and $\|F\| = \|f\|_M$. Since $\mathcal{X}^*$ is reflexive, then $\phi_*(\mathcal{X}^*) = \mathcal{X}^{***}$. Then, $\exists y_* \in \mathcal{X}^*$ such that $F = \phi_*(y_*)$. Then, $\forall x_{**} \in \mathcal{X}^{**}$, we have $F(x_{**}) = \phi_*(y_*)(x_{**}) = \langle\langle x_{**}, y_* \rangle\rangle$. $\forall m_{**} \in \phi(\mathcal{X})$, we have $\langle\langle m_{**}, y_* \rangle\rangle = F(m_{**}) = f(m_{**}) = 0$. Hence, $\forall x \in \mathcal{X}$, $\langle\langle y_*, x \rangle\rangle = \phi(x)(y_*) = \langle\langle \phi(x), y_* \rangle\rangle = 0$. Hence, $y_* = \vartheta_*$. Then, $F = \phi_*(\vartheta_*) = \vartheta_{***}$. This contradicts with the fact that $F(y_{**}) = f(y_{**}) = 1 \neq 0$. Therefore, $\mathcal{X}$ must be reflexive.

This completes the proof of the proposition. $\square$

Example 7.91 Clearly, the normed linear space $(\{\vartheta_{\mathcal{X}}\}, \mathbb{K}, \|\cdot\|)$ as defined in Example 7.73 is reflexive. Clearly $\mathbb{K}^n$, $n \in \mathbb{N}$, are reflexive. Let $\mathcal{Y}$ be a reflexive Banach space. Then, $l_p(\mathcal{Y})$, $p \in (1, \infty) \subset \mathbb{R}$, are reflexive. $l_1(\mathcal{Y})$ is not reflexive by Example 7.86. Then, by Proposition 7.90, $l_{\infty}(\mathcal{Y}^*)$ is not reflexive either. $\diamond$

Proposition 7.92 Let $\mathcal{X}$ be a finite-dimensional normed linear space over the field $\mathbb{K}$. Then, $\mathcal{X}$ is reflexive.

Proof Let $n \in \mathbb{Z}_+$ be the dimension of $\mathcal{X}$ and $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ be the natural mapping. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $n = 0$; Case 2: $n \in \mathbb{N}$. Case 1: $n = 0$. Then, $\mathcal{X} = (\{\vartheta\}, \mathbb{K}, \|\cdot\|)$. By Example 7.73, $\mathcal{X}^* = (\{\vartheta_*\}, \mathbb{K}, \|\cdot\|_*)$. Then $\mathcal{X}^{**} = (\{\vartheta_{**}\}, \mathbb{K}, \|\cdot\|_{**})$. Clearly, $\phi(\vartheta) = \vartheta_{**}$. Hence, $\phi(\mathcal{X}) = \mathcal{X}^{**}$ and $\mathcal{X}$ is reflexive.

Case 2: $n \in \mathbb{N}$. Let $\{e_1, \dots, e_n\} \subseteq \mathcal{X}$ be a basis of $\mathcal{X}$. $\forall x \in \mathcal{X}$, by Corollary 6.47, x can be uniquely expressed as $\sum_{j=1}^n \alpha_j e_j$ for some $\alpha_1, \dots, \alpha_n \in \mathbb{K}$. $\forall i = 1, \dots, n$, define $f_i : \mathcal{X} \rightarrow \mathbb{K}$ by $f_i(x) = \alpha_i$, $\forall x = \sum_{j=1}^n \alpha_j e_j \in \mathcal{X}$. Clearly, f_i is well-defined and a linear functional. By Proposition 7.67, f_i is continuous and $f_i \in \mathcal{X}^*$. Denote f_i by $e_{*i} \in \mathcal{X}^*$. $\forall x_* \in \mathcal{X}^*$, $\forall i = 1, \dots, n$, let $\beta_i = \langle\langle x_*, e_i \rangle\rangle \in \mathbb{K}$. $\forall x = \sum_{j=1}^n \alpha_j e_j \in \mathcal{X}$, we have

$$\begin{aligned} \langle\langle x_*, x \rangle\rangle &= \sum_{j=1}^n \alpha_j \langle\langle x_*, e_j \rangle\rangle = \sum_{j=1}^n \beta_j \alpha_j = \sum_{k=1}^n \beta_j \langle\langle e_{*j}, x \rangle\rangle \\ &= \langle\langle \sum_{k=1}^n \beta_j e_{*j}, x \rangle\rangle \end{aligned}$$

Hence, $x_* = \sum_{j=1}^n \beta_j e_{*j}$. Therefore, $\mathcal{X}^* = \text{span}(\{e_{*1}, \dots, e_{*n}\})$. $\forall x_{**} \in \mathcal{X}^{**}$, $\forall i = 1, \dots, n$, let $\gamma_i = \langle\langle x_{**}, e_{*i} \rangle\rangle \in \mathbb{K}$. $\forall x_* = \sum_{j=1}^n \beta_j e_{*j} \in \mathcal{X}^*$, we have

$$\begin{aligned} \langle\langle x_{**}, x_* \rangle\rangle &= \sum_{j=1}^n \beta_j \langle\langle x_{**}, e_{*j} \rangle\rangle = \sum_{j=1}^n \beta_j \gamma_j = \sum_{j=1}^n \beta_j \langle\langle e_{*j}, \sum_{k=1}^n \gamma_k e_k \rangle\rangle \\ &= \langle\langle x_*, \sum_{k=1}^n \gamma_k e_k \rangle\rangle = \langle\langle \phi(\sum_{k=1}^n \gamma_k e_k), x_* \rangle\rangle \end{aligned}$$

Hence, $x_{**} = \phi(\sum_{k=1}^n \gamma_k e_k)$ and $\mathcal{X}^{**} \subseteq \phi(\mathcal{X})$. This shows that $\mathcal{X}$ is reflexive.

This completes the proof of the proposition. $\square$

7.9.5 Alignment and orthogonal complements

Definition 7.93 Let $\mathcal{X}$ be a normed linear space, $x \in \mathcal{X}$, $\mathcal{X}^*$ be the dual, and $x_* \in \mathcal{X}^*$. We say that x_* is aligned with x if $\langle\langle x_*, x \rangle\rangle = \|x_*\| \|x\|$.

Clearly, ϑ aligns with any vector in the dual and ϑ_* is aligned with any vector in $\mathcal{X}$. By Proposition 7.85, $\forall x \in \mathcal{X}$ with $x \neq \vartheta$, there exists a $x_* \in \mathcal{X}^*$ with $x_* \neq \vartheta_*$ that is aligned with it.

Definition 7.94 Let $\mathcal{X}$ be a normed linear space, $x \in \mathcal{X}$, $\mathcal{X}^*$ be the dual, and $x_* \in \mathcal{X}^*$. We say that x and x_* are orthogonal if $\langle\langle x_*, x \rangle\rangle = 0$.

Definition 7.95 Let $\mathcal{X}$ be a normed linear space, $S \subseteq \mathcal{X}$, and $\mathcal{X}^*$ be the dual. The orthogonal complement of S , denoted by $S^\perp$, consists of all $x_* \in \mathcal{X}^*$ that is orthogonal to every vector in S , that is $S^\perp := \{x_* \in \mathcal{X}^* \mid \langle\langle x_*, x \rangle\rangle = 0, \forall x \in S\}$.

Definition 7.96 Let $\mathcal{X}$ be a normed linear space, $\mathcal{X}^*$ be the dual, and $S \subseteq \mathcal{X}^*$. The pre-orthogonal complement of S , denoted by ${}^\perp S$, consists of all $x \in \mathcal{X}$ that is orthogonal to every vector in S , that is ${}^\perp S := \{x \in \mathcal{X} \mid \langle\langle x_*, x \rangle\rangle = 0, \forall x_* \in S\}$.

Proposition 7.97 Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $M \subseteq \mathcal{X}$ be a subspace, and $y \in \mathcal{X}$. Then, the following statements hold.

- (i) $\delta := \inf_{m \in M} \|y - m\| = \max_{x_* \in M^\perp, \|x_*\| \leq 1} \text{Re}(\langle\langle x_*, y \rangle\rangle)$, where the maximum is achieved at some $x_{*0} \in M^\perp$ with $\|x_{*0}\| \leq 1$. If the infimum is achieved at $m_0 \in M$ then $y - m_0$ is aligned with x_{*0} .
- (ii) If $\exists m_0 \in M$ and $\exists x_{*0} \in M^\perp$ with $\|x_{*0}\| = 1$ such that $y - m_0$ is aligned with x_{*0} , then the infimum is achieved at m_0 and the maximum is achieved at x_{*0} , that is $\delta = \|y - m_0\| = \langle\langle x_{*0}, y \rangle\rangle = \text{Re}(\langle\langle x_{*0}, y \rangle\rangle)$.

Proof (i) $\forall x_* \in M^\perp$ with $\|x_*\| \leq 1$, $\forall m \in M$, we have

$$\begin{aligned} \operatorname{Re}(\langle x_*, y \rangle) &= \operatorname{Re}(\langle x_*, y - m \rangle) \leq |\operatorname{Re}(\langle x_*, y - m \rangle)| \\ &\leq |\langle x_*, y - m \rangle| \leq \|x_*\| \|y - m\| \leq \|y - m\| \end{aligned}$$

where we have applied Proposition 7.72 in the third inequality. Then, we have

$$\sup_{x_* \in M^\perp, \|x_*\| \leq 1} \operatorname{Re}(\langle x_*, y \rangle) \leq \inf_{m \in M} \|y - m\| = \delta$$

We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\delta = 0$; Case 2: $\delta > 0$. Case 1: $\delta = 0$. Take $x_{*0} := \vartheta_*$. Clearly, $x_{*0} \in M^\perp$ and $\|x_{*0}\| = 0 \leq 1$. Then, $\delta = \operatorname{Re}(\langle x_{*0}, y \rangle) \leq \sup_{x_* \in M^\perp, \|x_*\| \leq 1} \operatorname{Re}(\langle x_*, y \rangle)$. Therefore, we have $\delta = \inf_{m \in M} \|y - m\| = \max_{x_* \in M^\perp, \|x_*\| \leq 1} \operatorname{Re}(\langle x_*, y \rangle)$ and the maximum is achieved at some $x_{*0} \in M^\perp$ with $\|x_{*0}\| \leq 1$. If the infimum is achieved at $m_0 \in M$, then, $0 = \delta = \|y - m_0\|$ and $y = m_0$. Then, clearly $y - m_0$ is aligned with x_{*0} . This case is proved.

Case 2: $\delta > 0$. Then, $y \in \mathcal{X} \setminus M$. Consider the subspace $\operatorname{span}(M \cup \{y\}) =: \bar{M}$. $\forall x \in \bar{M}$, there exists a unique $\alpha \in \mathbb{K}$ and a unique $m \in M$ such that $x = \alpha y + m$. Then, we may define a functional $f : \bar{M} \rightarrow \mathbb{K}$ by $f(x) = f(\alpha y + m) = \alpha \delta$. Clearly, f is a linear functional on $\bar{M}$. $\|f\|_{\bar{M}} := \sup_{x \in \bar{M}, \|x\| \leq 1} |f(x)| \leq \sup_{\alpha \in \mathbb{K}, m \in M, \|\alpha y + m\| \leq 1} |\alpha| \delta \leq \sup_{\alpha \in \mathbb{K}, m \in M, \|\alpha y + m\| \leq 1} \|\alpha y + m\| \leq 1$. By the simple version of Hahn-Banach Theorem, $\exists x_{*0} \in \mathcal{X}^*$ such that $x_{*0}|_{\bar{M}} = f$ and $\|x_{*0}\| = \|f\|_{\bar{M}} \leq 1$. $\forall m \in M$, we have $\langle x_{*0}, m \rangle = f(m) = 0$. Hence, $x_{*0} \in M^\perp$. Note that $\delta = f(y) = \langle x_{*0}, y \rangle$. Hence, $\delta = \inf_{m \in M} \|y - m\| = \max_{x_* \in M^\perp, \|x_*\| \leq 1} \operatorname{Re}(\langle x_*, y \rangle)$ and the maximum is achieved at some $x_{*0} \in M^\perp$ with $\|x_{*0}\| \leq 1$. If the infimum is achieved at $m_0 \in M$, then $\delta = \|y - m_0\| = \langle x_{*0}, y \rangle = \langle x_{*0}, y - m_0 \rangle \leq \|x_{*0}\| \|y - m_0\| \leq \|y - m_0\|$, where the first inequality follows from Proposition 7.72. Hence, $\langle x_{*0}, y - m_0 \rangle = \|x_{*0}\| \|y - m_0\|$ and x_{*0} is aligned with $y - m_0$. This case is proved.

(ii) Note that $\delta \leq \|y - m_0\| = \|x_{*0}\| \|y - m_0\| = \langle x_{*0}, y - m_0 \rangle = \langle x_{*0}, y \rangle = \operatorname{Re}(\langle x_{*0}, y \rangle) \leq \delta$. Then, the result follows.

This completes the proof of the proposition. $\square$

Proposition 7.98 *Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $A \subseteq \mathcal{X}$, and $B \subseteq \mathcal{X}^*$. Then, the following statements holds.*

(i) $A^\perp \subseteq \mathcal{X}^*$ is a closed subspace.

(ii) ${}^\perp B \subseteq \mathcal{X}$ is a closed subspace.

(iii) ${}^\perp(A^\perp) = \overline{\operatorname{span}(A)}$.

Proof (i) $\vartheta_* \in A^\perp \neq \emptyset$. $\forall x_{*1}, x_{*2} \in A^\perp$, $\forall \alpha, \beta \in \mathbb{K}$, $\forall x \in A$, we have $\langle \alpha x_{*1} + \beta x_{*2}, x \rangle = \alpha \langle x_{*1}, x \rangle + \beta \langle x_{*2}, x \rangle = 0$. Hence, $\alpha x_{*1} + \beta x_{*2} \in A^\perp$ and $A^\perp$ is a subspace. $\forall x_* \in \overline{A^\perp}$, by Proposition 4.13, $\exists (x_{*n})_{n=1}^\infty \subseteq A^\perp$ such that $\lim_{n \in \mathbb{N}} x_{*n} = x_*$. Then, $\forall x \in A$, by Propositions 7.72 and 3.66, we have $\langle x_*, x \rangle = \lim_{n \in \mathbb{N}} \langle x_{*n}, x \rangle = 0$. Then, $x_* \in A^\perp$ and $\overline{A^\perp} \subseteq A^\perp$. By Proposition 3.3, $A^\perp = \overline{A^\perp}$ and $A^\perp$ is closed. Hence, (i) follows.

(ii) This can be shown by a similar argument as (i).

(iii) Clearly, $A \subseteq {}^\perp(A^\perp)$. Then, by (ii), we have $M := \overline{\text{span}(A)} \subseteq {}^\perp(A^\perp)$. On the other hand, $A \subseteq M$, then we have $A^\perp \supseteq M^\perp$ and ${}^\perp(A^\perp) \subseteq {}^\perp(M^\perp)$. $\forall x_0 \in {}^\perp(M^\perp)$, by Proposition 7.97, we have

$$\inf_{m \in M} \|x_0 - m\| = \max_{x_* \in M^\perp, \|x_*\| \leq 1} \text{Re}(\langle x_*, x_0 \rangle) = 0$$

Since M is closed, by Proposition 4.10, $x_0 \in M$. Hence, we have ${}^\perp(A^\perp) \subseteq {}^\perp(M^\perp) \subseteq M$. Hence, $M = {}^\perp(A^\perp)$. Hence, (iii) follows.

This completes the proof of the proposition. $\square$

Proposition 7.99 *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $M \subseteq \mathcal{X}$ be a subspace, and $y_* \in \mathcal{X}^*$. Then, the following statements holds.*

- (i) $\delta := \min_{x_* \in M^\perp} \|y_* - x_*\| = \sup_{m \in M, \|m\| \leq 1} \text{Re}(\langle y_*, m \rangle) =: \bar{\delta}$, where the minimum is achieved at some $x_{*0} \in M^\perp$. If the supremum is achieved at $m_0 \in M$ with $\|m_0\| \leq 1$, then m_0 is aligned with $y_* - x_{*0}$.
- (ii) If $\exists m_0 \in M$ with $\|m_0\| = 1$ and $\exists x_{*0} \in M^\perp$ such that $y_* - x_{*0}$ is aligned with m_0 , then the minimum is achieved at x_{*0} and the supremum is achieved at m_0 , that is $\delta = \|y_* - x_{*0}\| = \langle y_*, m_0 \rangle = \text{Re}(\langle y_*, m_0 \rangle)$.

Proof (i) $\forall x_* \in M^\perp$, $\forall m \in M$ with $\|m\| \leq 1$, we have

$$\begin{aligned} \text{Re}(\langle y_*, m \rangle) &= \text{Re}(\langle y_* - x_*, m \rangle) \leq |\text{Re}(\langle y_* - x_*, m \rangle)| \\ &\leq |\langle y_* - x_*, m \rangle| \leq \|y_* - x_*\| \|m\| \leq \|y_* - x_*\| \end{aligned}$$

where we have applied Proposition 7.72 in the third inequality. Then, we have

$$\delta = \inf_{x_* \in M^\perp} \|y_* - x_*\| \geq \sup_{m \in M, \|m\| \leq 1} \text{Re}(\langle y_*, m \rangle) = \bar{\delta}$$

We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\delta = 0$; Case 2: $\delta > 0$. Case 1: $\delta = 0$. By Proposition 7.98, $M^\perp$ is a closed subspace. By Proposition 4.10, we have $y_* \in M^\perp$ the minimum is achieved at a unique vector $x_{*0} := y_*$. Take $m_0 = \vartheta$ with $\|m_0\| = 0 \leq 1$. Then, $\delta = \text{Re}(\langle y_*, m_0 \rangle) \leq \sup_{m \in M, \|m\| \leq 1} \text{Re}(\langle y_*, m \rangle) = \bar{\delta} \leq \delta$. Hence, $\delta = \min_{x_* \in M^\perp} \|y_* - x_*\| = \sup_{m \in M, \|m\| \leq 1} \text{Re}(\langle y_*, m \rangle) = \bar{\delta}$ and the

minimum is achieved at $x_{*0} = y_* \in M^\perp$. If the supremum is achieved at $m_0 \in M$ with $\|m_0\| \leq 1$. Then, clearly m_0 is aligned with $y_* - x_{*0}$. This case is proved.

Case 2: $\delta > 0$. Then, $y_* \notin M^\perp$. Consider the subspace M . Then, we may define a functional $f : M \rightarrow \mathbb{K}$ by $f(m) = \langle y_*, m \rangle$. Clearly, f is a linear functional on M . $\forall m \in M$ with $\|m\| \leq 1$, we have either $f(m) = 0$, then $|f(m)| = 0 \leq \sup_{m \in M, \|m\| \leq 1} \operatorname{Re}(\langle y_*, m \rangle) = \bar{\delta}$; or $f(m) \neq 0$, then $|f(m)| = \frac{|f(m)|}{f(m)} f(m) = \operatorname{Re}\left(\frac{|f(m)|}{f(m)} f(m)\right) = \operatorname{Re}\left(f\left(\frac{|f(m)|}{f(m)} m\right)\right) = \operatorname{Re}\left(\left\langle y_*, \frac{|f(m)|}{f(m)} m \right\rangle\right) \leq \bar{\delta}$. Hence, $\|f\|_M := \sup_{m \in M, \|m\| \leq 1} |f(m)| \leq \bar{\delta} \leq \delta < +\infty$. By the simple version of Hahn-Banach Theorem, $\exists y_{*0} \in \mathcal{X}^*$ such that $y_{*0}|_M = f$ and $\|y_{*0}\| = \|f\|_M \leq \bar{\delta}$. Let $x_{*0} := y_* - y_{*0}$. $\forall m \in M$, we have $\langle x_{*0}, m \rangle = \langle y_*, m \rangle - \langle y_{*0}, m \rangle = 0$. Hence, $x_{*0} \in M^\perp$. Note that $\delta \geq \bar{\delta} \geq \|f\|_M = \|y_{*0}\| = \|y_* - x_{*0}\| \geq \inf_{x_* \in M^\perp} \|y_* - x_*\| = \delta$. Hence, $\delta = \min_{x_* \in M^\perp} \|y_* - x_*\| = \sup_{m \in M, \|m\| \leq 1} \operatorname{Re}(\langle y_*, m \rangle) = \bar{\delta}$ and the minimum is achieved at some $x_{*0} \in M^\perp$. If the supremum is achieved at $m_0 \in M$ with $\|m_0\| \leq 1$, then $\delta = \|y_* - x_{*0}\| \geq \|y_* - x_{*0}\| \|m_0\| \geq |\langle y_* - x_{*0}, m_0 \rangle| = |\langle y_*, m_0 \rangle| \geq \operatorname{Re}(\langle y_*, m_0 \rangle) = \bar{\delta}$, where the second inequality follows from Proposition 7.72. Hence, $\langle y_* - x_{*0}, m_0 \rangle = \|y_* - x_{*0}\| \|m_0\|$ and $y_* - x_{*0}$ is aligned with m_0 . This case is proved.

(ii) Note that $\delta \leq \|y_* - x_{*0}\| = \|y_* - x_{*0}\| \|m_0\| = \langle y_* - x_{*0}, m_0 \rangle = \langle y_*, m_0 \rangle = \operatorname{Re}(\langle y_*, m_0 \rangle) \leq \bar{\delta}$. Hence, the result follows.

This completes the proof of the proposition. $\square$

Proposition 7.100 *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$ and $S \subseteq \mathcal{X}$ be a subspace. By Proposition 7.13, S is a normed linear space over $\mathbb{K}$. Then, the following statement holds.*

(i) S^* is isometrically isomorphic to $\mathcal{X}^*/S^\perp$.

(ii) If $\mathcal{X}$ is reflexive and S is closed, then S is reflexive.

Proof (i) Define a mapping $A : \mathcal{X}^* \rightarrow S^*$ by $\langle A(x_*), s \rangle = \langle x_*, s \rangle$, $\forall x_* \in \mathcal{X}^*$ and $\forall s \in S$. $\forall x_* \in \mathcal{X}^*$, we will show that $A(x_*) \in S^*$. Clearly, $A(x_*)$ is a linear functional on S . $\|A(x_*)\| := \sup_{s \in S, \|s\| \leq 1} |\langle A(x_*), s \rangle| = \sup_{s \in S, \|s\| \leq 1} |\langle x_*, s \rangle| \leq \sup_{s \in S, \|s\| \leq 1} \|x_*\| \|s\| \leq \|x_*\| < +\infty$. Hence, $A(x_*) \in S^*$ and A is well-defined.

$\forall x_{*1}, x_{*2} \in \mathcal{X}^*$, $\forall \alpha, \beta \in \mathbb{K}$, $\forall s \in S$, we have

$$\begin{aligned} \langle A(\alpha x_{*1} + \beta x_{*2}), s \rangle &= \langle \alpha x_{*1} + \beta x_{*2}, s \rangle = \alpha \langle x_{*1}, s \rangle + \beta \langle x_{*2}, s \rangle \\ &= \alpha \langle A(x_{*1}), s \rangle + \beta \langle A(x_{*2}), s \rangle = \langle \alpha A(x_{*1}) + \beta A(x_{*2}), s \rangle \end{aligned}$$

Hence, A is a linear function. Since $\|Ax_*\| \leq \|x_*\|$, $\forall x_* \in \mathcal{X}^*$, then $A \in B(\mathcal{X}^*, S^*)$ with $\|A\| \leq 1$.

By Proposition 7.68, $\mathcal{N}(A) \subseteq \mathcal{X}^*$ is a closed subspace. $\forall x_* \in \mathcal{N}(A)$, we have $Ax_* = \vartheta_{S^*}$. $\forall s \in S$, $0 = \langle Ax_*, s \rangle = \langle x_*, s \rangle$. Hence, $x_* \in S^\perp$.

Therefore, $\mathcal{N}(A) \subseteq S^\perp$. On the other hand, $\forall x_* \in S^\perp, \forall s \in S$, we have $0 = \langle x_*, s \rangle = \langle Ax_*, s \rangle$. Then, $Ax_* = \vartheta_{S^*}$. Hence, $x_* \in \mathcal{N}(A)$. Thus, $S^\perp \subseteq \mathcal{N}(A)$. In conclusion, $S^\perp = \mathcal{N}(A)$.

By Proposition 7.45, $\mathcal{X}^*/S^\perp$ is a Banach space. Let $\phi : \mathcal{X}^* \rightarrow \mathcal{X}^*/S^\perp$ be the natural homomorphism. By Proposition 7.70, there exists $A_D \in \mathcal{B}(\mathcal{X}^*/S^\perp, S^*)$ such that $A = A_D \circ \phi$, A_D is injective, and $\|A_D\| = \|A\| \leq 1$.

$\forall s_* \in S^*$, by the simple version of Hahn-Banach Theorem, there exists $x_* \in \mathcal{X}^*$ such that $x_*|_S = s_*$ and $\|x_*\| = \|s_*\|$. $\forall s \in S$, we have $\langle s_*, s \rangle = \langle x_*, s \rangle = \langle Ax_*, s \rangle = \langle A_D(\phi(x_*)), s \rangle = \langle A_D([x_*]), s \rangle$. Hence, $s_* = A_D([x_*])$. Then, A_D is surjective. Thus, A_D is bijective.

$\forall [x_*] \in \mathcal{X}^*/S^\perp$, $\|[x_*]\| = \inf_{y_* \in S^\perp} \|x_* - y_*\|$, by Proposition 7.44. By Proposition 7.99, we have

$$\begin{aligned}
\|[x_*]\| &= \min_{y_* \in S^\perp} \|x_* - y_*\| = \sup_{s \in S, \|s\| \leq 1} \operatorname{Re}(\langle x_*, s \rangle) \\
&= \sup_{s \in S, \|s\| \leq 1} \operatorname{Re}(\langle Ax_*, s \rangle) = \sup_{s \in S, \|s\| \leq 1} \operatorname{Re}(\langle A_D(\phi(x_*)), s \rangle) \\
&= \sup_{s \in S, \|s\| \leq 1} \operatorname{Re}(\langle A_D[x_*], s \rangle) \\
&\leq \sup_{s \in S, \|s\| \leq 1} |\operatorname{Re}(\langle A_D[x_*], s \rangle)| \leq \sup_{s \in S, \|s\| \leq 1} |\langle A_D[x_*], s \rangle| \\
&\leq \sup_{s \in S, \|s\| \leq 1} \|A_D[x_*]\| \|s\| \leq \|A_D[x_*]\| = \|Ax_*\| \\
&= \inf_{y_* \in \mathcal{N}(A)} \|A(x_* - y_*)\| = \inf_{y_* \in S^\perp} \|A(x_* - y_*)\| \\
&\leq \inf_{y_* \in S^\perp} \|A\| \|x_* - y_*\| \leq \|[x_*]\|
\end{aligned}$$

where we have applied Proposition 7.72 in the third inequality and Proposition 7.64 in the fifth inequality. Therefore, we have $\|[x_*]\| = \|A_D[x_*]\|$ and A_D is an isometry. Thus, A_D is an isometrical isomorphism between $\mathcal{X}^*/S^\perp$ and S^* . Hence, (i) is established.

(ii) Let $\mathcal{X}$ be reflexive and S be a closed subspace. Let $\psi : S \rightarrow S^{**}$ be the natural mapping. All we need to show is that $\psi(S) = S^{**}$. Fix a $s_{**} \in S^{**}$. Define a functional $\tau : \mathcal{X}^* \rightarrow \mathbb{K}$ by $\tau(x_*) = \langle s_{**}, Ax_* \rangle$, $\forall x_* \in \mathcal{X}^*$. It is easy to show that τ is a linear functional on $\mathcal{X}^*$. $\forall x_* \in \mathcal{X}^*$ with $\|x_*\| \leq 1$, we have $|\tau(x_*)| = |\langle s_{**}, Ax_* \rangle| \leq \|s_{**}\| \|Ax_*\| \leq \|s_{**}\| \|A\| \|x_*\| \leq \|s_{**}\| < +\infty$, where we have applied Propositions 7.72 and 7.64. Hence, $\tau \in \mathcal{X}^{**}$.

Since $\mathcal{X}$ is reflexive, then, by Remark 7.88 and Definition 7.89, $\exists x_0 \in \mathcal{X}$ such that $\tau(x_*) = \langle x_*, x_0 \rangle$, $\forall x_* \in \mathcal{X}^*$. $\forall y_* \in S^\perp$, we have $\langle y_*, x_0 \rangle = \tau(y_*) = \langle s_{**}, Ay_* \rangle = \langle s_{**}, \vartheta_{S^*} \rangle = 0$, where the third equality follows from the fact that $S^\perp = \mathcal{N}(A)$. Hence, $x_0 \in {}^\perp(S^\perp) = S$, by Proposition 7.98. Hence, $\forall s_* \in S^*, \exists x_* \in \mathcal{X}^*$ such that $Ax_* = A_D[x_*] = s_*$ since A_D is an isometrical isomorphism. Then, we have $\langle s_{**}, s_* \rangle =$

$\langle\langle s_{**}, Ax_* \rangle\rangle = \tau(x_*) = \langle\langle x_*, x_0 \rangle\rangle = \langle\langle Ax_*, x_0 \rangle\rangle = \langle\langle s_*, x_0 \rangle\rangle$. This implies that $s_{**} = \psi(x_0)$ and $\psi(S) = S^{**}$. Hence, (ii) is established.

This completes the proof of the proposition. $\square$

7.10 The Open Mapping Theorem

Definition 7.101 Let $\mathcal{X} := (X, \mathcal{O}_X)$ and $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be topological spaces and $A : \mathcal{X} \rightarrow \mathcal{Y}$. A is called an open mapping if $\forall O_X \in \mathcal{O}_X$, we have $A(O_X) \in \mathcal{O}_Y$, that is the image of each open set is open.

Proposition 7.102 Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $S, T \subseteq \mathcal{X}$, and $\alpha \in \mathbb{K}$. Then, the following statements hold.

(i) $\overline{\alpha S} = \alpha \overline{S}$.

(ii) If $\alpha \neq 0$, then $\widetilde{\alpha S} = \alpha \widetilde{S}$.

(iii) If $\alpha \neq 0$, then $(\alpha S)^\circ = \alpha S^\circ$.

(iv) $\overline{S} + \overline{T} \subseteq \overline{S + T}$.

(v) $S^\circ + T^\circ \subseteq (S + T)^\circ$.

Proof (i) We will distinguish three exhaustive and mutually exclusive cases: Case 1: $\alpha = 0$ and $S = \emptyset$; Case 2: $\alpha = 0$ and $S \neq \emptyset$; Case 3: $\alpha \neq 0$. Case 1: $\alpha = 0$ and $S = \emptyset$. Then, $\overline{S} = \emptyset$ and $\overline{\alpha S} = \emptyset = \alpha \overline{S}$. Case 2: $\alpha = 0$ and $S \neq \emptyset$. Then, $\overline{S} \neq \emptyset$ and $\overline{\alpha S} = \{\emptyset\} = \alpha \overline{S}$.

Case 3: $\alpha \neq 0$. $\forall x \in \overline{\alpha S}$, by Proposition 4.13, $\exists (x_n)_{n=1}^\infty \subseteq \alpha S$ such that $\lim_{n \in \mathbb{N}} x_n = x$. Note that $(\alpha^{-1} x_n)_{n=1}^\infty \subseteq S$ and $\lim_{n \in \mathbb{N}} \alpha^{-1} x_n = \alpha^{-1} x$ by Propositions 3.66 and 7.23. By Proposition 4.13, $\alpha^{-1} x \in \overline{S}$. Then, $x \in \alpha \overline{S}$. This shows that $\overline{\alpha S} \subseteq \alpha \overline{S}$.

On the other hand, $\forall x \in \alpha \overline{S}$, then $\alpha^{-1} x \in \overline{S}$. By Proposition 4.13, $\exists (x_n)_{n=1}^\infty \subseteq S$ such that $\lim_{n \in \mathbb{N}} x_n = \alpha^{-1} x$. Note that $(\alpha x_n)_{n=1}^\infty \subseteq \alpha S$ and $\lim_{n \in \mathbb{N}} \alpha x_n = x$ by Propositions 3.66 and 7.23. By Proposition 4.13, $x \in \overline{\alpha S}$. Then, $\alpha \overline{S} \subseteq \overline{\alpha S}$. Hence, $\overline{\alpha S} = \alpha \overline{S}$.

(ii) $\forall x \in \widetilde{\alpha S}$ if, and only if, $\alpha^{-1} x \in \widetilde{S}$ if, and only if, $x \in \alpha \widetilde{S}$. Hence, $\widetilde{\alpha S} = \alpha \widetilde{S}$.

(iii) $\forall x \in (\alpha S)^\circ$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(x, \delta) \subseteq \alpha S$. Then, $\mathcal{B}(\alpha^{-1} x, \delta/|\alpha|) = \alpha^{-1} \mathcal{B}(x, \delta) \subseteq S$. Hence, $\alpha^{-1} x \in S^\circ$. Then, $x \in \alpha S^\circ$. This shows that $(\alpha S)^\circ \subseteq \alpha S^\circ$.

On the other hand, $\forall x \in \alpha S^\circ$, we have $\alpha^{-1} x \in S^\circ$. $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(\alpha^{-1} x, \delta) \subseteq S$. Then, $\mathcal{B}(x, |\alpha| \delta) = \alpha \mathcal{B}(\alpha^{-1} x, \delta) \subseteq \alpha S$. Hence, $x \in (\alpha S)^\circ$. This shows that $\alpha S^\circ \subseteq (\alpha S)^\circ$. Hence, we have $(\alpha S)^\circ = \alpha S^\circ$.

(iv) $\forall \bar{x} \in \overline{S} + \overline{T}$, $\exists \bar{s} \in \overline{S}$ and $\exists \bar{t} \in \overline{T}$ such that $\bar{x} = \bar{s} + \bar{t}$. By Proposition 4.13, $\exists (s_n)_{n=1}^\infty \subseteq S$ and $\exists (t_n)_{n=1}^\infty \subseteq T$ such that $\lim_{n \in \mathbb{N}} s_n = \bar{s}$ and

$\lim_{n \in \mathbb{N}} t_n = \bar{t}$. Then, $(s_n + t_n)_{n=1}^\infty \subseteq S + T$ and $\lim_{n \in \mathbb{N}} (s_n + t_n) = \bar{s} + \bar{t} = \bar{x}$, by Propositions 7.23, 3.66, and 3.67. By Proposition 4.13, $\bar{x} \in \overline{S + T}$. Hence, $\overline{S} + \overline{T} \subseteq \overline{S + T}$.

(v) $\forall x \in S^\circ + T^\circ$, $\exists s_0 \in S^\circ$ and $\exists t_0 \in T^\circ$ such that $x = s_0 + t_0$. Then, $\exists r_s, r_t \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(s_0, r_s) \subseteq S$ and $\mathcal{B}(t_0, r_t) \subseteq T$. Thus, we have $\mathcal{B}(x, r_s + r_t) = \mathcal{B}(s_0, r_s) + \mathcal{B}(t_0, r_t) \subseteq S + T$ and $x \in (S + T)^\circ$. Hence, $S^\circ + T^\circ \subseteq (S + T)^\circ$.

This completes the proof of the proposition. $\square$

Theorem 7.103 (Open Mapping Theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over the field $\mathbb{K}$ and $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be surjective. Then, A is an open mapping. Furthermore, if A is injective, then $A_{\text{inv}} \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$.*

Proof We need the following claim.

Claim 7.103.1 *The image of the unit ball in $\mathcal{X}$ under A contains an open ball centered at the origin in $\mathcal{Y}$, that is $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, \delta) \subseteq A(\mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 1))$.*

Proof of claim: Let $S_n := \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 2^{-n})$, $\forall n \in \mathbb{Z}_+$. Since A is linear and surjective and $\bigcup_{k=1}^\infty kS_1 = \mathcal{X}$, then, by Proposition 2.5, we have $\mathcal{Y} = \bigcup_{k=1}^\infty A(kS_1) = \bigcup_{k=1}^\infty kA(S_1)$. Since $\mathcal{Y}$ is a complete metric space, then, by Baire Category Theorem, $\mathcal{Y}$ is second category everywhere. Then, $\mathcal{Y}$ is not of first category, that is $\mathcal{Y}$ is not countable union of nowhere dense sets. Then, by Proposition 7.102, $A(S_1) \subseteq \mathcal{Y}$ is not nowhere dense. Then, $\overline{A(S_1)} \subseteq \mathcal{Y}$ is not nowhere dense. By Proposition 3.40, $\exists \bar{y} \in \mathcal{Y}$ and $\exists \bar{\delta} \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{Y}}(\bar{y}, \bar{\delta}) \subseteq \overline{A(S_1)}$. Then, we have $\mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, \bar{\delta}) = \mathcal{B}_{\mathcal{Y}}(\bar{y}, \bar{\delta}) - \bar{y} \subseteq \overline{A(S_1)} - \bar{y} \subseteq \overline{A(S_1) - \bar{y}}$. Note that, by Proposition 7.102 and the linearity of A , $-A(S_1) = \overline{-A(S_1)} = \overline{A(-S_1)} = \overline{A(S_1)}$. Then, again by Proposition 7.102 and the linearity of A , $\mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, \bar{\delta}) \subseteq \overline{A(S_1)} + A(S_1) \subseteq \overline{A(S_1) + A(S_1)} = \overline{A(S_1 + S_1)} = \overline{A(S_0)}$. Thus, $\overline{A(S_0)} \subseteq \mathcal{Y}$ contains a ball center at the origin with radius $\bar{\delta}$. By Proposition 7.102 and linearity of A , $\mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, 2^{-n}\bar{\delta}) = 2^{-n}\mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, \bar{\delta}) \subseteq 2^{-n}\overline{A(S_0)} = \overline{A(2^{-n}S_0)} = \overline{A(S_n)} \subseteq \mathcal{Y}$, $\forall n \in \mathbb{Z}_+$.

Now, we proceed to show that $\mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, \bar{\delta}/2) \subseteq A(S_0)$. Fix an arbitrary vector $y \in \mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, \bar{\delta}/2)$. Then, $y \in \overline{A(S_1)}$ and $\exists x_1 \in S_1$, such that $\|y - Ax_1\| < 2^{-2}\bar{\delta}$. This implies that $y - Ax_1 \in \mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, 2^{-2}\bar{\delta}) \subseteq \overline{A(S_2)}$. Recursively, $\forall n \in \mathbb{N}$ with $n \geq 2$, $y - \sum_{k=1}^{n-1} Ax_k \in \overline{A(S_n)}$. Then, $\exists x_n \in S_n$ such that $\|y - \sum_{k=1}^n Ax_k\| < 2^{-n-1}\bar{\delta}$. This implies that $y - \sum_{k=1}^n Ax_k \in \mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, 2^{-n-1}\bar{\delta}) \subseteq \overline{A(S_{n+1})}$. Since $x_n \in S_n$ and $\|x_n\| < 2^{-n}$, $\forall n \in \mathbb{N}$, then $\sum_{k=1}^\infty \|x_k\| < 1$. By Proposition 7.27, we have $\sum_{k=1}^\infty x_k = x \in \mathcal{Y}$. It is easy to show that $x \in S_0$. Then, $y = \lim_{n \in \mathbb{N}} \sum_{k=1}^n Ax_k = \lim_{n \in \mathbb{N}} A(\sum_{k=1}^n x_k) = Ax$, where we have applied Proposition 3.66. Thus, $y \in A(S_0)$ and $\mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, \bar{\delta}/2) \subseteq A(S_0)$.

This completes the proof of the claim. $\square$

Fix any open set $O \subseteq \mathcal{X}$ and any $y \in A(O)$. Let $x \in O$ be such that $Ax = y$. Then, there exists $r \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{X}}(x, r) \subseteq O$. By Claim 7.103.1, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{Y}}(\vartheta_y, \delta) \subseteq A(\mathcal{B}_{\mathcal{X}}(\vartheta_x, r))$. By the linearity of A , we have $\mathcal{B}_{\mathcal{Y}}(y, \delta) = y + \mathcal{B}_{\mathcal{Y}}(\vartheta_y, \delta) \subseteq Ax + A(\mathcal{B}_{\mathcal{X}}(\vartheta_x, r)) = A(\mathcal{B}_{\mathcal{X}}(x, r)) \subseteq A(O)$. By the arbitrariness of y , $A(O) \subseteq \mathcal{Y}$ is open.

If, in addition, A is injective, then A is bijective and A_{inv} exists. Since A is open mapping, then A_{inv} is continuous. It is obvious that $A_{\text{inv}} : \mathcal{Y} \rightarrow \mathcal{X}$ is linear since A is linear. Therefore, $A_{\text{inv}} \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$.

This completes the proof of the theorem. $\square$

Proposition 7.104 *Let $\mathcal{X}$ be a vector space over $\mathbb{K}$ and $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on $\mathcal{X}$ such that $\mathcal{X}_1 := (\mathcal{X}, \mathbb{K}, \|\cdot\|_1)$ and $\mathcal{X}_2 := (\mathcal{X}, \mathbb{K}, \|\cdot\|_2)$ are Banach spaces. If $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|x\|_2 \leq M\|x\|_1, \forall x \in \mathcal{X}$. Then, the two norms are equivalent.*

Proof Consider the mapping $A := \text{id}_{\mathcal{X}} : \mathcal{X}_1 \rightarrow \mathcal{X}_2$. Clearly A is a linear bijective function. By the assumption of the proposition, $A \in \mathcal{B}(\mathcal{X}_1, \mathcal{X}_2)$. By Open Mapping Theorem, $A_{\text{inv}} \in \mathcal{B}(\mathcal{X}_2, \mathcal{X}_1)$. Then, $\exists \bar{M} \in [0, \infty) \subset \mathbb{R}$ such that $\|x\|_1 = \|Ax\|_1 \leq \bar{M}\|x\|_2, \forall x \in \mathcal{X}$. Now, take $K = \max\{M, \bar{M}\} + 1 \in (0, \infty) \subset \mathbb{R}$. Then, we have $\|x\|_1/K \leq \|x\|_2 \leq K\|x\|_1, \forall x \in \mathcal{X}$. Hence, the two norms are equivalent. This completes the proof of the proposition. $\square$

Theorem 7.105 (Closed Graph Theorem) *Let $\mathcal{X} := (\mathcal{X}, \mathbb{K}, \|\cdot\|)$ and $\mathcal{Y}$ be Banach spaces over the field $\mathbb{K}$ and $A : \mathcal{X} \rightarrow \mathcal{Y}$ be a linear operator. Assume that A satisfies that $\forall (x_n)_{n=1}^{\infty} \subseteq \mathcal{X}$ with $\lim_{n \in \mathbb{N}} x_n = x_0 \in \mathcal{X}$ and $\lim_{n \in \mathbb{N}} Ax_n = y_0 \in \mathcal{Y}$, we have $y_0 = Ax_0$. Then, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.*

Proof Define a functional $\|\cdot\|_1 : \mathcal{X} \rightarrow \mathbb{R}$ by $\|x\|_1 := \|x\| + \|Ax\|, \forall x \in \mathcal{X}$. It is easy to show that $\|\cdot\|_1$ defines a norm on $\mathcal{X}$. Let $\mathcal{X}_1 := (\mathcal{X}, \mathbb{K}, \|\cdot\|_1)$ be the normed linear space. We will show that $\mathcal{X}_1$ is complete. Fix any Cauchy sequence $(x_n)_{n=1}^{\infty} \subseteq \mathcal{X}_1$. It is easy to see that $(x_n)_{n=1}^{\infty} \subseteq \mathcal{X}$ is a Cauchy sequence; and that $(Ax_n)_{n=1}^{\infty} \subseteq \mathcal{Y}$ is a Cauchy sequence. By the completeness of $\mathcal{X}$ and $\mathcal{Y}$, there exists $x_0 \in \mathcal{X}$ and $y_0 \in \mathcal{Y}$ such that $\lim_{n \in \mathbb{N}} \|x_n - x_0\| = 0$ and $\lim_{n \in \mathbb{N}} \|Ax_n - y_0\| = 0$. By the assumption of the theorem, we have $y_0 = Ax_0$. Hence, we have $\lim_{n \in \mathbb{N}} \|x_n - x_0\|_1 = 0$. Then, $\lim_{n \in \mathbb{N}} x_n = x_0$ in $\mathcal{X}_1$. Thus, $\mathcal{X}_1$ is a Banach space. Clearly, $\forall x \in \mathcal{X}$, we have $\|x\| \leq \|x\|_1$. By Proposition 7.104, $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent. Then, $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|x\|_1 \leq M\|x\|, \forall x \in \mathcal{X}$. Then, we have $\|Ax\| \leq M\|x\|, \forall x \in \mathcal{X}$. Hence, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. This completes the proof of the theorem. $\square$

Recall that the graph of a function $f : X \rightarrow Y$ is the set $\{(x, y) \in X \times Y \mid x \in X, y = f(x)\}$. Then, we have the following alternative statement of the Closed Graph Theorem.

Theorem 7.106 (Closed Graph Theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over the field $\mathbb{K}$ and $A : \mathcal{X} \rightarrow \mathcal{Y}$ be a linear operator. Then, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ if, and only if, the graph of A is a closed set in $\mathcal{X} \times \mathcal{Y}$.*

Proof “Sufficiency” Let the graph of A be $\text{graph}(A)$. Fix any sequence $(x_n)_{n=1}^\infty \subseteq \mathcal{X}$ with $\lim_{n \in \mathbb{N}} x_n = x_0 \in \mathcal{X}$ and $\lim_{n \in \mathbb{N}} Ax_n = y_0 \in \mathcal{Y}$. Note that $((x_n, Ax_n))_{n=1}^\infty \subseteq \text{graph}(A) \subseteq \mathcal{X} \times \mathcal{Y}$ and $\lim_{n \in \mathbb{N}} (x_n, Ax_n) = (x_0, y_0)$, by Proposition 3.67. Then, by Proposition 4.13, we have $(x_0, y_0) \in \overline{\text{graph}(A)} = \text{graph}(A)$, where the equality follows from Proposition 3.3. Hence, we have $y_0 = Ax_0$. By the Closed Graph Theorem, Theorem 7.105, we have $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.

“Necessity” $\forall (x_0, y_0) \in \overline{\text{graph}(A)}$, by Proposition 4.13, there exists $((x_n, Ax_n))_{n=1}^\infty \subseteq \text{graph}(A)$ such that $\lim_{n \in \mathbb{N}} (x_n, Ax_n) = (x_0, y_0)$. By Proposition 3.67, we have $\lim_{n \in \mathbb{N}} x_n = x_0$ and $\lim_{n \in \mathbb{N}} Ax_n = y_0$. By Proposition 3.66, we have $y_0 = \lim_{n \in \mathbb{N}} Ax_n = Ax_0$. Hence, $(x_0, y_0) \in \text{graph}(A)$. By Proposition 3.3, $\text{graph}(A)$ is closed in $\mathcal{X} \times \mathcal{Y}$.

This completes the proof of the theorem. $\square$

Proposition 7.107 *Let $\mathcal{X}$ be a Banach space over the field $\mathbb{K}$, $\mathcal{Y}$ be a normed linear space over the same field, and $\mathcal{F} \subseteq \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Assume that $\forall x \in \mathcal{X}, \exists M_x \in [0, \infty) \subset \mathbb{R}$ such that $\|Tx\| \leq M_x, \forall T \in \mathcal{F}$. Then, $\exists M \in [0, \infty) \subset \mathbb{R}$, such that $\|T\| \leq M, \forall T \in \mathcal{F}$.*

Proof By Baire Category Theorem, $\mathcal{X}$ is second category everywhere. $\forall T \in \mathcal{F}$, let $f : \mathcal{X} \rightarrow \mathbb{R}$ be given by $f(x) = \|Tx\|, \forall x \in \mathcal{X}$. By Propositions 7.21 and 3.12, f is a continuous real-valued function. By Uniform Boundedness Principle, there exist an open set $O \subseteq \mathcal{X}$ with $O \neq \emptyset$ and $\bar{M} \in [0, \infty) \subset \mathbb{R}$ such that $\|Tx\| \leq \bar{M}, \forall T \in \mathcal{F}, \forall x \in O$. Since O is nonempty and open, then $\exists \mathcal{B}_{\mathcal{X}}(x_0, r) \subseteq O$ for some $x_0 \in \mathcal{X}$ and some $r \in (0, \infty) \subset \mathbb{R}$. $\forall x \in \mathcal{X}$ with $\|x\| < r, x + x_0 \in \mathcal{B}_{\mathcal{X}}(x_0, r) \subseteq O$ and $\forall T \in \mathcal{F}$, we have

$$\|Tx\| = \|T(x + x_0) - Tx_0\| \leq \|T(x + x_0)\| + \|Tx_0\| \leq \bar{M} + M_{x_0}$$

Hence, $\forall T \in \mathcal{F}$, we have, $\forall \epsilon \in (0, r) \subset \mathbb{R}$,

$$\|T\| = \sup_{x \in \mathcal{X}, \|x\| \leq 1} \|Tx\| = \sup_{x \in \mathcal{X}, \|x\| \leq 1} (r - \epsilon)^{-1} \|T((r - \epsilon)x)\| \leq \frac{\bar{M} + M_{x_0}}{r - \epsilon}$$

By the arbitrariness of ϵ , we have $\|T\| \leq (\bar{M} + M_{x_0})/r =: M < +\infty$. This completes the proof of the proposition. $\square$

Proposition 7.108 *Let $\mathcal{X}$ be a Banach space over the field $\mathbb{K}$, $\mathcal{Y}$ be a normed linear space over the same field, and $(T_n)_{n=1}^\infty \subseteq \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Assume that $\forall x \in \mathcal{X}, \lim_{n \in \mathbb{N}} T_n x = T(x) \in \mathcal{Y}$. Then, $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.*

Proof $\forall x_1, x_2 \in \mathcal{X}, \forall \alpha, \beta \in \mathbb{K}$, we have

$$T(\alpha x_1 + \beta x_2) = \lim_{n \in \mathbb{N}} T_n(\alpha x_1 + \beta x_2) = \lim_{n \in \mathbb{N}} (\alpha T_n x_1 + \beta T_n x_2) = \alpha T x_1 + \beta T x_2$$

by Propositions 7.23 and 3.66. Hence, T is linear. $\forall x \in \mathcal{X}$, by Propositions 7.21 and 3.66, $\lim_{n \in \mathbb{N}} \|T_n x\| = \|Tx\| < +\infty$. Then, $\exists M_x \in [0, \infty) \subset \mathbb{R}$ such that $\|T_n x\| \leq M_x, \forall n \in \mathbb{N}$. By Proposition 7.107, $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|T_n\| \leq M, \forall n \in \mathbb{N}$. $\forall x \in \mathcal{X}$ with $\|x\| \leq 1$, by Proposition 7.64, $\|Tx\| = \lim_{n \in \mathbb{N}} \|T_n x\| \leq \limsup_{n \in \mathbb{N}} \|T_n\| \|x\| \leq M < +\infty$. Hence, $\|T\| \leq M < +\infty$. Therefore, $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. This completes the proof of the proposition. $\square$

7.11 The Adjoint of Linear Operators

Proposition 7.109 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$ and $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. The adjoint operator of A is $A' : \mathcal{Y}^* \rightarrow \mathcal{X}^*$ defined by*

$$\langle\langle A'(y_*), x \rangle\rangle = \langle\langle y_*, Ax \rangle\rangle; \quad \forall x \in \mathcal{X}, \forall y_* \in \mathcal{Y}^*$$

Then, $A' \in \mathcal{B}(\mathcal{Y}^*, \mathcal{X}^*)$ with $\|A'\| = \|A\|$.

Proof First, we will show that A' is well defined. $\forall y_* \in \mathcal{Y}^*, \forall x \in \mathcal{X}$. $\langle\langle y_*, Ax \rangle\rangle \in \mathbb{K}$. Hence, $f : \mathcal{X} \rightarrow \mathbb{K}$ defined by $f(x) = \langle\langle y_*, Ax \rangle\rangle, \forall x \in \mathcal{X}$, is a functional on $\mathcal{X}$. By the linearity of A and y_* , f is a linear functional. $\forall x \in \mathcal{X}$, we have, by Proposition 7.64

$$|f(x)| \leq \|y_*\| \|Ax\| \leq \|y_*\| \|A\| \|x\|$$

Hence, f is a bounded linear functional with $\|f\| \leq \|A\| \|y_*\|$. The above shows that $A'(y_*) = f \in \mathcal{X}^*$. Hence, A' is well-defined.

It is straightforward to show that A' is a linear operator. By the fact that $\|A'y_*\| = \|f\| \leq \|A\| \|y_*\|, \forall y_* \in \mathcal{Y}^*$, we have $A' \in \mathcal{B}(\mathcal{Y}^*, \mathcal{X}^*)$ and $\|A'\| \leq \|A\|$.

On the other hand, $\forall x \in \mathcal{X}$, we have either $Ax = \vartheta_{\mathcal{Y}}$, then $\|Ax\| = 0 \leq \|A'\| \|x\|$; or $Ax \neq \vartheta_{\mathcal{Y}}$, then, by Proposition 7.85, $\exists y_* \in \mathcal{Y}^*$ with $\|y_*\| = 1$ such that $\|Ax\| = \langle\langle y_*, Ax \rangle\rangle$, which implies that $\|Ax\| = \langle\langle A'y_*, x \rangle\rangle \leq \|A'y_*\| \|x\| \leq \|A'\| \|y_*\| \|x\| = \|A'\| \|x\|$. Hence, we must have $\|Ax\| \leq \|A'\| \|x\|$. This implies that $\|A\| \leq \|A'\|$.

Then, $\|A\| = \|A'\|$. This completes the proof of the proposition. $\square$

Proposition 7.110 *Let $\mathcal{X}, \mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over the field $\mathbb{K}$. Then, the following statements hold.*

(i) $\text{id}'_{\mathcal{X}} = \text{id}_{\mathcal{X}^*}$.

(ii) If $A_1, A_2 \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, then $(A_1 + A_2)' = A'_1 + A'_2$.

(iii) If $A \in B(\mathcal{X}, \mathcal{Y})$ and $\alpha \in \mathbb{K}$, then $(\alpha A)' = \alpha A'$.

(iv) If $A_1 \in B(\mathcal{X}, \mathcal{Y})$ and $A_2 \in B(\mathcal{Y}, \mathcal{Z})$, then $(A_2 A_1)' = A_1' A_2'$.

(v) If $A \in B(\mathcal{X}, \mathcal{Y})$ and A has a bounded inverse, then $(A^{-1})' = (A')^{-1}$.

Proof (i) $\forall x_* \in \mathcal{X}^*, \forall x \in \mathcal{X}$, $\langle \langle \text{id}_{\mathcal{X}}^* x_*, x \rangle \rangle = \langle \langle x_*, \text{id}_{\mathcal{X}} x \rangle \rangle = \langle \langle x_*, x \rangle \rangle = \langle \langle \text{id}_{\mathcal{X}^*} x_*, x \rangle \rangle$. Hence, the result follows.

(ii) $\forall y_* \in \mathcal{Y}^*, \forall x \in \mathcal{X}$, we have $\langle \langle (A_1 + A_2)' y_*, x \rangle \rangle = \langle \langle y_*, (A_1 + A_2)x \rangle \rangle = \langle \langle y_*, A_1 x \rangle \rangle + \langle \langle y_*, A_2 x \rangle \rangle = \langle \langle A_1' y_*, x \rangle \rangle + \langle \langle A_2' y_*, x \rangle \rangle = \langle \langle A_1' y_* + A_2' y_*, x \rangle \rangle = \langle \langle (A_1' + A_2') y_*, x \rangle \rangle$. Hence, the result follows.

(iii) $\forall y_* \in \mathcal{Y}^*, \forall x \in \mathcal{X}$, we have $\langle \langle (\alpha A)' y_*, x \rangle \rangle = \langle \langle y_*, (\alpha A)x \rangle \rangle = \alpha \langle \langle y_*, Ax \rangle \rangle = \alpha \langle \langle A' y_*, x \rangle \rangle = \langle \langle \alpha(A' y_*), x \rangle \rangle = \langle \langle (\alpha A') y_*, x \rangle \rangle$. Hence, the result follows.

(iv) $\forall z_* \in \mathcal{Z}^*, \forall x \in \mathcal{X}$, we have $\langle \langle (A_2 A_1)' z_*, x \rangle \rangle = \langle \langle z_*, (A_2 A_1)x \rangle \rangle = \langle \langle z_*, A_2(A_1 x) \rangle \rangle = \langle \langle A_2' z_*, A_1 x \rangle \rangle = \langle \langle A_1'(A_2' z_*), x \rangle \rangle = \langle \langle (A_1' A_2') z_*, x \rangle \rangle$. Hence, the result follows.

(v) By (i) and (iv), we have $(A^{-1})' A' = (A A^{-1})' = \text{id}_{\mathcal{Y}}' = \text{id}_{\mathcal{Y}^*}$ and $A' (A^{-1})' = (A^{-1} A)' = \text{id}_{\mathcal{X}}' = \text{id}_{\mathcal{X}^*}$. By Proposition 2.4, we have $(A^{-1})' = (A')^{-1}$.

This completes the proof of the proposition. $\square$

Proposition 7.111 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$, $A \in B(\mathcal{X}, \mathcal{Y})$, $\phi_{\mathcal{X}} : \mathcal{X} \rightarrow \mathcal{X}^{**}$ and $\phi_{\mathcal{Y}} : \mathcal{Y} \rightarrow \mathcal{Y}^{**}$ be the natural mappings on $\mathcal{X}$ and $\mathcal{Y}$, respectively, and $A'' : \mathcal{X}^{**} \rightarrow \mathcal{Y}^{**}$ be the adjoint of the adjoint of A . Then, we have $A'' \circ \phi_{\mathcal{X}} = \phi_{\mathcal{Y}} \circ A$.

Proof $\forall x \in \mathcal{X}, \forall y_* \in \mathcal{Y}^*$, we have

$$\begin{aligned} \langle \langle A''(\phi_{\mathcal{X}}(x)), y_* \rangle \rangle &= \langle \langle \phi_{\mathcal{X}}(x), A' y_* \rangle \rangle = \langle \langle A' y_*, x \rangle \rangle = \langle \langle y_*, Ax \rangle \rangle \\ &= \langle \langle \phi_{\mathcal{Y}}(Ax), y_* \rangle \rangle \end{aligned}$$

Hence, the desired result follows. This completes the proof of the proposition. $\square$

By Proposition 7.111 and Remark 7.88, $\forall A, B \in B(\mathcal{X}, \mathcal{Y})$, we have $A' = B'$ if, and only if, $A = B$.

Proposition 7.112 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$ and $A \in B(\mathcal{X}, \mathcal{Y})$. Then,

$$(\mathcal{R}(A))^{\perp} = \mathcal{N}(A'); \quad {}^{\perp}(\mathcal{R}(A')) = \mathcal{N}(A)$$

Proof Fix a vector $y_* \in \mathcal{N}(A')$. $\forall y \in \mathcal{R}(A)$, $\exists x \in \mathcal{X}$ such that $y = Ax$. Then, $\langle \langle y_*, y \rangle \rangle = \langle \langle y_*, Ax \rangle \rangle = \langle \langle A' y_*, x \rangle \rangle = \langle \langle \vartheta_{\mathcal{X}^*}, x \rangle \rangle = 0$. Hence, $y_* \in (\mathcal{R}(A))^{\perp}$. This shows that $\mathcal{N}(A') \subseteq (\mathcal{R}(A))^{\perp}$.

On the other hand, fix a vector $y_* \in (\mathcal{R}(A))^{\perp}$. $\forall x \in \mathcal{X}$, we have $\langle \langle A' y_*, x \rangle \rangle = \langle \langle y_*, Ax \rangle \rangle = 0$. Hence, we have $A' y_* = \vartheta_{\mathcal{X}^*}$. This shows that $(\mathcal{R}(A))^{\perp} \subseteq \mathcal{N}(A')$.

Hence, we have $(\mathcal{R}(A))^\perp = \mathcal{N}(A')$.

Fix a vector $x \in \mathcal{N}(A)$. $\forall x_* \in \mathcal{R}(A')$, $\exists y_* \in \mathcal{Y}^*$ such that $x_* = A'y_*$. Then, $\langle\langle x_*, x \rangle\rangle = \langle\langle y_*, Ax \rangle\rangle = \langle\langle y_*, \vartheta y \rangle\rangle = 0$. Hence, $x \in {}^\perp(\mathcal{R}(A'))$. This shows that $\mathcal{N}(A) \subseteq {}^\perp(\mathcal{R}(A'))$.

On the other hand, fix a vector $x \in {}^\perp(\mathcal{R}(A'))$. $\forall y_* \in \mathcal{Y}^*$, we have $0 = \langle\langle A'y_*, x \rangle\rangle = \langle\langle y_*, Ax \rangle\rangle$. Hence, by Proposition 7.85, we have $Ax = \vartheta y$ and $x \in \mathcal{N}(A)$. This shows that ${}^\perp(\mathcal{R}(A')) \subseteq \mathcal{N}(A)$.

Hence, we have ${}^\perp(\mathcal{R}(A')) = \mathcal{N}(A)$. This completes the proof of the proposition. $\square$

The dual version of the above proposition is deeper, which requires both Open Mapping Theorem and Hahn-Banach Theorem. Towards this end, we need the following result.

Proposition 7.113 *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over the field $\mathbb{K}$ and $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Assume that $\mathcal{R}(A) \subseteq \mathcal{Y}$ is closed. Then, there exists $K \in [0, \infty) \subset \mathbb{R}$ such that, $\forall y \in \mathcal{R}(A)$, there exists $x \in \mathcal{X}$ such that $y = Ax$ and $\|x\| \leq K\|y\|$.*

Proof Since A is continuous, then $\mathcal{N}(A) \subseteq \mathcal{X}$ is closed. By Proposition 7.45, $\mathcal{X}/\mathcal{N}(A)$ is a Banach space. Let $\phi : \mathcal{X} \rightarrow \mathcal{X}/\mathcal{N}(A)$ be the natural homomorphism, which is a bounded linear function by Proposition 7.69. By Proposition 7.70, there exists a bounded linear function $A_D : \mathcal{X}/\mathcal{N}(A) \rightarrow \mathcal{R}(A)$ such that $A = A_D \circ \phi$, $\|A\| = \|A_D\|$ and A_D is injective. By Proposition 4.39, $\mathcal{R}(A)$ is complete. Then, by Proposition 7.13, $\mathcal{R}(A)$ is a Banach space. The mapping A_D is surjective to $\mathcal{R}(A)$. Hence, $A_D : \mathcal{X}/\mathcal{N}(A) \rightarrow \mathcal{R}(A)$ is a bijective bounded linear operator. By Open Mapping Theorem, $A_D^{-1} \in \mathcal{B}(\mathcal{R}(A), \mathcal{X}/\mathcal{N}(A))$. $\forall y \in \mathcal{R}(A)$, Let $[x] := A_D^{-1}y \in \mathcal{X}/\mathcal{N}(A)$. Then, by Proposition 7.64, $\|[x]\| \leq \|A_D^{-1}\|\|y\|$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\|[x]\| = 0$; Case 2: $\|[x]\| > 0$. Case 1: $\|[x]\| = 0$. Then, $[x] = [\vartheta x]$. Take $x = \vartheta x$, we have $y = A_D[x] = Ax = \vartheta y$ and $\|x\| = 0 = 2\|A_D^{-1}\|\|y\|$. Case 2: $\|[x]\| > 0$. Note that $\|[x]\| = \inf_{x \in [x]} \|x\|$. Then, $\exists x \in [x]$ such that $\|x\| \leq 2\|[x]\|$. Then, $y = A_D[x] = Ax$ and $\|x\| \leq 2\|[x]\| \leq 2\|A_D^{-1}\|\|y\|$. Hence, the desired result holds in both cases with $K = 2\|A_D^{-1}\|$. This completes the proof of the proposition. $\square$

Proposition 7.114 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$ and $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Assume that $\mathcal{R}(A)$ is closed in $\mathcal{Y}$. Then,*

$$\mathcal{R}(A) = {}^\perp(\mathcal{N}(A'))$$

If, in addition, $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, then

$$\mathcal{R}(A') = (\mathcal{N}(A))^\perp$$

Proof Fix $y \in \mathcal{R}(A)$. Then $\exists x \in \mathcal{X}$ such that $y = Ax$. $\forall y_* \in \mathcal{N}(A')$, we have $\langle\langle y_*, y \rangle\rangle = \langle\langle A'y_*, x \rangle\rangle = \langle\langle \vartheta_{x^*}, x \rangle\rangle = 0$. Hence, we have $y \in {}^\perp(\mathcal{N}(A'))$ and $\mathcal{R}(A) \subseteq {}^\perp(\mathcal{N}(A'))$.

On the other hand, fix $y \in {}^\perp(\mathcal{N}(A'))$. By Proposition 7.97, $\delta := \inf_{k \in \mathcal{R}(A)} \|y - k\| = \max_{y_* \in (\mathcal{R}(A))^\perp, \|y_*\| \leq 1} \operatorname{Re}(\langle\langle y_*, y \rangle\rangle)$, where the maximum is achieved at some $y_{*0} \in (\mathcal{R}(A))^\perp$ with $\|y_{*0}\| \leq 1$. By Proposition 7.112, we have $(\mathcal{R}(A))^\perp = \mathcal{N}(A') \ni y_{*0}$. Then, $\delta = \operatorname{Re}(\langle\langle y_{*0}, y \rangle\rangle) = 0$. By Proposition 4.10, $y \in \mathcal{R}(A)$ since $\mathcal{R}(A) \subseteq \mathcal{Y}$ is closed. This shows that ${}^\perp(\mathcal{N}(A')) \subseteq \mathcal{R}(A)$.

Hence, we have $\mathcal{R}(A) = {}^\perp(\mathcal{N}(A'))$.

Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces. Fix an $x_* \in \mathcal{R}(A')$, then $\exists y_* \in \mathcal{Y}^*$ such that $x_* = A'y_*$. $\forall x \in \mathcal{N}(A)$, we have $\langle\langle x_*, x \rangle\rangle = \langle\langle A'y_*, x \rangle\rangle = \langle\langle y_*, Ax \rangle\rangle = \langle\langle y_*, \vartheta_{\mathcal{Y}} \rangle\rangle = 0$. Then, $x_* \in (\mathcal{N}(A))^\perp$. This shows that $\mathcal{R}(A') \subseteq (\mathcal{N}(A))^\perp$.

On the other hand, fix an $x_* \in (\mathcal{N}(A))^\perp$. Let $K \in [0, \infty) \subset \mathbb{R}$ be the constant described in Proposition 7.113. $\forall y \in \mathcal{R}(A)$, $\forall x \in \mathcal{X}$ with $Ax = y$, $\langle\langle x_*, x \rangle\rangle$ assumes a constant value that is dependent on y only. So, we may define a functional $f : \mathcal{R}(A) \rightarrow \mathbb{K}$ by $f(y) = \langle\langle x_*, x \rangle\rangle$, $\forall y \in \mathcal{R}(A)$ and $\forall x \in \mathcal{X}$ with $Ax = y$. Clearly, f is a linear functional. By Proposition 7.113, $\exists x_0 \in \mathcal{X}$ with $Ax_0 = y$, such that $\|x_0\| \leq K\|y\|$. Then, by Proposition 7.64, $|f(y)| \leq \|x_*\|\|x\| \leq \|x_*\|K\|y\|$. Hence, $\|f\|_{\mathcal{R}(A)} := \sup_{y \in \mathcal{R}(A), \|y\| \leq 1} |f(y)| \leq K\|x_*\| < +\infty$. By the simple version of Hahn-Banach Theorem 7.83, $\exists y_* \in \mathcal{Y}^*$ such that $y_*|_{\mathcal{R}(A)} = f$ and $\|y_*\| = \|f\|_{\mathcal{R}(A)}$. $\forall x \in \mathcal{X}$, we have

$$\langle\langle A'y_*, x \rangle\rangle = \langle\langle y_*, Ax \rangle\rangle = f(Ax) = \langle\langle x_*, x \rangle\rangle$$

This implies that $A'y_* = x_* \in \mathcal{R}(A')$. Hence, we have $(\mathcal{N}(A))^\perp \subseteq \mathcal{R}(A')$.

Therefore, we have $\mathcal{R}(A') = (\mathcal{N}(A))^\perp$. This completes the proof of the proposition. $\square$

7.12 Weak Topology

Definition 7.115 Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$. The weak topology on $\mathcal{X}$, denoted by $\mathcal{O}_{\text{weak}}(\mathcal{X}) \subseteq \mathcal{X}_2$, is the weak topology generated by $\mathcal{X}^*$, that is the weakest topology on $\mathcal{X}$ such that $x_* : \mathcal{X} \rightarrow \mathbb{K}$, $\forall x_* \in \mathcal{X}^*$, are continuous.

For a normed linear space $\mathcal{X}$, let $\mathcal{O}_{\mathcal{X}}$ be the natural topology induced by the norm on $\mathcal{X}$. Then, $\mathcal{O}_{\text{weak}}(\mathcal{X}) \subseteq \mathcal{O}_{\mathcal{X}}$. We usually call the topology $\mathcal{O}_{\mathcal{X}}$ the strong topology. A set $S \subseteq \mathcal{X}$ is weakly open, that is $S \in \mathcal{O}_{\text{weak}}(\mathcal{X})$, then S is strongly open, that is $S \in \mathcal{O}_{\mathcal{X}}$, and if S is weakly closed, then it

is strongly closed, but not conversely. We will denote the topological space $(\mathcal{X}, \mathcal{O}_{\text{weak}}(\mathcal{X}))$ by $\mathcal{X}_{\text{weak}}$.

Proposition 7.116 *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$. Then, the following statements hold.*

(i) *A basis for $\mathcal{O}_{\text{weak}}(\mathcal{X})$ consists of all sets of the form*

$$\{x \in \mathcal{X} \mid \langle x_{*i}, x \rangle \in O_i, i = 1, \dots, n\} \quad (7.8)$$

where $n \in \mathbb{N}$, $O_i \in \mathcal{O}_{\mathbb{K}}$ and $x_{*i} \in \mathcal{X}^*$, $i = 1, \dots, n$, and $\mathcal{O}_{\mathbb{K}}$ is the natural topology on $\mathbb{K}$. A basis at $x_0 \in \mathcal{X}$ for $\mathcal{O}_{\text{weak}}(\mathcal{X})$ consists of all sets of the form

$$\{x \in \mathcal{X} \mid |\langle x_{*i}, x - x_0 \rangle| < \epsilon, i = 1, \dots, n\} \quad (7.9)$$

where $\epsilon \in (0, \infty) \subset \mathbb{R}$, $n \in \mathbb{N}$, and $x_{*i} \in \mathcal{X}^*$, $i = 1, \dots, n$.

(ii) $\mathcal{X}_{\text{weak}}$ is completely regular $(\mathbf{T}_{3\frac{1}{2}})$.

(iii) *For a sequence $(x_k)_{k=1}^\infty \subseteq \mathcal{X}_{\text{weak}}$, $x_0 \in \mathcal{X}$ is the limit point of the sequence in the weak topology if, and only if, $\lim_{k \in \mathbb{N}} \langle x_*, x_k \rangle = \langle x_*, x_0 \rangle$, $\forall x_* \in \mathcal{X}^*$. In this case, we will write $\lim_{k \in \mathbb{N}} x_k = x_0$ weakly and say that $(x_k)_{k=1}^\infty$ converges weakly to x_0 .*

Proof (i) By Definition 7.115, $\mathcal{O}_{\text{weak}}(\mathcal{X})$ is the topology generated by sets of the form (7.8). We will show that these sets form a basis for the topology by Proposition 3.18. Take $B = \{x \in \mathcal{X} \mid \langle \vartheta_*, x \rangle \in \mathbb{K}\}$, which is of the form (7.8) and $B = \mathcal{X}$. $\forall x \in \mathcal{X}$, we have $x \in B$. For any $B_1, B_2 \subseteq \mathcal{X}$, which are of the form (7.8), clearly, $B_1 \cap B_2$ is again of the form (7.8). Hence, Proposition 3.18 applies and the sets of the form (7.8) form a basis for $\mathcal{O}_{\text{weak}}(\mathcal{X})$.

Let $B \subseteq \mathcal{X}$ be any set of the form (7.9). Clearly, $x_0 \in B \in \mathcal{O}_{\text{weak}}(\mathcal{X})$. $\forall O \in \mathcal{O}_{\text{weak}}(\mathcal{X})$ with $x_0 \in O$, by Definition 3.17, there exists a basis open set $B_1 := \{x \in \mathcal{X} \mid \langle x_{*i}, x \rangle \in O_i, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$ and some $x_{*i} \in \mathcal{X}^*$ and $O_i \in \mathcal{O}_{\mathbb{K}}$, $i = 1, \dots, n$, such that $x_0 \in B_1 \subseteq O$. For each $i = 1, \dots, n$, $c_i := \langle x_{*i}, x_0 \rangle \in O_i \in \mathcal{O}_{\mathbb{K}}$. Then, $\exists \epsilon_i \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathbb{K}}(c_i, \epsilon_i) \subseteq O_i$. Take $\epsilon = \min_{1 \leq i \leq n} \epsilon_i \in (0, \infty) \subset \mathbb{R}$ and $B_2 := \{x \in \mathcal{X} \mid |\langle x_{*i}, x - x_0 \rangle| < \epsilon, i = 1, \dots, n\}$. Clearly, B_2 is of the form (7.9) and $x_0 \in B_2 \subseteq B_1 \subseteq O$. Hence, sets of the form (7.9) form a basis at x_0 for $\mathcal{O}_{\text{weak}}(\mathcal{X})$. Thus, (i) holds.

(ii) $\forall x_1, x_2 \in \mathcal{X}$ with $x_1 \neq x_2$, we have $x_1 - x_2 \neq \vartheta$. By Proposition 7.85, $\exists x_* \in \mathcal{X}^*$ with $\|x_*\| = 1$ such that $\langle x_*, x_1 - x_2 \rangle = \|x_1 - x_2\| > 0$. Let $O_1 := \{a \in \mathbb{K} \mid \operatorname{Re}(a) > \operatorname{Re}(\langle x_*, x_2 \rangle) + \|x_1 - x_2\|/2\}$ and $O_2 := \{a \in \mathbb{K} \mid \operatorname{Re}(a) < \operatorname{Re}(\langle x_*, x_2 \rangle) + \|x_1 - x_2\|/2\}$. Then, $O_1, O_2 \in \mathcal{O}_{\mathbb{K}}$ and $O_1 \cap O_2 = \emptyset$. Let $B_1 := \{x \in \mathcal{X} \mid \langle x_*, x \rangle \in O_1\}$ and $B_2 := \{x \in \mathcal{X} \mid \langle x_*, x \rangle \in O_2\}$. Then, $B_1, B_2 \in \mathcal{O}_{\text{weak}}(\mathcal{X})$, $x_1 \in B_1$, $x_2 \in B_2$, and $B_1 \cap B_2 = \emptyset$. This shows that $\mathcal{X}_{\text{weak}}$ is Hausdorff.

Next, we show that $\mathcal{X}_{\text{weak}}$ is completely regular. Fix any weakly closed set $F \subseteq \mathcal{X}_{\text{weak}}$ and $x_0 \in \widetilde{F} \in \mathcal{O}_{\text{weak}}(\mathcal{X})$. Then, there exists a basis open set $B = \{x \in \mathcal{X} \mid \langle x_{*i}, x \rangle \in O_i, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$ and some $x_{*i} \in \mathcal{X}^*$ and $O_i \in \mathcal{O}_{\mathbb{K}}, i = 1, \dots, n$, such that $x_0 \in B \subseteq \widetilde{F}$. Let $O := \prod_{i=1}^n O_i \subseteq \mathbb{K}^n$ which is open. Let $p_0 := (\langle x_{*1}, x_0 \rangle, \dots, \langle x_{*n}, x_0 \rangle) \in \mathbb{K}^n$. Then, $p_0 \in O$. By Propositions 4.11 and 3.61, $\mathbb{K}^n$ is normal and therefore completely regular. Then, there exists a continuous function $f : \mathbb{K}^n \rightarrow [0, 1] \subset \mathbb{R}$ such that $f|_{\widetilde{O}} = 0$ and $f(p_0) = 1$. Define $g : \mathcal{X}_{\text{weak}} \rightarrow [0, 1] \subset \mathbb{R}$ by $g(x) = f(\langle x_{*1}, x \rangle, \dots, \langle x_{*n}, x \rangle), \forall x \in \mathcal{X}_{\text{weak}}$. By Propositions 3.12 and 3.32, g is a continuous real-valued function on $\mathcal{X}_{\text{weak}}$, $g(x_0) = f(p_0) = 1$, and $g|_F = 0$. Hence, $\mathcal{X}_{\text{weak}}$ is completely regular. Thus, (ii) holds.

(iii) “Only if” $\forall x_* \in \mathcal{X}^*, x_* : \mathcal{X} \rightarrow \mathbb{K}$ is weakly continuous. By Proposition 3.66, we have $\lim_{k \in \mathbb{N}} \langle x_*, x_k \rangle = \langle x_*, x_0 \rangle$.

“If” Let $(x_k)_{k=1}^\infty$ satisfy that $\lim_{k \in \mathbb{N}} \langle x_*, x_k \rangle = \langle x_*, x_0 \rangle, \forall x_* \in \mathcal{X}^*$. For any basis open set $B = \{x \in \mathcal{X} \mid \langle x_{*i}, x \rangle \in O_i, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$ and some $x_{*i} \in \mathcal{X}^*$ and $O_i \in \mathcal{O}_{\mathbb{K}}, i = 1, \dots, n$, with $x_0 \in B$, we have that $\forall i = 1, \dots, n, \exists N_i \in \mathbb{N}$ such that $\langle x_{*i}, x_k \rangle \in O_i, \forall k \geq N_i$. Take $N = \max_{1 \leq i \leq n} N_i \in \mathbb{N}$. $\forall k \geq N, x_k \in B$. This shows that $(x_k)_{k=1}^\infty$ converges weakly to x_0 . Hence, (iii) holds.

This completes the proof of the proposition. $\square$

Proposition 7.117 *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$ and $\mathcal{X}_{\text{weak}}$ be the topological space of $\mathcal{X}$ endowed with the weak topology. Then, vector addition $\oplus : \mathcal{X}_{\text{weak}} \times \mathcal{X}_{\text{weak}} \rightarrow \mathcal{X}_{\text{weak}}$ is continuous; and scalar multiplication $\otimes : \mathbb{K} \times \mathcal{X}_{\text{weak}} \rightarrow \mathcal{X}_{\text{weak}}$ is continuous.*

Proof Fix $(x_0, y_0) \in \mathcal{X}_{\text{weak}} \times \mathcal{X}_{\text{weak}}$. We will show that $\oplus$ is continuous at (x_0, y_0) . Fix a basis open set $B \in \mathcal{O}_{\text{weak}}(\mathcal{X})$ with $x_0 + y_0 \in B$. Then, by Proposition 7.116, $B = \{z \in \mathcal{X} \mid |\langle x_{*i}, z - x_0 - y_0 \rangle| < \epsilon, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$, for some $\epsilon \in (0, \infty) \subset \mathbb{R}$, and for some $x_{*i} \in \mathcal{X}^*, i = 1, \dots, n$. Let $B_1 := \{x \in \mathcal{X} \mid |\langle x_{*i}, x - x_0 \rangle| < \epsilon/2, i = 1, \dots, n\}$ and $B_2 := \{y \in \mathcal{X} \mid |\langle x_{*i}, y - y_0 \rangle| < \epsilon/2, i = 1, \dots, n\}$. Clearly, $B_1, B_2 \in \mathcal{O}_{\text{weak}}(\mathcal{X})$, $B_1 \times B_2 \in \mathcal{O}_{\mathcal{X}_{\text{weak}} \times \mathcal{X}_{\text{weak}}}$, and $(x_0, y_0) \in B_1 \times B_2$. $\forall (x, y) \in B_1 \times B_2, \forall i = 1, \dots, n, |\langle x_{*i}, x + y - x_0 - y_0 \rangle| \leq |\langle x_{*i}, x - x_0 \rangle| + |\langle x_{*i}, y - y_0 \rangle| < \epsilon$. Hence, $x + y \in B$. This shows that $\oplus$ is continuous at (x_0, y_0) . By the arbitrariness of (x_0, y_0) and Proposition 3.9, $\oplus$ is continuous.

Fix $(\alpha_0, x_0) \in \mathbb{K} \times \mathcal{X}_{\text{weak}}$. We will show that $\otimes$ is continuous at (α_0, x_0) . Fix a basis open set $B \in \mathcal{O}_{\text{weak}}(\mathcal{X})$ with $\alpha_0 x_0 \in B$. Then, by Proposition 7.116, $B = \{z \in \mathcal{X} \mid |\langle x_{*i}, z - \alpha_0 x_0 \rangle| < \epsilon, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$, for some $\epsilon \in (0, \infty) \subset \mathbb{R}$, and for some $x_{*i} \in \mathcal{X}^*, i = 1, \dots, n$. Let $M := \max_{1 \leq i \leq n} \|x_{*i}\| \in [0, \infty) \subset \mathbb{R}$. Let $B_1 := \{\alpha \in \mathbb{K} \mid |\alpha - \alpha_0| < \frac{\epsilon}{\epsilon + 2M\|x_0\|}\}$ and $B_2 := \{x \in \mathcal{X} \mid |\langle x_{*i}, x - x_0 \rangle| < \frac{\epsilon}{2(1+|\alpha_0|)}, i = 1, \dots, n\}$. Clearly, $B_1 \in \mathcal{O}_{\mathbb{K}}$ and $B_2 \in \mathcal{O}_{\text{weak}}(\mathcal{X})$, $B_1 \times B_2 \in \mathcal{O}_{\mathbb{K} \times \mathcal{X}_{\text{weak}}}$, and $(\alpha_0, x_0) \in B_1 \times B_2$. $\forall (\alpha, x) \in B_1 \times B_2, \forall i = 1, \dots, n, |\langle x_{*i}, \alpha x - \alpha_0 x_0 \rangle| \leq$

$|\langle x_{*i}, \alpha x - \alpha x_0 \rangle| + |\langle x_{*i}, \alpha x_0 - \alpha_0 x_0 \rangle| = |\alpha| |\langle x_{*i}, x - x_0 \rangle| + |\alpha - \alpha_0| |\langle x_{*i}, x_0 \rangle| \leq (|\alpha_0| + |\alpha - \alpha_0|) |\langle x_{*i}, x - x_0 \rangle| + |\alpha - \alpha_0| \|x_{*i}\| \|x_0\| \leq (|\alpha_0| + 1) |\langle x_{*i}, x - x_0 \rangle| + |\alpha - \alpha_0| M \|x_0\| < \epsilon$. Hence, $\alpha x \in B$. This shows that $\otimes$ is continuous at (α_0, x_0) . By the arbitrariness of (α_0, x_0) and Proposition 3.9, $\otimes$ is continuous.

This completes the proof of the proposition. $\square$

Proposition 7.118 *Let $\mathcal{X}$ be a finite-dimensional normed linear space over the field $\mathbb{K}$, $\mathcal{O}_{\mathcal{X}}$ be the strong topology on $\mathcal{X}$, and $\mathcal{O}_{\text{weak}}(\mathcal{X})$ be the weak topology on $\mathcal{X}$. Then, $\mathcal{O}_{\mathcal{X}} = \mathcal{O}_{\text{weak}}(\mathcal{X})$.*

Proof Clearly, $\mathcal{O}_{\text{weak}}(\mathcal{X}) \subseteq \mathcal{O}_{\mathcal{X}}$. Fix any basis open set $O = \mathcal{B}_{\mathcal{X}}(x_0, r) \in \mathcal{O}_{\mathcal{X}}$ with $x_0 \in \mathcal{X}$ and $r \in (0, \infty) \subset \mathbb{R}$. We will show that $O \in \mathcal{O}_{\text{weak}}(\mathcal{X})$. Let $n \in \mathbb{Z}_+$ be the dimension of $\mathcal{X}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $n = 0$; Case 2: $n \in \mathbb{N}$. Case 1: $n = 0$. Then, $\mathcal{X}$ is a singleton set. $O = \mathcal{X} \in \mathcal{O}_{\text{weak}}(\mathcal{X})$. This case is proved.

Case 2: $n \in \mathbb{N}$. Let $\{e_1, \dots, e_n\} \subseteq \mathcal{X}$ be a basis of $\mathcal{X}$ with $\|e_i\| = 1$, $i = 1, \dots, n$. $\forall i = 1, \dots, n$, let $f_i : \mathcal{X} \rightarrow \mathbb{K}$ be defined by $f_i(x) = \alpha_i$, $\forall x = \sum_{j=1}^n \alpha_j e_j \in \mathcal{X}$. Clearly, f_i is well-defined and is a linear functional. By Proposition 7.67, f_i is continuous. Denote f_i by $e_{*i} \in \mathcal{X}^*$. $\forall x_1 \in O$, let $\delta = (r - \|x_1 - x_0\|)/n \in (0, \infty) \subset \mathbb{R}$. Let $B := \{x \in \mathcal{X} \mid |\langle e_{*i}, x - x_1 \rangle| < \delta, i = 1, \dots, n\}$. By Proposition 7.116, $B \in \mathcal{O}_{\text{weak}}(\mathcal{X})$. $\forall x \in B$, let $x - x_1 = \sum_{j=1}^n \alpha_j e_j$. Then, $\forall i = 1, \dots, n$, $|\alpha_i| = |\langle e_{*i}, x - x_1 \rangle| < \delta$. Thus, $\|x - x_1\| \leq \sum_{j=1}^n |\alpha_j| \|e_j\| < n\delta = r - \|x_1 - x_0\|$. Then, we have $\|x - x_0\| \leq \|x - x_1\| + \|x_1 - x_0\| < r$ and $x \in O$. Hence, $x_1 \in B \subseteq O$. Therefore, $O \in \mathcal{O}_{\text{weak}}(\mathcal{X})$.

In both cases, we have shown that $O \in \mathcal{O}_{\text{weak}}(\mathcal{X})$. Then, $\mathcal{O}_{\mathcal{X}} \subseteq \mathcal{O}_{\text{weak}}(\mathcal{X})$. Hence, $\mathcal{O}_{\mathcal{X}} = \mathcal{O}_{\text{weak}}(\mathcal{X})$. This completes the proof of the proposition. $\square$

Given a normed linear space $\mathcal{X}$, its dual $\mathcal{X}^*$ is a Banach space. On $\mathcal{X}^*$, we can also talk about the notion of weak topology. The weak topology of $\mathcal{X}^*$ is the weakest topology on $\mathcal{X}^*$ such that all of the functional in $\mathcal{X}^{**}$ are continuous. This topology turns out to be less useful than the weak topology for $\mathcal{X}^*$ generated by $\mathcal{X}$ (or more precisely by $\phi(\mathcal{X})$, where $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ is the natural mapping). This leads us to the following definition.

Definition 7.119 *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $\mathcal{X}^*$ be its dual, and $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ be the natural mapping. The weak* topology on $\mathcal{X}^*$, denoted by $\mathcal{O}_{\text{weak}^*}(\mathcal{X}^*) \subseteq \mathcal{X}^{*2}$, is the weak topology generated by $\phi(\mathcal{X})$, that is the weakest topology on $\mathcal{X}^*$ such that $\phi(x) : \mathcal{X}^* \rightarrow \mathbb{K}$, $\forall x \in \mathcal{X}$, are continuous.*

For a normed linear space $\mathcal{X}$, let $\mathcal{O}_{\mathcal{X}^*}$ be the natural topology induced by the norm on $\mathcal{X}^*$. Then, $\mathcal{O}_{\text{weak}^*}(\mathcal{X}^*) \subseteq \mathcal{O}_{\text{weak}}(\mathcal{X}^*) \subseteq \mathcal{O}_{\mathcal{X}^*}$. Therefore, the

weak* topology is weaker than the weak topology on $\mathcal{X}^*$, which is further weaker than the strong topology on $\mathcal{X}^*$, $\mathcal{O}_{\mathcal{X}^*}$. Clearly, if $\mathcal{X}$ is reflexive, then the weak topology and the weak* topology coincide. We will denote the topological space $(\mathcal{X}^*, \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*))$ by $\mathcal{X}_{\text{weak}^*}^*$. We have the following two basic results for the weak* topology, which are counterpart results for Propositions 7.116 and 7.117.

Proposition 7.120 *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$ and $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ be the natural mapping. Then, the following statements hold.*

(i) *A basis for $\mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$ consists of all sets of the form*

$$\{x_* \in \mathcal{X}^* \mid \langle x_*, x_i \rangle = \langle \phi(x_i), x_* \rangle \in O_i, i = 1, \dots, n\} \quad (7.10)$$

where $n \in \mathbb{N}$, $O_i \in \mathcal{O}_{\mathbb{K}}$ and $x_i \in \mathcal{X}$, $i = 1, \dots, n$, and $\mathcal{O}_{\mathbb{K}}$ is the natural topology on $\mathbb{K}$. A basis at $x_{*0} \in \mathcal{X}^*$ for $\mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$ consists of all sets of the form

$$\{x_* \in \mathcal{X}^* \mid |\langle x_* - x_{*0}, x_i \rangle| = |\langle \phi(x_i), x_* - x_{*0} \rangle| < \epsilon, i = 1, \dots, n\} \quad (7.11)$$

where $\epsilon \in (0, \infty) \subset \mathbb{R}$, $n \in \mathbb{N}$, and $x_i \in \mathcal{X}$, $i = 1, \dots, n$.

(ii) $\mathcal{X}_{\text{weak}^*}^*$ is completely regular $(\mathbf{T}_{3\frac{1}{2}})$.

(iii) *For a sequence $(x_{*k})_{k=1}^\infty \subseteq \mathcal{X}_{\text{weak}^*}^*$, $x_{*0} \in \mathcal{X}^*$ is the limit point of the sequence in the weak* topology if, and only if, $\lim_{k \in \mathbb{N}} \langle x_{*k}, x \rangle = \langle x_{*0}, x \rangle$, $\forall x \in \mathcal{X}$. In this case, we will write $\lim_{k \in \mathbb{N}} x_{*k} = x_{*0}$ weak* and say that $(x_{*k})_{k=1}^\infty$ converges weak* to x_{*0} .*

Proof (i) By Definition 7.119, $\mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$ is the topology generated by sets of the form (7.10). We will show that these sets form a basis for the topology by Proposition 3.18. Take $B = \{x_* \in \mathcal{X}^* \mid \langle x_*, \vartheta \rangle \in \mathbb{K}\}$, which is of the form (7.10) and $B = \mathcal{X}^*$. $\forall x_* \in \mathcal{X}^*$, we have $x_* \in B$. For any $B_1, B_2 \subseteq \mathcal{X}^*$, which are of the form (7.10), clearly, $B_1 \cap B_2$ is again of the form (7.10). Hence, Proposition 3.18 applies and the sets of the form (7.10) form a basis for $\mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$.

Let $B \subseteq \mathcal{X}^*$ be any set of the form (7.11). Clearly, $x_{*0} \in B \in \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$. $\forall O \in \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$ with $x_{*0} \in O$, by Definition 3.17, there exists a basis open set $B_1 := \{x_* \in \mathcal{X}^* \mid \langle x_*, x_i \rangle \in O_i, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$ and some $x_i \in \mathcal{X}$ and $O_i \in \mathcal{O}_{\mathbb{K}}$, $i = 1, \dots, n$, such that $x_{*0} \in B_1 \subseteq O$. For each $i = 1, \dots, n$, $c_i := \langle x_{*0}, x_i \rangle \in O_i \in \mathcal{O}_{\mathbb{K}}$. Then, $\exists \epsilon_i \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathbb{K}}(c_i, \epsilon_i) \subseteq O_i$. Take $\epsilon = \min_{1 \leq i \leq n} \epsilon_i \in (0, \infty) \subset \mathbb{R}$ and $B_2 := \{x_* \in \mathcal{X}^* \mid |\langle x_* - x_{*0}, x_i \rangle| < \epsilon, i = 1, \dots, n\}$. Clearly, B_2 is of the form (7.11) and $x_{*0} \in B_2 \subseteq B_1 \subseteq O$. Hence, sets of the form (7.11) form a basis at x_{*0} for $\mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$. Thus, (i) holds.

(ii) $\forall x_{*1}, x_{*2} \in \mathcal{X}^*$ with $x_{*1} \neq x_{*2}$, we have $x_{*1} - x_{*2} \neq \vartheta_*$. By Lemma 7.75, $\exists x \in \mathcal{X}$ with $\|x\| \leq 1$ such that $\langle x_{*1} - x_{*2}, x \rangle \in \mathbb{R}$ and

$\langle\langle x_{*1} - x_{*2}, x \rangle\rangle \geq \|x_{*1} - x_{*2}\|/2 > 0$. Let $O_1 := \{a \in \mathbb{K} \mid \operatorname{Re}(a) > \operatorname{Re}(\langle\langle x_{*2}, x \rangle\rangle) + \|x_{*1} - x_{*2}\|/4\}$ and $O_2 := \{a \in \mathbb{K} \mid \operatorname{Re}(a) < \operatorname{Re}(\langle\langle x_{*2}, x \rangle\rangle) + \|x_{*1} - x_{*2}\|/4\}$. Then, $O_1, O_2 \in \mathcal{O}_{\mathbb{K}}$ and $O_1 \cap O_2 = \emptyset$. Let $B_1 := \{x_* \in \mathcal{X}^* \mid \langle\langle x_*, x \rangle\rangle \in O_1\}$ and $B_2 := \{x_* \in \mathcal{X}^* \mid \langle\langle x_*, x \rangle\rangle \in O_2\}$. Then, $B_1, B_2 \in \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$, $x_{*1} \in B_1$, $x_{*2} \in B_2$, and $B_1 \cap B_2 = \emptyset$. This shows that $\mathcal{X}_{\text{weak}^*}^*$ is Hausdorff.

Next, we show that $\mathcal{X}_{\text{weak}^*}^*$ is completely regular. Fix any weak* closed set $F \subseteq \mathcal{X}_{\text{weak}^*}^*$ and $x_{*0} \in F \in \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$. Then, there exists a basis open set $B = \{x_* \in \mathcal{X}^* \mid \langle\langle x_*, x_i \rangle\rangle \in O_i, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$ and some $x_i \in \mathcal{X}$ and $O_i \in \mathcal{O}_{\mathbb{K}}$, $i = 1, \dots, n$, such that $x_{*0} \in B \subseteq \tilde{F}$. Let $O := \prod_{i=1}^n O_i \subseteq \mathbb{K}^n$ which is open. Let $p_0 := (\langle\langle x_{*0}, x_1 \rangle\rangle, \dots, \langle\langle x_{*0}, x_n \rangle\rangle) \in \mathbb{K}^n$. Then, $p_0 \in O$. By Propositions 4.11 and 3.61, $\mathbb{K}^n$ is normal and therefore completely regular. Then, there exists a continuous function $f : \mathbb{K}^n \rightarrow [0, 1] \subset \mathbb{R}$ such that $f|_{\tilde{O}} = 0$ and $f(p_0) = 1$. Define $g : \mathcal{X}_{\text{weak}^*}^* \rightarrow [0, 1] \subset \mathbb{R}$ by $g(x_*) = f(\langle\langle x_*, x_1 \rangle\rangle, \dots, \langle\langle x_*, x_n \rangle\rangle)$, $\forall x_* \in \mathcal{X}_{\text{weak}^*}^*$. By Propositions 3.12 and 3.32, g is a continuous real-valued function on $\mathcal{X}_{\text{weak}^*}^*$, $g(x_{*0}) = f(p_0) = 1$, and $g|_F = 0$. Hence, $\mathcal{X}_{\text{weak}^*}^*$ is completely regular. Thus, (ii) holds.

(iii) “Only if” $\forall x \in \mathcal{X}$, $\phi(x) : \mathcal{X}^* \rightarrow \mathbb{K}$ is weak* continuous. By Proposition 3.66, we have $\lim_{k \in \mathbb{N}} \langle\langle x_{*k}, x \rangle\rangle = \lim_{k \in \mathbb{N}} \langle\langle \phi(x), x_{*k} \rangle\rangle = \langle\langle \phi(x), x_{*0} \rangle\rangle = \langle\langle x_{*0}, x \rangle\rangle$.

“If” Let $(x_{*k})_{k=1}^{\infty}$ satisfy that $\lim_{k \in \mathbb{N}} \langle\langle x_{*k}, x \rangle\rangle = \langle\langle x_{*0}, x \rangle\rangle$, $\forall x \in \mathcal{X}$. For any basis open set $B = \{x_* \in \mathcal{X}^* \mid \langle\langle x_*, x_i \rangle\rangle \in O_i, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$ and some $x_i \in \mathcal{X}$ and $O_i \in \mathcal{O}_{\mathbb{K}}$, $i = 1, \dots, n$, with $x_{*0} \in B$, we have that $\forall i = 1, \dots, n$, $\exists N_i \in \mathbb{N}$ such that $\langle\langle x_{*k}, x_i \rangle\rangle \in O_i$, $\forall k \geq N_i$. Take $N = \max_{1 \leq i \leq n} N_i \in \mathbb{N}$. $\forall k \geq N$, $x_{*k} \in B$. This shows that $(x_{*k})_{k=1}^{\infty}$ converges weak* to x_{*0} . Hence, (iii) holds.

This completes the proof of the proposition. $\square$

Proposition 7.121 *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$ and $\mathcal{X}_{\text{weak}^*}^*$ be the topological space of $\mathcal{X}^*$ endowed with the weak* topology. Then, vector addition $\oplus : \mathcal{X}_{\text{weak}^*}^* \times \mathcal{X}_{\text{weak}^*}^* \rightarrow \mathcal{X}_{\text{weak}^*}^*$ is continuous; and scalar multiplication $\otimes : \mathbb{K} \times \mathcal{X}_{\text{weak}^*}^* \rightarrow \mathcal{X}_{\text{weak}^*}^*$ is continuous.*

Proof Fix $(x_{*0}, y_{*0}) \in \mathcal{X}_{\text{weak}^*}^* \times \mathcal{X}_{\text{weak}^*}^*$. We will show that $\oplus$ is continuous at (x_{*0}, y_{*0}) . Fix a basis open set $B \in \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$ with $x_{*0} + y_{*0} \in B$. Then, by Proposition 7.120, $B = \{z_* \in \mathcal{X}^* \mid |\langle\langle z_* - x_{*0} - y_{*0}, x_i \rangle\rangle| < \epsilon, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$, for some $\epsilon \in (0, \infty) \subset \mathbb{R}$, and for some $x_i \in \mathcal{X}$, $i = 1, \dots, n$. Let $B_1 := \{x_* \in \mathcal{X}^* \mid |\langle\langle x_* - x_{*0}, x_i \rangle\rangle| < \epsilon/2, i = 1, \dots, n\}$ and $B_2 := \{y_* \in \mathcal{X}^* \mid |\langle\langle y_* - y_{*0}, x_i \rangle\rangle| < \epsilon/2, i = 1, \dots, n\}$. Clearly, $B_1, B_2 \in \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$, $B_1 \times B_2 \in \mathcal{O}_{\mathcal{X}_{\text{weak}^*}^* \times \mathcal{X}_{\text{weak}^*}^*}$, and $(x_{*0}, y_{*0}) \in B_1 \times B_2$. $\forall (x_*, y_*) \in B_1 \times B_2$, $\forall i = 1, \dots, n$, $|\langle\langle x_* + y_* - x_{*0} - y_{*0}, x_i \rangle\rangle| \leq |\langle\langle x_* - x_{*0}, x_i \rangle\rangle| + |\langle\langle y_* - y_{*0}, x_i \rangle\rangle| < \epsilon$. Hence, $x_* + y_* \in B$. This shows that $\oplus$ is continuous at (x_{*0}, y_{*0}) . By the arbitrariness of (x_{*0}, y_{*0}) and Proposition 3.9, $\oplus$ is continuous.

Fix $(\alpha_0, x_{*0}) \in \mathbb{K} \times \mathcal{X}_{\text{weak}^*}^*$. We will show that $\otimes$ is continuous at (α_0, x_{*0}) . Fix a basis open set $B \in \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$ with $\alpha_0 x_{*0} \in B$. Then, by Proposition 7.120, $B = \{z_* \in \mathcal{X}^* \mid |\langle z_* - \alpha_0 x_{*0}, x_i \rangle| < \epsilon, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$, for some $\epsilon \in (0, \infty) \subset \mathbb{R}$, and for some $x_i \in \mathcal{X}, i = 1, \dots, n$. Let $M := \max_{1 \leq i \leq n} \|x_i\| \in [0, \infty) \subset \mathbb{R}$. Let $B_1 := \{\alpha \in \mathbb{K} \mid |\alpha - \alpha_0| < \frac{\epsilon}{\epsilon + 2M\|x_{*0}\|}\}$ and $B_2 := \{x_* \in \mathcal{X}^* \mid |\langle x_* - x_{*0}, x_i \rangle| < \frac{\epsilon}{2(1+|\alpha_0|)}, i = 1, \dots, n\}$. Clearly, $B_1 \in \mathcal{O}_{\mathbb{K}}$ and $B_2 \in \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*)$, $B_1 \times B_2 \in \mathcal{O}_{\mathbb{K} \times \mathcal{X}_{\text{weak}^*}^*}$, and $(\alpha_0, x_{*0}) \in B_1 \times B_2$. $\forall (\alpha, x_*) \in B_1 \times B_2, \forall i = 1, \dots, n, |\langle \alpha x_* - \alpha_0 x_{*0}, x_i \rangle| \leq |\langle \alpha x_* - \alpha x_{*0}, x_i \rangle| + |\langle \alpha x_{*0} - \alpha_0 x_{*0}, x_i \rangle| = |\alpha| |\langle x_* - x_{*0}, x_i \rangle| + |\alpha - \alpha_0| |\langle x_{*0}, x_i \rangle| \leq (|\alpha| + |\alpha - \alpha_0|) |\langle x_* - x_{*0}, x_i \rangle| + |\alpha - \alpha_0| \|x_{*0}\| \|x_i\| \leq (|\alpha| + 1) |\langle x_* - x_{*0}, x_i \rangle| + |\alpha - \alpha_0| M \|x_{*0}\| < \epsilon$. Hence, $\alpha x_* \in B$. This shows that $\otimes$ is continuous at (α_0, x_{*0}) . By the arbitrariness of (α_0, x_{*0}) and Proposition 3.9, $\otimes$ is continuous.

This completes the proof of the proposition. $\square$

The importance of the weak* topology stems from the following Theorem.

Theorem 7.122 (Alaoglu Theorem) *Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$ and $S = \overline{\mathcal{B}_{\mathcal{X}^*}(\vartheta_*, r)}$, for some $r \in \mathbb{Z}_+$. Then, S is weak* compact.*

Proof Let the weak* topology on S be $\mathcal{O}_{\text{weak}^*}(S)$. Denote the topological space $(S, \mathcal{O}_{\text{weak}^*}(S))$ by $\mathcal{S}$. $\forall x \in \mathcal{X}, \forall x_* \in S$, by Proposition 7.72, we have $|\langle x_*, x \rangle| \leq \|x_*\| \|x\| \leq r \|x\|$. Let $I_x := \overline{\mathcal{B}_{\mathbb{K}}(0, r\|x\|)}$. Then, $\langle x_*, x \rangle \in I_x$. Let $I_x \subseteq \mathbb{K}$ be endowed with the subset topology $\mathcal{O}_x$. Denote the topological space $(I_x, \mathcal{O}_x)$ by $\mathcal{I}_x$. By Proposition 5.40 or 7.41, $\mathcal{I}_x$ is compact. Let $\mathcal{P} := \prod_{x \in \mathcal{X}} \mathcal{I}_x$, whose topology is denoted by $\mathcal{O}_{\mathcal{P}}$. By Tychonoff Theorem, $\mathcal{P}$ is compact. Define the equivalence map $E : \mathcal{S} \rightarrow \mathcal{P}$ by $\pi_x(E(x_*)) = \langle x_*, x \rangle, \forall x \in \mathcal{X}, \forall x_* \in S$.

Claim 7.122.1 $E : \mathcal{S} \rightarrow E(\mathcal{S}) \subseteq \mathcal{P}$ is a homeomorphism.

Proof of claim: Clearly, $E : \mathcal{S} \rightarrow E(\mathcal{S})$ is surjective. $\forall x_{*1}, x_{*2} \in S$ with $x_{*1} \neq x_{*2}$, then $\exists x \in \mathcal{X}$ such that $\pi_x(E(x_{*1})) = \langle x_{*1}, x \rangle \neq \langle x_{*2}, x \rangle = \pi_x(E(x_{*2}))$. Then, $E(x_{*1}) \neq E(x_{*2})$. Hence, E is injective. Therefore, $E : \mathcal{S} \rightarrow E(\mathcal{S})$ is bijective and admits inverse $E_{\text{inv}} : E(\mathcal{S}) \rightarrow \mathcal{S}$.

Fix any $x_{*0} \in S$. Fix any basis open set $O \in \mathcal{O}_{\mathcal{P}}$ with $E(x_{*0}) \in O$. By Proposition 3.25, $O = \prod_{x \in \mathcal{X}} O_x$ where $O_x \in \mathcal{O}_x, \forall x \in \mathcal{X}$, and $O_x = I_x$ for all x 's except finitely many x 's, say $x \in X_N$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $X_N = \emptyset$; Case 2: $X_N \neq \emptyset$. Case 1: $X_N = \emptyset$. Take $x_1 = \vartheta$ and $O_1 = \mathbb{K}$. Let $B := \{x_* \in S \mid \langle x_*, x_1 \rangle \in O_1\} = S$. Clearly, $x_{*0} \in B \in \mathcal{O}_{\text{weak}^*}(S)$ and $\forall x_* \in B, \forall x \in \mathcal{X}, \pi_x(E(x_*)) = \langle x_*, x \rangle \in I_x = O_x$. Then, $E(x_*) \in O$ and $E(B) \subseteq O$. Case 2: $X_N \neq \emptyset$. $\forall x \in X_N$, by Proposition 3.4, let $\bar{O}_x \in \mathcal{O}_{\mathbb{K}}$ be such that $\bar{O}_x \cap I_x = O_x$. Let $B := \{x_* \in S \mid \langle x_*, x \rangle \in \bar{O}_x, \forall x \in X_N\}$. Clearly, $x_{*0} \in B \in \mathcal{O}_{\text{weak}^*}(S)$. $\forall x_* \in B, \forall x \in X_N, \pi_x(E(x_*)) = \langle x_*, x \rangle \in$

$\bar{O}_x \cap I_x = O_x$. Hence, $E(x_*) \in O$ and $E(B) \subseteq O$. In both cases, we have shown that $\exists B \in \mathcal{O}_{\text{weak}^*}(S)$ with $x_{*0} \in B$ such that $E(B) \subseteq O$. Hence, E is continuous at x_{*0} . By the arbitrariness of x_{*0} and Proposition 3.9, E is continuous.

Fix any $p_0 \in E(S)$. Let $x_{*0} = E_{\text{inv}}(p_0) \in S$. Fix any basis open set $B \in \mathcal{O}_{\text{weak}^*}(S)$ with $x_{*0} \in B$. By Proposition 7.120, $B = \{x_* \in S \mid \langle\langle x_*, x_i \rangle\rangle \in O_i, i = 1, \dots, n\}$, for some $n \in \mathbb{N}$ and some $x_i \in X$ and $O_i \in \mathcal{O}_{\mathbb{K}}, i = 1, \dots, n$. Without loss of generality, we may assume that $x_1, \dots, x_n$ are distinct (otherwise, say $x_n = x_{n-1}$, then set $\bar{O}_{n-1} = O_n \cap O_{n-1}$ and $\bar{O}_i = O_i, i = 1, \dots, n-2$, and consider $x_1, \dots, x_{n-1}$ with $\bar{O}_1, \dots, \bar{O}_{n-1}$). $\forall x \in X$, if $x = x_i$ for some $i \in \{1, \dots, n\}$, then let $O_x = O_i \cap I_x \in \mathcal{O}_x$; otherwise, let $O_x = I_x \in \mathcal{O}_x$. Let $O := \prod_{x \in X} O_x$. Clearly, $p_0 \in O$. $\forall p \in O \cap E(S)$, let $x_* = E_{\text{inv}}(p) \in S$. $\forall i = 1, \dots, n, \langle\langle x_*, x_i \rangle\rangle = \pi_{x_i}(E(x_*)) \in O_{x_i} = O_i \cap I_{x_i}$. Then, $x_* \in B$ and $E_{\text{inv}}(O \cap E(S)) \subseteq B$. This shows that E_{inv} is continuous at p_0 . By the arbitrariness of p_0 and Proposition 3.9, E_{inv} is continuous.

Hence, $E : S \rightarrow E(S)$ is a homeomorphism. This completes the proof of the claim. $\square$

Next, we will show that $E(S)$ is closed. $\forall p_0 \in \overline{E(S)} \subseteq \mathcal{P}$, define $f_0 : X \rightarrow \mathbb{K}$ by $f_0(x) = \pi_x(p_0), \forall x \in X$. $\forall x_1, x_2 \in X, \forall \alpha, \beta \in \mathbb{K}$, let $z = \alpha x_1 + \beta x_2$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall x \in X$, if $x \in \{x_1, x_2, z\}$, let $O_x = \mathcal{B}_{\mathbb{K}}(\pi_x(p_0), \epsilon) \cap I_x \in \mathcal{O}_x$, otherwise, let $O_x = I_x \in \mathcal{O}_x$. Let $O := \prod_{x \in X} O_x \in \mathcal{O}_{\mathcal{P}}$. Clearly, $p_0 \in O$. By Proposition 3.3, $\exists p_1 \in E(S) \cap O$. Let $x_* = E_{\text{inv}}(p_1) \in S$. $\forall x \in \{x_1, x_2, z\}, \langle\langle x_*, x \rangle\rangle = \pi_x(p_1) \in O_x = \mathcal{B}_{\mathbb{K}}(\pi_x(p_0), \epsilon) \cap I_x \subseteq \mathcal{B}_{\mathbb{K}}(f_0(x), \epsilon)$. Hence, we have $|\langle\langle x_*, x_1 \rangle\rangle - f_0(x_1)| < \epsilon, |\langle\langle x_*, x_2 \rangle\rangle - f_0(x_2)| < \epsilon$, and $|\langle\langle x_*, z \rangle\rangle - f_0(z)| < \epsilon$. This implies that

$$\begin{aligned} |f_0(z) - \alpha f_0(x_1) - \beta f_0(x_2)| &= |f_0(z) - \langle\langle x_*, z \rangle\rangle - \alpha f_0(x_1) \\ &\quad + \alpha \langle\langle x_*, x_1 \rangle\rangle - \beta f_0(x_2) + \beta \langle\langle x_*, x_2 \rangle\rangle| < \epsilon + |\alpha| \epsilon + |\beta| \epsilon \end{aligned}$$

By the arbitrariness of ϵ , we have $f_0(z) = \alpha f_0(x_1) + \beta f_0(x_2)$. Hence, f_0 is linear. $\forall x \in X, f_0(x) = \pi_x(p_0) \in I_x$, then $|f_0(x)| \leq r \|x\|$. Therefore, $f_0 \in X^*$ and $\|f_0\| \leq r$. Hence, $f_0 \in S$. $\forall x \in X, \pi_x(E(f_0)) = \underline{f_0}(x) = \pi_x(p_0)$. Hence, $E(f_0) = p_0$ and $p_0 \in E(S)$. This shows that $\overline{E(S)} \subseteq E(S)$. By Proposition 3.3, $E(S)$ is closed.

By Proposition 5.5 and the compactness of $\mathcal{P}$, we have $E(S)$ is compact. By Proposition 5.7, S is compact. This completes the proof of the theorem. $\square$

Proposition 7.123 *Let X and Y be normed linear spaces over $\mathbb{K}$, $A \in B(X, Y)$, and X_{weak} and Y_{weak} be the topological spaces of X and Y endowed with the weak topology. Then, $A : X_{\text{weak}} \rightarrow Y_{\text{weak}}$ is continuous.*

Proof Fix any basis open set $O_Y \in \mathcal{O}_{\text{weak}}(Y)$, we will show that $A_{\text{inv}}(O_Y) \in \mathcal{O}_{\text{weak}}(X)$. By Proposition 7.116, $O_Y = \{y \in Y \mid \langle\langle y_{*i}, y \rangle\rangle \in O_i, i = 1, \dots, n\}$, where $n \in \mathbb{N}, O_i \in \mathcal{O}_{\mathbb{K}}, y_{*i} \in Y^*, i = 1, \dots, n$. Then, $A_{\text{inv}}(O_Y) = \{x \in X \mid \langle\langle y_{*i}, Ax \rangle\rangle \in O_i, i = 1, \dots, n\} = \{x \in$

$\mathcal{X} \mid \langle \langle A'y_{*i}, x \rangle \rangle \in O_i, i = 1, \dots, n \} \in \mathcal{O}_{\text{weak}}(\mathcal{X})$. Hence, $A : \mathcal{X}_{\text{weak}} \rightarrow \mathcal{Y}_{\text{weak}}$ is continuous. This completes the proof of the proposition. $\square$

Definition 7.124 Let $\mathcal{X}$ be a normed linear space and $D \subseteq \mathcal{X}$. D is said to be locally convex at $x_0 \in D$ if $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$, such that $D_0 := D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$ is convex. D is said to be locally convex if it is locally convex at any $x \in D$.

Then, $D \subseteq \mathcal{X}$ is locally convex at $x_0 \in D$ if, and only if, there exists a basis $\mathcal{B} \subseteq \mathcal{O}$ at x_0 for the subset topological space $(D, \mathcal{O})$ such that each $B \in \mathcal{B}$ is convex. Furthermore, $D \subseteq \mathcal{X}$ is locally convex if, and only if, there exists a basis $\mathcal{B} \subseteq \mathcal{O}$ for the subset topological space $(D, \mathcal{O})$ such that each $B \in \mathcal{B}$ is convex.

Proposition 7.125 Let $\mathcal{X}$ be a Banach space over $\mathbb{K}$. Then, $\mathcal{X}$ is σ -compact if, and only if, $\mathcal{X}$ is finite dimensional.

Proof “Sufficiency” Assume that $\mathcal{X}$ is finite dimensional. Let $K_n := \overline{\mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, n)}$, $\forall n \in \mathbb{N}$. By Proposition 7.42, K_n is compact, $\forall n \in \mathbb{N}$. Clearly, $\bigcup_{n=1}^{\infty} K_n = \mathcal{X}$. Hence, $\mathcal{X}$ is σ -compact.

“Necessity” Assume that $\mathcal{X}$ is σ -compact. Then, $\mathcal{X} = \bigcup_{n=1}^{\infty} K_n$, where $K_n \subseteq \mathcal{X}$ is compact, $\forall n \in \mathbb{N}$. We will show that $\overline{\mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 1)}$ is compact by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\exists n_0 \in \mathbb{N}$, $\exists x_0 \in K_{n_0}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, such that $\mathcal{B}_{\mathcal{X}}(x_0, \delta) \subseteq K_{n_0}$; Case 2: $\forall n \in \mathbb{N}$, K_n does not contain any nonempty open ball as its subset. Case 1: $\exists n_0 \in \mathbb{N}$, $\exists x_0 \in K_{n_0}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, such that $\mathcal{B}_{\mathcal{X}}(x_0, \delta) \subseteq K_{n_0}$. Introduce the transformation $F : \mathcal{X} \rightarrow \mathcal{X}$ by $F(x) = \frac{2}{\delta}(x - x_0)$, $\forall x \in \mathcal{X}$. Clearly, F is a homeomorphism. Then, $\overline{\mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 1)} \subseteq F(K_{n_0})$. By K_{n_0} being compact, we have $F(K_{n_0})$ is compact since F is a homeomorphism. Then, $\overline{\mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 1)}$ is a closed subset of $F(K_{n_0})$ and therefore compact by Proposition 5.5. This proves our intended conclusion. Case 2: $\forall n \in \mathbb{N}$, K_n does not contain any nonempty open ball as its subset. By Proposition 5.5, K_n is closed, $\forall n \in \mathbb{N}$. Then, $\forall x \in \overline{K_n}$, $x \in \widetilde{K_n}$, $\forall n \in \mathbb{N}$, since x is not an interior point of K_n . Then, $\mathcal{X} \supseteq \overline{K_n} \supseteq K_n \cup \widetilde{K_n} = \mathcal{X}$. This implies that $\mathcal{X} = \overline{K_n} = \widetilde{K_n}$. Hence, K_n is a nowhere dense subset in $\mathcal{X}$. This then leads to the conclusion that $\mathcal{X} = \bigcup_{n=1}^{\infty} K_n$ is of first category. By Baire Category Theorem 4.40, $\mathcal{X}$ is second category everywhere. Then, $\mathcal{X} \neq \emptyset$ is not of first category. This is a contradiction. Hence, this case is impossible.

We conclude from the previous paragraph that $K := \overline{\mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 1)}$ is compact. Suppose that $\mathcal{X}$ is not finite dimensional. Then, there exists $(y_n)_{n=1}^{\infty} \subseteq \mathcal{X}$ such that y_n 's are distinct and the set $E := \{y_n \mid n \in \mathbb{N}\} \subseteq \mathcal{X}$ is a linearly independent set. Without loss of generality, we may assume that $\|y_n\| = 1$, $\forall n \in \mathbb{N}$. Let $x_1 := y_1 \in K$, $z_n \in \arg \min_{z \in \text{span}(\{x_1, \dots, x_{n-1}\})} \|y_n - z\|$, and $x_n := \frac{1}{\|y_n - z_n\|}(y_n - z_n) \in K$, $n = 2, 3, \dots$. Clearly, z_n exists since the minimization is over finite dimensional

subspaces of $\mathcal{X}$. By E being a linearly independent set, we have $\|y_n - z_n\| > 0$, $\forall n \in \mathbb{N}$. Hence, $x_n \in \mathcal{X}$ is well-defined, $\forall n \in \mathbb{N}$. Clearly, $\|x_n\| = 1$ and $x_n \in K$, $\forall n \in \mathbb{N}$. By K being compact and Borel-Lebesgue Theorem 5.37, $\exists x_0 \in K$ and a subsequence $(x_{n_i})_{i=1}^\infty$ of $(x_n)_{n=1}^\infty$ such that $\lim_{i \in \mathbb{N}} x_{n_i} = x_0$. Since $x_{n_i} \in \text{span}(E)$, $\forall i \in \mathbb{N}$, then $x_0 \in \overline{\text{span}(E)}$. This implies that $x_0 = \sum_{j=1}^\infty a_j x_j$, where $a_j \in \mathbb{K}$, $\forall j \in \mathbb{N}$. Note that $x_0 = \lim_{i \in \mathbb{N}} x_{n_i} = \lim_{i \in \mathbb{N}} \sum_{j=1}^{n_i-1} a_j x_j$ and thus $\lim_{i \in \mathbb{N}} (x_{n_i} - \sum_{j=1}^{n_i-1} a_j x_j) = \vartheta_{\mathcal{X}}$. But, $\forall i \in \mathbb{N}$, $\|x_{n_i} - \sum_{j=1}^{n_i-1} a_j x_j\| = \frac{1}{\|y_{n_i} - z_{n_i}\|} \|y_{n_i} - \sum_{j=1}^{n_i-1} (\|y_{n_i} - z_{n_i}\|) a_j x_j - z_{n_i}\|$. The vector $\sum_{j=1}^{n_i-1} (\|y_{n_i} - z_{n_i}\|) a_j x_j - z_{n_i} \in \text{span}(\{x_1, \dots, x_{n_i-1}\})$. By the definition of z_{n_i} , we have $\|x_{n_i} - \sum_{j=1}^{n_i-1} a_j x_j\| = \frac{1}{\|y_{n_i} - z_{n_i}\|} \|y_{n_i} - \sum_{j=1}^{n_i-1} (\|y_{n_i} - z_{n_i}\|) a_j x_j - z_{n_i}\| \geq \frac{1}{\|y_{n_i} - z_{n_i}\|} \|y_{n_i} - z_{n_i}\| = 1$. This contradicts with $\lim_{i \in \mathbb{N}} (x_{n_i} - \sum_{j=1}^{n_i-1} a_j x_j) = \vartheta_{\mathcal{X}}$. Hence, the hypothesis does not hold, and $\mathcal{X}$ must be finite dimensional.

This completes the proof of the proposition. $\square$

Chapter 8

Global Theory of Optimization

In this chapter, we are going to develop a number of tools for optimization in real normed linear spaces, including the geometric form of Hahn-Banach Theorem, minimum norm duality for convex sets, Fenchel Duality Theorem, and Lagrange multiplier theory for convex programming. In this chapter, we will restrict our attention to real spaces, rather than complex ones.

8.1 Hyperplanes and Convex Sets

Definition 8.1 *Let $\mathcal{X}$ be a real vector space. A hyperplane H is a maximal proper linear variety in $\mathcal{X}$, that is, a linear variety $H \subset \mathcal{X}$, and if $V \supseteq H$ is a linear variety, then either $V = H$ or $V = \mathcal{X}$.*

Proposition 8.2 *Let $\mathcal{X}$ be a real vector space. $H \subseteq \mathcal{X}$ is a hyperplane if, and only if, there exist a linear functional $f : \mathcal{X} \rightarrow \mathbb{R}$ and $c \in \mathbb{R}$ with f being not identically equal to zero, such that $H = \{x \in \mathcal{X} \mid f(x) = c\}$.*

Proof “Necessity” Let H be a hyperplane. Then, H is a linear variety. There exists a subspace M and $x_0 \in \mathcal{X}$ such that $H = x_0 + M$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x_0 \notin M$; Case 2: $x_0 \in M$. Case 1: $x_0 \notin M$. Let $\bar{M} := \text{span}(M \cup \{x_0\})$. Clearly, $\bar{M} \supset H$ and is a linear variety. Then, $\bar{M} = \mathcal{X}$ by Definition 8.1. $\forall x \in \mathcal{X}$, x can be uniquely written as $\alpha x_0 + m$, where $\alpha \in \mathbb{R}$ and $m \in M$. Define $f : \mathcal{X} \rightarrow \mathbb{R}$ by $f(x) = f(\alpha x_0 + m) = \alpha$, $\forall x \in \mathcal{X}$. Clearly, f is a linear functional and is not identically equal to zero. It is straightforward to verify that $H = \{x \in \mathcal{X} \mid f(x) = 1\}$. Case 2: $x_0 \in M$. Then, $H = M$. By Definition 8.1, $\exists x_1 \in \mathcal{X} \setminus H$. Let $\bar{M} := \text{span}(M \cup \{x_1\})$. Clearly, $\bar{M} \supset H$ and is a linear variety. Then, $\bar{M} = \mathcal{X}$ by Definition 8.1. $\forall x \in \mathcal{X}$, x can be uniquely written as $\alpha x_1 + m$, where $\alpha \in \mathbb{R}$ and $m \in M$. Define $f : \mathcal{X} \rightarrow \mathbb{R}$

by $f(x) = f(\alpha x_1 + m) = \alpha$, $\forall x \in \mathcal{X}$. Clearly, f is a linear functional and is not identically equal to zero. It is straightforward to verify that $H = \{x \in \mathcal{X} \mid f(x) = 0\}$.

“Sufficiency” Let $H = \{x \in \mathcal{X} \mid f(x) = c\}$, where $f : \mathcal{X} \rightarrow \mathbb{R}$ is a linear functional, $c \in \mathbb{R}$, and f is not identically equal to zero. Let $M := \mathcal{N}(f)$. Clearly, M is a proper subspace of $\mathcal{X}$. Since f is not identically equal to zero, then $\exists x_0 \in \mathcal{X} \setminus M$ such that $f(x_0) = 1$. $\forall x \in \mathcal{X}$, $f(x - f(x)x_0) = 0$. Then, $x - f(x)x_0 \in M$ and $x \in \text{span}(M \cup \{x_0\})$. Hence, $\mathcal{X} = \text{span}(M \cup \{x_0\})$ and M is a maximal proper subspace. Then, $H = cx_0 + M$ is a hyperplane.

This completes the proof of the proposition. $\square$

Consider a hyperplane H in a real normed linear space $\mathcal{X}$. By Proposition 7.17, $\overline{H}$ is a linear variety. By Definition 8.1, $\overline{H} = \mathcal{X}$ or $\overline{H} = H$. Thus, a hyperplane in a real normed linear space must either be dense or closed.

Proposition 8.3 *Let $\mathcal{X}$ be a real normed linear space and $H \subseteq \mathcal{X}$. H is a closed hyperplane if, and only if, there exist $x_* \in \mathcal{X}^*$ and $c \in \mathbb{R}$ with $x_* \neq \vartheta_*$ such that $H = \{x \in \mathcal{X} \mid \langle x_*, x \rangle = c\}$.*

Proof “Necessity” Let H be a closed hyperplane. By Proposition 8.2, there exists a linear functional $f : \mathcal{X} \rightarrow \mathbb{R}$ and $c \in \mathbb{R}$ with f being not identically equal to zero such that $H = \{x \in \mathcal{X} \mid f(x) = c\}$. All we need to show is that $f \in \mathcal{X}^*$. Since H is a linear variety, then $H \neq \emptyset$. Fix $x_0 \in H$. It is easy to show that $M := H - x_0 = \mathcal{N}(f)$. By Proposition 7.16, M is closed. By Proposition 7.72, $f \in \mathcal{X}^*$.

“Sufficiency” By Proposition 8.2, H is a hyperplane. By the continuity of x_* and Proposition 3.10, H is closed. Hence, H is a closed hyperplane.

This completes the proof of the proposition. $\square$

For a real normed linear space $\mathcal{X}$ and a closed hyperplane $H \subset \mathcal{X}$. Then, $H = \{x \in \mathcal{X} \mid \langle x_*, x \rangle = c\}$, where $\vartheta_* \neq x_* \in \mathcal{X}^*$ and $c \in \mathbb{R}$. We associate four sets with H : (a) $\{x \in \mathcal{X} \mid \langle x_*, x \rangle \leq c\}$; (b) $\{x \in \mathcal{X} \mid \langle x_*, x \rangle < c\}$; (c) $\{x \in \mathcal{X} \mid \langle x_*, x \rangle \geq c\}$; (d) $\{x \in \mathcal{X} \mid \langle x_*, x \rangle > c\}$, which are called half-spaces. The first two are negative half-spaces. The last two are positive half-spaces. The first and third are closed. The second and fourth are open.

Definition 8.4 *Let $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathcal{X}$ be convex with $\vartheta \in K^\circ$. The Minkowski functional $p : \mathcal{X} \rightarrow \mathbb{R}$ of K is defined by $p(x) = \inf \{r \in \mathbb{R} \mid r^{-1}x \in K, r > 0\}$, $\forall x \in \mathcal{X}$.*

Proposition 8.5 *Let $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathcal{X}$ be convex with $\vartheta \in K^\circ$. Then, the Minkowski functional $p : \mathcal{X} \rightarrow \mathbb{R}$ of K satisfies*

- (i) $0 \leq p(x) < +\infty$, $\forall x \in \mathcal{X}$;
- (ii) $p(\alpha x) = \alpha p(x)$, $\forall x \in \mathcal{X}$, $\forall \alpha \in [0, \infty) \subset \mathbb{R}$;

(iii) $p(x_1 + x_2) \leq p(x_1) + p(x_2)$, $\forall x_1, x_2 \in \mathcal{X}$;

(iv) p is uniformly continuous;

(v) $\overline{K} = \{x \in \mathcal{X} \mid p(x) \leq 1\}$; $K^\circ = \{x \in \mathcal{X} \mid p(x) < 1\}$.

Furthermore, (ii) and (iii) implies that p is a sublinear functional.

Proof (i) Since $\vartheta \in K^\circ$, then $\exists \epsilon_0 \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(\vartheta, \epsilon_0) \subseteq K$. $\forall x \in \mathcal{X}$, we have either $x = \vartheta$, then $p(\vartheta) = 0$; or $x \neq \vartheta$, then $0 \leq p(x) \leq \|x\|/\epsilon_0 < +\infty$.

(ii) $\forall x \in \mathcal{X}$, $\forall \alpha \in [0, \infty) \subset \mathbb{R}$, we will distinguish three exhaustive and mutually exclusive cases: Case 1: $x = \vartheta$; Case 2: $x \neq \vartheta$ and $\alpha = 0$; Case 3: $x \neq \vartheta$ and $\alpha > 0$. Case 1: $x = \vartheta$. We have $p(\alpha x) = p(\vartheta) = 0 = \alpha p(x)$. Case 2: $x \neq \vartheta$ and $\alpha = 0$. Then, we have $p(\alpha x) = p(\vartheta) = 0 = \alpha p(x)$. Case 3: $x \neq \vartheta$ and $\alpha > 0$. $\forall r \in \{r \in \mathbb{R} \mid r^{-1}x \in K, r > 0\}$, we have $\alpha r > 0$ and $\alpha r \in \{r \in \mathbb{R} \mid r^{-1}(\alpha x) \in K, r > 0\}$. Hence, $\alpha p(x) \geq p(\alpha x)$. On the other hand, $\forall r \in \{r \in \mathbb{R} \mid r^{-1}(\alpha x) \in K, r > 0\}$, we have $r/\alpha \in \{r \in \mathbb{R} \mid r^{-1}x \in K, r > 0\}$. Hence, $\alpha^{-1}p(\alpha x) \geq p(x)$. Therefore, we have $\alpha p(x) = p(\alpha x)$.

(iii) $\forall x_1, x_2 \in \mathcal{X}$, $\forall r_1 \in \{r \in \mathbb{R} \mid r^{-1}x_1 \in K, r > 0\}$, $\forall r_2 \in \{r \in \mathbb{R} \mid r^{-1}x_2 \in K, r > 0\}$, we have $r_1^{-1}x_1, r_2^{-1}x_2 \in K$. By the convexity of K , we have

$$(r_1 + r_2)^{-1}(x_1 + x_2) = \frac{r_1}{r_1 + r_2}r_1^{-1}x_1 + \frac{r_2}{r_1 + r_2}r_2^{-1}x_2 \in K$$

Then, $r_1 + r_2 \in \{r \in \mathbb{R} \mid r^{-1}(x_1 + x_2) \in K, r > 0\}$. Hence, we have $r_1 + r_2 \geq p(x_1 + x_2)$ and $p(x_1) + p(x_2) \geq p(x_1 + x_2)$.

(iv) $\forall x \in \mathcal{X}$, $0 \leq p(x) \leq \|x\|/\epsilon_0$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta = \epsilon_0 \epsilon \in (0, \infty) \subset \mathbb{R}$. $\forall x_1, x_2 \in \mathcal{X}$ with $\|x_1 - x_2\| < \delta$, we have $p(x_1) \leq p(x_2) + p(x_1 - x_2) \leq p(x_2) + \|x_1 - x_2\|/\epsilon_0 < p(x_2) + \epsilon$ and $p(x_2) \leq p(x_1) + p(x_2 - x_1) \leq p(x_1) + \|x_2 - x_1\|/\epsilon_0 < p(x_1) + \epsilon$. Hence, $|p(x_1) - p(x_2)| < \epsilon$. This shows that p is uniformly continuous.

(v) $\forall x \in K^\circ$, we have either $x = \vartheta$, then $p(x) = 0 < 1$; or $x \neq \vartheta$, then $\exists \epsilon \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}(x, \epsilon\|x\|) \subseteq K$, which implies that $p(x) \leq 1/(1 + \epsilon) < 1$. Therefore, $K^\circ \subseteq \{x \in \mathcal{X} \mid p(x) < 1\}$.

$\forall x \in \mathcal{X}$ with $p(x) < 1$, by the continuity of p , $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $p(y) < 1$, $\forall y \in \mathcal{B}(x, \delta)$. $\forall y \in \mathcal{B}(x, \delta)$, $y \in K$ by the convexity of K , $p(y) < 1$, and the fact that $\vartheta \in K$. Then, $\mathcal{B}(x, \delta) \subseteq K$ and $x \in K^\circ$. Therefore, we have $\{x \in \mathcal{X} \mid p(x) < 1\} \subseteq K^\circ$. Thus, $K^\circ = \{x \in \mathcal{X} \mid p(x) < 1\}$.

$\forall x \in K$, we have $p(x) \leq 1$. Then, $K \subseteq \{x \in \mathcal{X} \mid p(x) \leq 1\}$. By the continuity of p and Proposition 3.10, $\{x \in \mathcal{X} \mid p(x) \leq 1\}$ is closed. Then, $\overline{K} \subseteq \{x \in \mathcal{X} \mid p(x) \leq 1\}$.

$\forall x \in \mathcal{X}$ with $p(x) \leq 1$, $\forall \rho \in (0, 1) \subset \mathbb{R}$, $p(\rho x) = \rho p(x) < 1$. Then, $\rho x \in K^\circ$. Then, by Proposition 4.13, $x \in \overline{K^\circ} \subseteq \overline{K}$. Hence, $\{x \in \mathcal{X} \mid p(x) \leq 1\} \subseteq \overline{K}$. Therefore, $\overline{K} = \{x \in \mathcal{X} \mid p(x) \leq 1\}$.

This completes the proof of the proposition. $\square$

Proposition 8.6 *Let $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathcal{X}$ be convex with $K^\circ \neq \emptyset$. Then, $\overline{K} = \overline{K^\circ}$.*

Proof Clearly, $K^\circ \subseteq K$. Then, $\overline{K^\circ} \subseteq \overline{K}$.

To show that $\overline{K} \subseteq \overline{K^\circ}$, we will distinguish two exhaustive and mutually exclusive cases: Case 1: $\vartheta \in K^\circ$; Case 2: $\vartheta \notin K^\circ$. Case 1: $\vartheta \in K^\circ$. Let $p : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$ be the Minkowski functional of K . By Proposition 8.5, p is a continuous sublinear functional, $K^\circ = \{x \in \mathcal{X} \mid p(x) < 1\}$, and $\overline{K} = \{x \in \mathcal{X} \mid p(x) \leq 1\}$. $\forall x \in \overline{K}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $p(x) \leq 1$. By the convexity of K , $y := \frac{\|x\|}{\epsilon + \|x\|}x \in K$. Then, $p(y) = \frac{\|x\|}{\epsilon + \|x\|}p(x) < 1$. Then, $y \in K^\circ$. Note that $\|x - y\| \leq \frac{\epsilon}{\epsilon + \|x\|} \|x\| < \epsilon$. Hence, $x \in \overline{K^\circ}$. Therefore, $\overline{K} \subseteq \overline{K^\circ}$. This case is proved.

Case 2: $\vartheta \notin K^\circ$. Let $x_0 \in K^\circ \neq \emptyset$. Let $K_1 = K - x_0$. By Propositions 7.16 and 6.39, $K_1^\circ = K^\circ - x_0 \neq \emptyset$ and K_1 is convex. By Case 1, $\overline{K_1} \subseteq \overline{K_1^\circ}$. By Proposition 7.16, $\overline{K} = \overline{x_0 + K_1} = x_0 + \overline{K_1} \subseteq x_0 + \overline{K_1^\circ} = \overline{x_0 + K_1^\circ} = \overline{(x_0 + K_1)^\circ} = \overline{K^\circ}$. This case is proved.

Hence, $\overline{K} = \overline{K^\circ}$. This completes the proof of the proposition. $\square$

8.2 Geometric Form of Hahn-Banach Theorem

The next theorem is the geometric form of Hahn-Banach Theorem.

Theorem 8.7 (Mazur's Theorem) *Let $\mathcal{X}$ be a real normed linear space, $K \subseteq \mathcal{X}$ be a convex set with nonempty interior, and $V \subseteq \mathcal{X}$ is a linear variety with $V \cap K^\circ = \emptyset$. Then, there exists a closed hyperplane H containing V but no interior point of K and K is contained in one of the closed half-spaces associated with H ; that is, $\exists c \in \mathbb{R}$ and $\exists x_* \in \mathcal{X}^*$ with $x_* \neq \vartheta_*$ such that $\langle x_*, v \rangle = c$, $\forall v \in V$, $\langle x_*, k \rangle < c$, $\forall k \in K^\circ$, and $\langle x_*, k \rangle \leq c$, $\forall k \in \overline{K}$.*

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\vartheta \in K^\circ$; Case 2: $\vartheta \notin K^\circ$. Case 1: $\vartheta \in K^\circ$. Let $p : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$ be the Minkowski functional of K . By Proposition 8.5, p is a continuous sublinear functional. Let $M := \text{span}(V)$. Since V is a linear variety and $\vartheta \notin V$, then $\exists x_1 \in V$ such that $V - x_1 =: \bar{M}$ is a subspace in M and $x_1 \notin \bar{M}$. $\forall m \in M$, $\exists! \alpha \in \mathbb{R}$ and $\exists! \bar{m} \in \bar{M}$ such that $m = \alpha x_1 + \bar{m}$. Then, we may define a functional $f : M \rightarrow \mathbb{R}$ by $f(m) = f(\alpha x_1 + \bar{m}) = \alpha$, $\forall m \in M$. Clearly, f is a linear functional on M and $f(v) = 1$, $\forall v \in V$. $\forall v \in V$, we have $v \notin K^\circ$ and $p(v) \geq 1 = f(v)$. $\forall m = \alpha x_1 + \bar{m} \in M$, we have either $\alpha > 0$, then $f(m) = \alpha f(x_1 + \alpha^{-1}\bar{m}) = \alpha \leq \alpha p(x_1 + \alpha^{-1}\bar{m}) = p(m)$, where the inequality follows from the fact that $x_1 + \alpha^{-1}\bar{m} \in V$; or $\alpha \leq 0$, then $f(m) = \alpha \leq 0 \leq p(m)$. Thus, $\forall m \in M$, we have $f(m) \leq p(m)$. By the extension form of Hahn-Banach Theorem, there exists $x_* \in \mathcal{X}^*$ such

that $x_*|_M = f$ and $\langle x_*, x \rangle \leq p(x)$, $\forall x \in \mathcal{X}$. Clearly, $x_* \neq \vartheta_*$. $\forall v \in V$, $\langle x_*, v \rangle = f(v) = 1 =: c \in \mathbb{R}$. $\forall k \in \overline{K}$, we have $\langle x_*, k \rangle \leq p(k) \leq 1$. $\forall k \in K^\circ$, we have $\langle x_*, k \rangle \leq p(k) < 1$. Hence, the closed hyperplane $H := \{x \in \mathcal{X} \mid \langle x_*, x \rangle = c\}$ is the one that we seek.

Case 2: $\vartheta \notin K^\circ$. Let $x_0 \in K^\circ$. Then, by Proposition 6.39, $K_1 := K - x_0$ is a convex set. $V_1 := V - x_0$ is a linear variety. By Proposition 7.16, $K_1^\circ = K^\circ - x_0$. Then, $\vartheta \in K_1^\circ$ and $V_1 \cap K_1^\circ = \emptyset$. By Case 1, there exist $c_1 \in \mathbb{R}$ and $x_* \in \mathcal{X}^*$ with $x_* \neq \vartheta_*$ such that $\langle x_*, v_1 \rangle = c_1$, $\forall v_1 \in V_1$, $\langle x_*, k_1 \rangle < c_1$, $\forall k_1 \in K_1^\circ$, and $\langle x_*, k_1 \rangle \leq c_1$, $\forall k_1 \in \overline{K_1}$. Let $c := c_1 + \langle x_*, x_0 \rangle \in \mathbb{R}$. Then, $\forall v \in V$, $\langle x_*, v \rangle = c$. $\forall k \in K^\circ$, $\langle x_*, k \rangle < c$. By Proposition 7.16, $\overline{K_1} = \overline{K} - x_0$. Then, $\forall k \in \overline{K}$, $\langle x_*, k \rangle \leq c$. Thus, the closed hyperplane $H := \{x \in \mathcal{X} \mid \langle x_*, x \rangle = c\}$ is the one that we seek.

This completes the proof of the theorem. $\square$

Definition 8.8 Let $\mathcal{X}$ be a real normed linear space. A closed hyperplane $H := \{x \in \mathcal{X} \mid \langle x_*, x \rangle = c\}$ with $x_* \in \mathcal{X}^*$, $x_* \neq \vartheta_*$, and $c \in \mathbb{R}$ is called a supporting hyperplane of a convex set $K \subseteq \mathcal{X}$ if either $\inf_{k \in K} \langle x_*, k \rangle = c$ or $\sup_{k \in K} \langle x_*, k \rangle = c$.

Clearly, for a convex set K in a real normed linear space, if K admits interior points, then, by Mazur's Theorem, there exists a supporting hyperplane of K passing through each boundary point of K .

Theorem 8.9 (Eidelheit Separation Theorem) Let $\mathcal{X}$ be a real normed linear space and $K_1, K_2 \subseteq \mathcal{X}$ be nonempty convex sets with $K_1^\circ \neq \emptyset$ and $K_1^\circ \cap K_2 = \emptyset$. Then, there exists a closed hyperplane H that separates K_1 and K_2 , that is $\exists x_* \in \mathcal{X}^*$ with $x_* \neq \vartheta_*$ and $\exists c \in \mathbb{R}$ such that

$$\sup_{k_1 \in K_1} \langle x_*, k_1 \rangle \leq c \leq \inf_{k_2 \in K_2} \langle x_*, k_2 \rangle$$

Proof Let $K := K_1^\circ - K_2$. By Propositions 7.15 and 6.39, K is convex. Clearly, $K^\circ \neq \emptyset$ since $K_1^\circ \neq \emptyset$ and $K_2 \neq \emptyset$. Since $K_1^\circ \cap K_2 = \emptyset$, then $\vartheta \notin K$. Let $V = \{\vartheta\}$. Then, V is a linear variety and $V \cap K^\circ = \emptyset$. By Mazur's Theorem, $\exists x_* \in \mathcal{X}^*$ with $x_* \neq \vartheta_*$ such that $\langle x_*, k \rangle \leq \langle x_*, \vartheta \rangle = 0$, $\forall k \in K$. $\forall k_1 \in K_1^\circ$, $\forall k_2 \in K_2$, $k_1 - k_2 \in K$ and $\langle x_*, k_1 - k_2 \rangle \leq 0$. Then, $\langle x_*, k_1 \rangle \leq \langle x_*, k_2 \rangle$. Hence, $-\infty < \sup_{k_1 \in K_1^\circ} \langle x_*, k_1 \rangle \leq c \leq \inf_{k_2 \in K_2} \langle x_*, k_2 \rangle < +\infty$ for some $c \in \mathbb{R}$. By Proposition 8.6, $\overline{K_1} = \overline{K_1^\circ}$. By Proposition 4.13, $\forall k_1 \in \overline{K_1}$, $\langle x_*, k_1 \rangle \leq \sup_{\bar{k}_1 \in K_1^\circ} \langle x_*, \bar{k}_1 \rangle \leq c$. Hence, we have

$$\sup_{k_1 \in K_1} \langle x_*, k_1 \rangle = \sup_{k_1 \in K_1^\circ} \langle x_*, k_1 \rangle \leq c \leq \inf_{k_2 \in K_2} \langle x_*, k_2 \rangle$$

Thus, the closed hyperplane $H := \{x \in \mathcal{X} \mid \langle x_*, x \rangle = c\}$ is the one we seek. This completes the proof of the theorem. $\square$

Proposition 8.10 *Let $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathcal{X}$ be a closed convex set. Then, $\forall x_0 \in \mathcal{X} \setminus K$. Then, there exists a $x_* \in \mathcal{X}^*$ such that $\langle x_*, x_0 \rangle < \inf_{k \in K} \langle x_*, k \rangle$. Hence, K is weakly closed.*

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $K = \emptyset$; Case 2: $K \neq \emptyset$. Case 1: $K = \emptyset$. Let $x_* = \vartheta_*$. Then, $\langle x_*, x_0 \rangle = 0 < +\infty = \inf_{k \in K} \langle x_*, k \rangle$. Clearly, K is weakly closed. This case is proved.

Case 2: $K \neq \emptyset$. $\forall x_0 \in \mathcal{X} \setminus K$, by Proposition 4.10, $\text{dist}(x_0, K) = \inf_{k \in K} \|x_0 - k\| =: d \in (0, \infty) \subset \mathbb{R}$. Then, $K_1 := \mathcal{B}_{\mathcal{X}}(x_0, d)$ is a convex set with nonempty interior and $K_1 \cap K = \emptyset$. By Eidelheit Separation Theorem, there exists $x_* \in \mathcal{X}^*$ with $x_* \neq \vartheta_*$ and $\exists c \in \mathbb{R}$ such that $\sup_{k_1 \in K_1} \langle x_*, k_1 \rangle \leq c \leq \inf_{k \in K} \langle x_*, k \rangle$. By Lemma 7.75, $\langle x_*, x_0 \rangle < \sup_{k_1 \in K_1} \langle x_*, k_1 \rangle \leq c$. Then, $\langle x_*, x_0 \rangle < \inf_{k \in K} \langle x_*, k \rangle$.

Let $\bar{K}$ be the closure of K in the weak topology $\mathcal{O}_{\text{weak}}(\mathcal{X})$. $x_0 \in \{x \in \mathcal{X} \mid \langle x_*, x \rangle < c\} =: O$. Clearly, O is a weakly open set and $O \cap K = \emptyset$. Then, by Proposition 3.3, $x_0 \notin \bar{K}$. Thus, we have shown that $\forall x_0 \in \bar{K}$, $x_0 \in K$. Hence, $\bar{K} \subseteq K$. Hence, K is weakly closed by Proposition 3.3. This case is proved.

This completes the proof of the proposition. $\square$

Proposition 8.11 *Let $\mathcal{X}$ be a reflexive real normed linear space and $K \subseteq \mathcal{X}^*$ be a bounded closed convex set. Then, K is weak* compact.*

Proof Since K is bounded, then there exists $n \in \mathbb{Z}_+$ such that $K \subseteq \bar{\mathcal{B}}_{\mathcal{X}^*}(\vartheta_*, n) =: S$. By Alaoglu Theorem, S is weak* compact. Since K is closed and convex, then, by Proposition 8.10, K is weakly closed. Since $\mathcal{X}$ is reflexive, then the weak topology and the weak* topology on $\mathcal{X}^*$ coincide. Then, K is weak* closed. By Proposition 5.5, K is weak* compact. This completes the proof of the proposition. $\square$

8.3 Duality in Minimum Norm Problems

Definition 8.12 *Let $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathcal{X}$ be a nonempty convex set. The support of K is the set $K_{\text{supp}} := \{x_* \in \mathcal{X}^* \mid \sup_{k \in K} \langle x_*, k \rangle < +\infty\}$.*

Proposition 8.13 *Let $\mathcal{X}$ be a real normed linear space and $K, G \subseteq \mathcal{X}$ be nonempty convex sets. Then, the following statements hold.*

(i) $K_{\text{supp}} \subseteq \mathcal{X}^*$ is a convex cone.

(ii) *If K is closed, then $K = \bigcap_{x_* \in K_{\text{supp}}} \{x \in \mathcal{X} \mid \langle x_*, x \rangle \leq \sup_{k \in K} \langle x_*, k \rangle\}$*

$$= \bigcap_{x_* \in K_{\text{supp}}, x_* \neq \vartheta_*} \{x \in \mathcal{X} \mid \langle x_*, x \rangle \leq \sup_{k \in K} \langle x_*, k \rangle\}, \text{ that is, } K$$

equals to the intersection of all closed half-spaces containing K .

$$(iii) (K + G)_{\text{supp}} = K_{\text{supp}} \cap G_{\text{supp}}.$$

$$(iv) (K \cap G)_{\text{supp}} \supseteq K_{\text{supp}} + G_{\text{supp}}.$$

Proof (i) Clearly, $\vartheta_* \in K_{\text{supp}}$. $\forall x_{*1}, x_{*2} \in K_{\text{supp}}, \forall \alpha \in [0, \infty) \subset \mathbb{R}$, we have either $\alpha = 0$, then $\alpha x_{*1} = \vartheta_* \in K_{\text{supp}}$; or $\alpha > 0$, then $\sup_{k \in K} \langle \alpha x_*, k \rangle = \sup_{k \in K} \alpha \langle x_*, k \rangle = \alpha \sup_{k \in K} \langle x_*, k \rangle < +\infty$, where we have applied Proposition 3.81, which further implies that $\alpha x_* \in K_{\text{supp}}$. This shows that K_{supp} is a cone. Note that $\sup_{k \in K} \langle x_{*1} + x_{*2}, k \rangle = \sup_{k \in K} (\langle x_{*1}, k \rangle + \langle x_{*2}, k \rangle) \leq \sup_{k \in K} \langle x_{*1}, k \rangle + \sup_{k \in K} \langle x_{*2}, k \rangle < +\infty$, where we have applied Proposition 3.81. Then, $x_{*1} + x_{*2} \in K_{\text{supp}}$. This coupled with the fact that K_{supp} is a cone implies that K_{supp} is a convex cone.

(ii) Clearly $K \subseteq \bigcap_{x_* \in K_{\text{supp}}} \{x \in \mathcal{X} \mid \langle x_*, x \rangle \leq \sup_{k \in K} \langle x_*, k \rangle\} =: \bar{K}$. $\forall x \in \mathcal{X} \setminus K$, by Proposition 8.10, $\exists x_{*0} \in \mathcal{X}^*$ such that $\langle x_{*0}, x \rangle < \inf_{k \in K} \langle x_{*0}, k \rangle$. Then, $\sup_{k \in K} \langle -x_{*0}, k \rangle < \langle -x_{*0}, x \rangle < +\infty$. Hence, $-x_{*0} \in K_{\text{supp}}$ and $x \in \mathcal{X} \setminus \bar{K}$. This shows that $\mathcal{X} \setminus K \subseteq \mathcal{X} \setminus \bar{K}$ and $\bar{K} \subseteq K$. Hence, $K = \bar{K}$. Note that, for $\vartheta_* \in K_{\text{supp}}$, $\{x \in \mathcal{X} \mid \langle \vartheta_*, x \rangle \leq \sup_{k \in K} \langle \vartheta_*, k \rangle\} = \mathcal{X}$. Then, $K = \bigcap_{x_* \in K_{\text{supp}}, x_* \neq \vartheta_*} \{x \in \mathcal{X} \mid \langle x_*, x \rangle \leq \sup_{k \in K} \langle x_*, k \rangle\}$. Let $\dot{K}$ be the intersection of all closed half-spaces containing K . Clearly, $K \subseteq \dot{K}$. Any closed half-space containing K can be expressed as $\{x \in \mathcal{X} \mid \langle x_*, x \rangle \leq c\}$ for some $x_* \in \mathcal{X}^*$ with $x_* \neq \vartheta_*$ and for some $c \in \mathbb{R}$. Then, $c \geq \sup_{k \in K} \langle x_*, k \rangle \in \mathbb{R}$. Hence, $x_* \in K_{\text{supp}}$ and $\dot{K} \subseteq \bigcap_{x_* \in K_{\text{supp}}, x_* \neq \vartheta_*} \bigcap_{c \geq \sup_{k \in K} \langle x_*, k \rangle} \{x \in \mathcal{X} \mid \langle x_*, x \rangle \leq c\} = \bigcap_{x_* \in K_{\text{supp}}, x_* \neq \vartheta_*} \{x \in \mathcal{X} \mid \langle x_*, x \rangle \leq \sup_{k \in K} \langle x_*, k \rangle\} = K$. Hence, $K = \dot{K}$.

(iii) $\forall x_* \in K_{\text{supp}} \cap G_{\text{supp}}$, then $\sup_{k \in K} \langle x_*, k \rangle =: c_1 < +\infty$ and $\sup_{g \in G} \langle x_*, g \rangle =: c_2 < +\infty$. $\forall x \in K + G$, $x = k + g$ for some $k \in K$ and some $g \in G$. Then, $\langle x_*, x \rangle = \langle x_*, k \rangle + \langle x_*, g \rangle \leq c_1 + c_2 < +\infty$. This shows that $x_* \in (K + G)_{\text{supp}}$. Hence, $K_{\text{supp}} \cap G_{\text{supp}} \subseteq (K + G)_{\text{supp}}$.

On the other hand, $\forall x_* \in (K + G)_{\text{supp}}$, $\sup_{x \in K + G} \langle x_*, x \rangle =: c < +\infty$. Fix $k_0 \in K$ and $g_0 \in G$, since K and G are nonempty. Then, $\sup_{k \in K} \langle x_*, k \rangle = \sup_{x \in g_0 + K} \langle x_*, x \rangle - \langle x_*, g_0 \rangle \leq c - \langle x_*, g_0 \rangle < +\infty$ and $\sup_{g \in G} \langle x_*, g \rangle = \sup_{x \in k_0 + G} \langle x_*, x \rangle - \langle x_*, k_0 \rangle \leq c - \langle x_*, k_0 \rangle < +\infty$. Hence, $x_* \in K_{\text{supp}} \cap G_{\text{supp}}$. This shows that $(K + G)_{\text{supp}} \subseteq K_{\text{supp}} \cap G_{\text{supp}}$. Therefore, we have $(K + G)_{\text{supp}} = K_{\text{supp}} \cap G_{\text{supp}}$.

(iv) $\forall x_* \in K_{\text{supp}} + G_{\text{supp}}$, let $x_* = x_{*1} + x_{*2}$ with $x_{*1} \in K_{\text{supp}}$ and $x_{*2} \in G_{\text{supp}}$. $\forall x \in K \cap G$, we have $\langle x_*, x \rangle = \langle x_{*1}, x \rangle + \langle x_{*2}, x \rangle \leq \sup_{k \in K} \langle x_{*1}, k \rangle + \sup_{g \in G} \langle x_{*2}, g \rangle < +\infty$. Hence, $x_* \in (K \cap G)_{\text{supp}}$. Hence, we have $K_{\text{supp}} + G_{\text{supp}} \subseteq (K \cap G)_{\text{supp}}$.

This completes the proof of the proposition. $\square$

Definition 8.14 Let $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathcal{X}$ be a nonempty convex set. The support functional of K is $h : K_{\text{supp}} \rightarrow \mathbb{R}$ given by $h(x_*) = \sup_{k \in K} \langle x_*, k \rangle$, $\forall x_* \in K_{\text{supp}}$.

In the above definition, h takes value in $\mathbb{R}$ since $K \neq \emptyset$ and $x_* \in K_{\text{supp}}$.

Proposition 8.15 *Let $\mathcal{X}$ be a real normed linear space, $K \subseteq \mathcal{X}$ be a nonempty convex set, $h : K_{\text{supp}} \rightarrow \mathbb{R}$ be the support functional of K , and $y \in \mathcal{X}$. Then, the following statements holds.*

- (i) $\delta := \inf_{k \in K} \|y - k\| = \max_{x_* \in K_{\text{supp}}, \|x_*\| \leq 1} (\langle x_*, y \rangle - h(x_*))$, where the maximum is achieved at some $x_{*0} \in K_{\text{supp}}$ with $\|x_{*0}\| \leq 1$. If the infimum is achieved at $k_0 \in K$ then $y - k_0$ is aligned with x_{*0} and $\langle x_{*0}, k_0 \rangle = h(x_{*0})$.
- (ii) If $\exists k_0 \in K$ and $\exists x_{*0} \in K_{\text{supp}}$ with $\|x_{*0}\| = 1$ such that $y - k_0$ is aligned with x_{*0} and $\langle x_{*0}, k_0 \rangle = h(x_{*0})$, then the infimum is achieved at k_0 and the maximum is achieved at x_{*0} , that is $\delta = \|y - k_0\| = \langle x_{*0}, y \rangle - h(x_{*0})$.

Proof (i) $\forall x_* \in K_{\text{supp}}$ with $\|x_*\| \leq 1$, $\forall k \in K$, we have $\|y - k\| \geq \|x_*\| \|y - k\| \geq \langle x_*, y - k \rangle = \langle x_*, y \rangle - \langle x_*, k \rangle \geq \langle x_*, y \rangle - h(x_*)$. Hence, $\delta = \inf_{k \in K} \|y - k\| \geq \sup_{x_* \in K_{\text{supp}}, \|x_*\| \leq 1} (\langle x_*, y \rangle - h(x_*))$.

We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\delta = 0$; Case 2: $\delta > 0$. Case 1: $\delta = 0$. Take $x_{*0} = \vartheta_* \in K_{\text{supp}}$. $\|x_{*0}\| = 0 \leq 1$. Then, $\delta = 0 = \langle x_{*0}, y \rangle - h(x_{*0})$. Then, $\delta := \inf_{k \in K} \|y - k\| = \max_{x_* \in K_{\text{supp}}, \|x_*\| \leq 1} (\langle x_*, y \rangle - h(x_*))$. If the infimum is achieved at $k_0 \in K$, then $y - k_0$ is aligned with x_{*0} and $\langle x_{*0}, k_0 \rangle = 0 = h(x_{*0})$. This case is proved.

Case 2: $\delta > 0$. Since $K \neq \emptyset$, then $\delta < +\infty$. Take $K_1 = \mathcal{B}_{\mathcal{X}}(y, \delta)$. Then, K_1 is a convex set with nonempty interior. $K \cap K_1 = \emptyset$. By Eidelheit Separation Theorem, there exists $x_* \in \mathcal{X}^*$ with $x_* \neq \vartheta_*$ and $\exists c \in \mathbb{R}$ such that $\sup_{k_1 \in K_1} \langle x_*, k_1 \rangle \leq c \leq \inf_{k \in K} \langle x_*, k \rangle$. Take $x_{*0} = -\|x_*\|^{-1} x_*$. Then, by Proposition 3.81, we have $\inf_{k_1 \in K_1} \langle x_{*0}, k_1 \rangle \geq -c/\|x_*\| \geq \sup_{k \in K} \langle x_{*0}, k \rangle$. This shows, $x_{*0} \in K_{\text{supp}}$ and $\|x_{*0}\| = 1$. The above inequality is equivalent to $\langle x_{*0}, y \rangle + \inf_{x \in \mathcal{B}_{\mathcal{X}}(\vartheta, \delta)} \langle x_{*0}, x \rangle \geq h(x_{*0})$. By Lemma 7.75, we have $\langle x_{*0}, y \rangle - h(x_{*0}) \geq \delta$. Hence, we have $\delta = \langle x_{*0}, y \rangle - h(x_{*0}) = \max_{x_* \in K_{\text{supp}}, \|x_*\| \leq 1} (\langle x_*, y \rangle - h(x_*))$. If the infimum is achieved at some $k_0 \in K$, then $\delta = \|y - k_0\| = \langle x_{*0}, y \rangle - h(x_{*0}) \leq \langle x_{*0}, y \rangle - \langle x_{*0}, k_0 \rangle = \langle x_{*0}, y - k_0 \rangle \leq \|x_{*0}\| \|y - k_0\| \leq \|y - k_0\|$, where the second inequality follows from Proposition 7.72. Hence, $\langle x_{*0}, y - k_0 \rangle = \|x_{*0}\| \|y - k_0\|$, x_{*0} is aligned with $y - k_0$, and $\langle x_{*0}, k_0 \rangle = h(x_{*0})$. This case is proved.

(ii) Note that $\delta \leq \|y - k_0\| = \|x_{*0}\| \|y - k_0\| = \langle x_{*0}, y - k_0 \rangle = \langle x_{*0}, y \rangle - h(x_{*0}) \leq \delta$. Then, the result follows.

This completes the proof of the proposition. $\square$

Now, we will state a proposition that guarantees the existence of a minimizing solution to a minimum norm problem. This proposition is based on the following result.

Proposition 8.16 *Let $\mathcal{X}$ be a real normed linear space. Then, $\|\cdot\| : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$ is weakly lower semicontinuous.*

Proof By Definition 3.14, all we need to show is that, $\forall a \in \mathbb{R}$, the set $S_a := \{x \in \mathcal{X} \mid \|x\| > a\}$ is weakly open. Note that $\widetilde{S}_a$ is strongly closed and convex. Then, by Proposition 8.10, $\widetilde{S}_a$ is weakly closed. Then, S_a is weakly open. Hence, the norm is a weakly lower semicontinuous functional on $\mathcal{X}$. This completes the proof of the proposition. $\square$

Proposition 8.17 *Let $\mathcal{X}$ be a reflexive real normed linear space, $x_* \in \mathcal{X}^*$, and $K \subseteq \mathcal{X}^*$ be a nonempty closed convex set. Then, $\delta = \min_{k_* \in K} \|x_* - k_*\|$ and the minimum is achieved at some $k_{*0} \in K$.*

Proof Fix a $k_{*1} \in K \neq \emptyset$. Let $\mu := \|x_* - k_{*1}\| \in [0, \infty) \subset \mathbb{R}$ and $d = \mu + \|x_*\| + 1 \in (0, \infty) \subset \mathbb{R}$. Then, $\forall k_* \in K$ with $\|k_*\| > d$, we have $\|x_* - k_*\| \geq \|k_*\| - \|x_*\| > \mu + 1$. Thus, $\delta = \inf_{k_* \in K} \|x_* - k_*\| = \inf_{k_* \in K \cap \overline{B}_{\mathcal{X}^*}(\vartheta_*, d)} \|x_* - k_*\|$ and any k_{*0} achieving the infimum for the original problem must be in the set $K \cap \overline{B}_{\mathcal{X}^*}(\vartheta_*, d) =: K_1$. By Proposition 6.40, K_1 is bounded closed and convex. Note that $k_{*1} \in K_1 \neq \emptyset$. Then, by Proposition 8.11, K_1 is weak* compact. By Propositions 7.121, 8.16, and 3.16, we have $f : \mathcal{X}^* \rightarrow \mathbb{R}$ given by $f(k_*) = \|x_* - k_*\|$, $\forall k_* \in \mathcal{X}^*$, is lower semicontinuous. By Proposition 5.30 and Definition 3.14, there exists $k_{*0} \in K_1$ that achieves the infimum on K_1 . Such k_{*0} achieves the infimum on K . This completes the proof of the proposition. $\square$

8.4 Convex and Concave Functionals

Definition 8.18 *Let $\mathcal{X}$ be a real vector space, $C \subseteq \mathcal{X}$ be a convex set, $f : C \rightarrow \mathbb{R}$ and $g : C \rightarrow \mathbb{R}$. f is said to be convex, if $\forall x_1, x_2 \in C$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we have*

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2)$$

f is said to be strictly convex if $\forall x_1, x_2 \in C$ with $x_1 \neq x_2$ and $\forall \alpha \in (0, 1) \subset \mathbb{R}$, we have that the strict inequality holds in the above. g is said to be (strictly) concave if $-g$ is (strictly) convex.

Definition 8.19 *Let $\mathcal{X}$ be a real vector space, $C \subseteq \mathcal{X}$, and $f : C \rightarrow \mathbb{R}$. The epigraph of f over C is the set*

$$[f, C] := \{(r, x) \in \mathbb{R} \times \mathcal{X} \mid x \in C, f(x) \leq r\}$$

Proposition 8.20 *Let $\mathcal{X}$ be a real vector space, $C \subseteq \mathcal{X}$ be convex, and $f : C \rightarrow \mathbb{R}$. Then, f is convex if, and only if, the epigraph $[f, C]$ is convex.*

Proof “Sufficiency” $\forall x_1, x_2 \in C, \forall \alpha \in [0, 1] \subset \mathbb{R}$, we have $(f(x_1), x_1), (f(x_2), x_2) \in [f, C]$. By the convexity of the epigraph, then $\alpha(f(x_1), x_1) + (1 - \alpha)(f(x_2), x_2) = (\alpha f(x_1) + (1 - \alpha)f(x_2), \alpha x_1 + (1 - \alpha)x_2) \in [f, C]$. Then, $\alpha f(x_1) + (1 - \alpha)f(x_2) \geq f(\alpha x_1 + (1 - \alpha)x_2)$. Hence, f is convex.

“Necessity” Let f be convex. $\forall (r_1, x_1), (r_2, x_2) \in [f, C], \forall \alpha \in [0, 1] \subset \mathbb{R}$, we have $r_1 \geq f(x_1)$ and $r_2 \geq f(x_2)$. By the convexity of f and C , we have $\alpha x_1 + (1 - \alpha)x_2 \in C$ and $\alpha r_1 + (1 - \alpha)r_2 \geq \alpha f(x_1) + (1 - \alpha)f(x_2) \geq f(\alpha x_1 + (1 - \alpha)x_2)$. Then, we have $\alpha(r_1, x_1) + (1 - \alpha)(r_2, x_2) = (\alpha r_1 + (1 - \alpha)r_2, \alpha x_1 + (1 - \alpha)x_2) \in [f, C]$. Hence, the epigraph is convex.

This completes the proof of the proposition. $\square$

Proposition 8.21 Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be nonempty, and $f : C \rightarrow \mathbb{R}$. Then, $V([f, C]) = \mathbb{R} \times V(C)$.

Proof We will first show that $v([f, C]) = \mathbb{R} \times v(C)$. Fix $x_0 \in C$. Then, $(f(x_0), x_0) \in [f, C]$. $\forall (r, x) \in v([f, C]) = (f(x_0), x_0) + \text{span}([f, C] - (f(x_0), x_0))$, where the equality follows from Proposition 6.37, then $\exists n \in \mathbb{Z}_+, \exists (r_1, x_1), \dots, (r_n, x_n) \in [f, C]$, and $\exists \alpha_1, \dots, \alpha_n \in \mathbb{R}$ such that $(r - f(x_0), x - x_0) = \sum_{i=1}^n \alpha_i (r_i - f(x_0), x_i - x_0)$. Then, $x \in v(C) = x_0 + \text{span}(C - x_0)$ and $(r, x) \in \mathbb{R} \times v(C)$. Hence, $v([f, C]) \subseteq \mathbb{R} \times v(C)$.

On the other hand, $\forall (r, x) \in \mathbb{R} \times v(C)$, then, by Proposition 6.37, $x \in v(C) = x_0 + \text{span}(C - x_0)$. Then, $\exists n \in \mathbb{Z}_+, \exists x_1, \dots, x_n \in C$, and $\exists \alpha_1, \dots, \alpha_n \in \mathbb{R}$ such that $x - x_0 = \sum_{i=1}^n \alpha_i (x_i - x_0)$. $\exists r_0 \in (0, \infty) \subset \mathbb{R}$ such that $r_0 > f(x_0)$. Then, $(r_0, x_0) \in [f, C]$. Let $\alpha_0 = (r - f(x_0) - \sum_{i=1}^n \alpha_i (f(x_i) - f(x_0))) / (r_0 - f(x_0))$. Now it is easy to check that $(r - f(x_0), x - x_0) = \sum_{i=1}^n \alpha_i (f(x_i) - f(x_0), x_i - x_0) + \alpha_0 (r_0 - f(x_0), x_0 - x_0)$. Note that $(f(x_i), x_i) \in [f, C], \forall i = 1, \dots, n$. Then, we have $(r, x) \in (f(x_0), x_0) + \text{span}([f, C] - (f(x_0), x_0)) = v([f, C])$. Hence, $\mathbb{R} \times v(C) \subseteq v([f, C])$. Therefore, we have $v([f, C]) = \mathbb{R} \times v(C)$.

By Proposition 7.18, $V([f, C]) = \overline{v([f, C])} = \mathbb{R} \times v(C)$. By Proposition 4.13 and 3.67, we have $\mathbb{R} \times v(C) = \mathbb{R} \times V(C)$. Therefore, we have $V([f, C]) = \mathbb{R} \times V(C)$. This completes the proof of the proposition. $\square$

Proposition 8.22 Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be convex, $f : C \rightarrow \mathbb{R}$ be convex, and ${}^\circ C \neq \emptyset$. Then, $[f, C]$ has a relative interior point (r_0, x_0) if, and only if, f is continuous at x_0 , $x_0 \in {}^\circ C$, and $r_0 \in (f(x_0), +\infty) \subset \mathbb{R}$.

Proof “Sufficiency” Fix $x_0 \in {}^\circ C$ and $r_0 \in (f(x_0), +\infty) \subset \mathbb{R}$. Let f be continuous at x_0 . Then, for $\epsilon = (r_0 - f(x_0))/2 \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, r_0 - f(x_0) - \epsilon) \subset \mathbb{R}$ such that $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta) \cap V(C)$, we have $x \in C$ and $|f(x) - f(x_0)| < \epsilon$. Clearly, $(r_0, x_0) \in [f, C]$. $\forall (r, x) \in \mathcal{B}_{\mathbb{R} \times \mathcal{X}}((r_0, x_0), \delta) \cap V([f, C]) = \mathcal{B}_{\mathbb{R} \times \mathcal{X}}((r_0, x_0), \delta) \cap (\mathbb{R} \times V(C))$, we have $x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta) \cap V(C)$ and $r \in \mathcal{B}(r_0, \delta)$. Then, $x \in C$ and $r > r_0 - \delta \geq f(x_0) + \epsilon > f(x)$. Hence, $(r, x) \in [f, C]$. This shows that $(r_0, x_0) \in {}^\circ [f, C]$.

“Necessity” Let $(r_0, x_0) \in {}^\circ[f, C]$. Then, $\exists \epsilon_0 \in (0, \infty) \subset \mathbb{R}$ such that $(\mathcal{B}_{\mathbf{R}}(r_0, \epsilon_0) \times \mathcal{B}_{\mathcal{X}}(x_0, \epsilon_0)) \cap V([f, C]) \subseteq [f, C]$. By Proposition 8.21, we have $\mathcal{B}_{\mathbf{R}}(r_0, \epsilon_0) \times (\mathcal{B}_{\mathcal{X}}(x_0, \epsilon_0) \cap V(C)) \subseteq [f, C]$. Therefore, $\mathcal{B}_{\mathcal{X}}(x_0, \epsilon_0) \cap V(C) \subseteq C$, $x_0 \in {}^\circ C$, $f(x) \leq r_0 - \epsilon_0$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \epsilon_0) \cap V(C)$, and $f(x_0) \leq r_0 - \epsilon_0 < r_0$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, Let $\delta = \min\{\epsilon/(r_0 - f(x_0)), 1\} \in (0, 1] \subset \mathbb{R}$. $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta\epsilon_0) \cap C$, we have $x, x_0, x_0 + \delta^{-1}(x - x_0), x_0 - \delta^{-1}(x - x_0) \in \mathcal{B}_{\mathcal{X}}(x_0, \epsilon_0) \cap V(C) \subseteq C$. By the convexity of f , we have

$$\begin{aligned} f(x) &= f((1 - \delta)x_0 + \delta(x_0 + \delta^{-1}(x - x_0))) \\ &\leq (1 - \delta)f(x_0) + \delta f(x_0 + \delta^{-1}(x - x_0)) \\ &\leq f(x_0) + (r_0 - \epsilon_0 - f(x_0))\delta < f(x_0) + \epsilon \\ f(x_0) &= f\left(\frac{1}{1 + \delta}x + \frac{\delta}{1 + \delta}(x_0 - \delta^{-1}(x - x_0))\right) \\ &\leq \frac{1}{1 + \delta}f(x) + \frac{\delta}{1 + \delta}f(x_0 - \delta^{-1}(x - x_0)) \\ &\leq \frac{1}{1 + \delta}f(x) + \frac{\delta}{1 + \delta}(r_0 - \epsilon_0) \\ \implies (1 + \delta)f(x_0) &\leq f(x) + (r_0 - \epsilon_0)\delta \\ \implies f(x) &\geq f(x_0) - (r_0 - f(x_0) - \epsilon_0)\delta > f(x_0) - \epsilon \end{aligned}$$

Hence, $|f(x) - f(x_0)| < \epsilon$. This shows that f is continuous at x_0 .

This completes the proof of the proposition. $\square$

Proposition 8.23 *Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be convex, $x_0 \in {}^\circ C \neq \emptyset$, and $f : C \rightarrow \mathbb{R}$ be convex. If f is continuous at x_0 , then f is continuous at x , $\forall x \in {}^\circ C$.*

Proof Fix any $x \in {}^\circ C$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x = x_0$; Case 2: $x \neq x_0$. Case 1: $x = x_0$. Then f is continuous at x . Case 2: $x \neq x_0$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the continuity of f at x_0 and the fact that $x_0 \in {}^\circ C$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $|f(y) - f(x_0)| < \epsilon$, $\forall y \in \mathcal{B}_{\mathcal{X}}(x_0, \delta) \cap V(C) \subseteq C$. Since $x \in {}^\circ C$, then $\exists \delta_1 \in (0, \delta] \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{X}}(x, \delta_1) \cap V(C) \subseteq C$. Take $\beta = \frac{\|x - x_0\| + \delta_1/2}{\|x - x_0\|} \in (1, \infty) \subset \mathbb{R}$. Then, $\beta(x - x_0) + x_0 \in \mathcal{B}_{\mathcal{X}}(x, \delta_1) \cap V(C) \subseteq C$. $\forall y \in \mathcal{B}_{\mathcal{X}}(x, (\beta - 1)\delta_1/\beta) \cap C$, $y = (1 - 1/\beta)(\beta/(\beta - 1)(y - x) + x_0) + (1/\beta)(\beta(x - x_0) + x_0)$. Note that $\beta/(\beta - 1)(y - x) + x_0 \in \mathcal{B}_{\mathcal{X}}(x_0, \delta) \cap V(C) \subseteq C$. By the convexity of f , we have

$$\begin{aligned} f(y) &\leq (1 - 1/\beta)f(\beta/(\beta - 1)(y - x) + x_0) + (1/\beta)f(\beta(x - x_0) + x_0) \\ &< (1 - 1/\beta)(f(x_0) + \epsilon) + (1/\beta)f(\beta(x - x_0) + x_0) =: r_1 \end{aligned}$$

Define $r := r_1 + (\beta - 1)\delta_1/\beta$. Then, by Proposition 8.21, $\mathcal{B}_{\mathbf{R} \times \mathcal{X}}((r, x), (\beta - 1)\delta_1/\beta) \cap V([f, C]) = \mathcal{B}_{\mathbf{R} \times \mathcal{X}}((r, x), (\beta - 1)\delta_1/\beta) \cap (\mathbb{R} \times V(C)) \subseteq [f, C]$. Hence, $(r, x) \in {}^\circ[f, C]$. By Proposition 8.22, f is continuous at x . This completes the proof of the proposition. $\square$

Proposition 8.24 *Let $\mathcal{X}$ be a finite-dimensional real normed linear space, $C \subseteq \mathcal{X}$ be a convex set, and $f : C \rightarrow \mathbb{R}$ be convex. Then, $\forall x_0 \in {}^\circ C$, f is continuous at x_0 .*

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: ${}^\circ C = \emptyset$; Case 2: ${}^\circ C \neq \emptyset$. Case 1: ${}^\circ C = \emptyset$. This is trivial.

Case 2: ${}^\circ C \neq \emptyset$. Fix any $x_0 \in {}^\circ C$. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\bar{K} := \bar{\mathcal{B}}_{\mathcal{X}}(x_0, \delta) \cap V(C) \subseteq C$. Let $M := V(C) - x_0$, which is a closed subspace. Let $n \in \mathbb{Z}_+$ be the dimension of M , which is well-defined by Theorem 6.51. We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: $n = 0$; Case 2b: $n \in \mathbb{N}$. Case 2a: $n = 0$. Then, $C = \{x_0\}$. Clearly, f is continuous at x_0 .

Case 2b: $n \in \mathbb{N}$. Let $\{e_1, \dots, e_n\} \subseteq M$ be a basis in M such that $\|e_i\| = 1$, $\forall i = 1, \dots, n$. Then, $K := \text{co}(\{x_0 \pm \delta e_i \mid i = 1, \dots, n\}) \subseteq \text{co}(\bar{K}) \subseteq C$ since C is convex. $\forall x \in K$, by Proposition 6.43, $x = \sum_{i=1}^n \alpha_i (x_0 + \delta e_i) + \sum_{i=1}^n \beta_i (x_0 - \delta e_i)$ for some $\alpha_i, \beta_i \in [0, 1] \subset \mathbb{R}$, $i = 1, \dots, n$, with $\sum_{i=1}^n (\alpha_i + \beta_i) = 1$. By the convexity of f , we have

$$\begin{aligned} f(x) &\leq \sum_{i=1}^n \alpha_i f(x_0 + \delta e_i) + \sum_{i=1}^n \beta_i f(x_0 - \delta e_i) \\ &\leq \sum_{i=1}^n |f(x_0 + \delta e_i)| + \sum_{i=1}^n |f(x_0 - \delta e_i)| =: r_1 \end{aligned}$$

It is easy to see that $p(m) := \sum_{i=1}^n |\alpha_i|$, $\forall m = \sum_{i=1}^n \alpha_i e_i \in M$ defines a norm on M . By Theorem 7.38, $\exists \xi \in [1, \infty) \subset \mathbb{R}$ such that $p(m)/\xi \leq \|m\| \leq \xi p(m)$, $\forall m \in M$. $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta/\xi) \cap V(C)$, we have $x = x_0 + \sum_{i=1}^n \alpha_i \delta e_i$ for some $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. Then, $\delta/\xi > \|x - x_0\| \geq p(\sum_{i=1}^n \alpha_i \delta e_i)/\xi = \delta/\xi \sum_{i=1}^n |\alpha_i|$. Then, we have $\sum_{i=1}^n |\alpha_i| < 1$. Then, x can be expressed as a convex combination of vectors in $\{x_0 \pm \delta e_i \mid i = 1, \dots, n\}$ and $x \in K$. Hence, $\mathcal{B}_{\mathcal{X}}(x_0, \delta/\xi) \cap V(C) \subseteq K \subseteq C$. Take $r = \delta/\xi + r_1$. It is easy to show that $\mathcal{B}_{\mathbb{R} \times \mathcal{X}}((r, x_0), \delta/\xi) \cap V([f, C]) \subseteq [f, C]$, by Proposition 8.21. Then, $(r, x_0) \in {}^\circ[f, C]$. By Proposition 8.22, f is continuous at x_0 .

This completes the proof of the proposition. $\square$

Proposition 8.25 *Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$, and $f : C \rightarrow \mathbb{R}$. Then, the following statements hold.*

- (i) *If $[f, C]$ is closed, then, f is lower semicontinuous.*
- (ii) *If C is closed and f is lower semicontinuous, then $[f, C]$ is closed.*

Proof (i) $\forall a \in \mathbb{R}$, $V_a := \{(a, x) \in \mathbb{R} \times \mathcal{X} \mid x \in \mathcal{X}\}$ is closed. Then, $[f, C] \cap V_a = \{(a, x) \in \mathbb{R} \times \mathcal{X} \mid x \in C, f(x) \leq a\}$ is closed. Hence, by Proposition 4.13, $T_a := \{x \in C \mid f(x) \leq a\}$ is closed. Note that

$$\{x \in C \mid -f(x) < a\} = C \setminus \{x \in C \mid f(x) \leq -a\} = C \setminus T_{-a}$$

Clearly, $C \setminus T_{-a}$ is open in the subset topology of C . Then, $-f$ is upper semicontinuous and f is lower semicontinuous.

(ii) $\forall (r_0, x_0) \in \overline{[f, C]}$, by Proposition 4.13, $\exists ((r_n, x_n))_{n=1}^\infty \subseteq [f, C]$ such that $\lim_{n \in \mathbb{N}} (r_n, x_n) = (r_0, x_0)$. By Proposition 3.67, we have $\lim_{n \in \mathbb{N}} r_n = r_0$ and $\lim_{n \in \mathbb{N}} x_n = x_0$. By Definition 8.19, we have $(x_n)_{n=1}^\infty \subseteq C$ and $r_n \geq f(x_n)$, $\forall n \in \mathbb{N}$. By Proposition 4.13, $x_0 \in \overline{C} = C$.

We will distinguish two exhaustive and mutually exclusive cases: Case 1: there exists infinitely many $n \in \mathbb{N}$ such that $x_n = x_0$; Case 2: there exists only finitely many $n \in \mathbb{N}$ such that $x_n = x_0$. Case 1: there exists infinitely many $n \in \mathbb{N}$ such that $x_n = x_0$. Then, without loss of generality, assume $x_n = x_0$, $\forall n \in \mathbb{N}$. Then, $r_n \geq f(x_0)$, $\forall n \in \mathbb{N}$. Hence, $f(x_0) \leq \lim_{n \in \mathbb{N}} r_n = r_0$. Then, $(r_0, x_0) \in [f, C]$.

Case 2: there exists only finitely many $n \in \mathbb{N}$ such that $x_n = x_0$. Then, without loss of generality, assume $x_n \neq x_0$, $\forall n \in \mathbb{N}$. Then, x_0 is an accumulation point of C . By Propositions 3.86 and 3.85 and Definition 3.14, $-f(x_0) \geq \limsup_{x \rightarrow x_0} (-f(x)) = -\liminf_{x \rightarrow x_0} f(x)$. By Proposition 3.87, we have $f(x_0) \leq \liminf_{x \rightarrow x_0} f(x) \leq \liminf_{n \in \mathbb{N}} f(x_n) \leq \liminf_{n \in \mathbb{N}} r_n = r_0$. Then, $(r_0, x_0) \in [f, C]$.

In both cases, we have $(r_0, x_0) \in [f, C]$. Then, $\overline{[f, C]} \subseteq [f, C]$. By Proposition 3.3, $[f, C]$ is closed.

This completes the proof of the proposition. $\square$

Proposition 8.26 *Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be convex, and $f : C \rightarrow \mathbb{R}$ be convex. Assume that $[f, C]$ is closed. Then, f is weakly lower semicontinuous.*

Proof $\forall a \in \mathbb{R}$, $V_a := \{(a, x) \in \mathbb{R} \times \mathcal{X} \mid x \in \mathcal{X}\}$ is closed. Then, $[f, C] \cap V_a = \{(a, x) \in \mathbb{R} \times \mathcal{X} \mid x \in C, f(x) \leq a\}$ is closed. Hence, by Proposition 4.13, $T_a := \{x \in C \mid f(x) \leq a\}$ is closed. Since f is convex, then T_a is convex. By Proposition 8.10, T_a is weakly closed. Note that

$$\{x \in C \mid -f(x) < a\} = C \setminus \{x \in C \mid f(x) \leq -a\} = C \setminus T_{-a}$$

Clearly, $C \setminus T_{-a}$ is weakly open in the subset topology of C . Then, $-f$ is weakly upper semicontinuous and f is weakly lower semicontinuous. This completes the proof of the proposition. $\square$

8.5 Conjugate Convex Functionals

Definition 8.27 *Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be a nonempty convex set, and $f : C \rightarrow \mathbb{R}$ be convex. The conjugate set C_{conj} is defined as*

$$C_{\text{conj}} := \{x_* \in \mathcal{X}^* \mid \sup_{x \in C} (\langle x_*, x \rangle - f(x)) < +\infty\}$$

and the functional $f_{\text{conj}} : C_{\text{conj}} \rightarrow \mathbb{R}$ conjugate to f is defined by

$$f_{\text{conj}}(x_*) = \sup_{x \in C} (\langle x_*, x \rangle - f(x)), \quad \forall x_* \in C_{\text{conj}}$$

We will use a compact notation $[f, C]_{\text{conj}}$ for $[f_{\text{conj}}, C_{\text{conj}}]$.

In the above definition, f_{conj} takes value in $\mathbb{R}$ since $x_* \in C_{\text{conj}}$ and $C \neq \emptyset$.

Proposition 8.28 *Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be a nonempty convex set, $f : C \rightarrow \mathbb{R}$ be convex. Then, $C_{\text{conj}} \subseteq \mathcal{X}^*$ is convex, $f_{\text{conj}} : C_{\text{conj}} \rightarrow \mathbb{R}$ is convex, and $[f_{\text{conj}}, C_{\text{conj}}] \subseteq \mathbb{R} \times \mathcal{X}^*$ is a closed convex set.*

Proof $\forall x_{*1}, x_{*2} \in C_{\text{conj}}, \forall \alpha \in [0, 1] \subset \mathbb{R}$, let $M_i := f_{\text{conj}}(x_{*i}) = \sup_{x \in C} (\langle x_{*i}, x \rangle - f(x)) \in \mathbb{R}$, $i = 1, 2$. $\forall x \in C$, we have $\langle \alpha x_{*1} + (1 - \alpha)x_{*2}, x \rangle - f(x) = \alpha (\langle x_{*1}, x \rangle - f(x)) + (1 - \alpha) (\langle x_{*2}, x \rangle - f(x)) \leq \alpha M_1 + (1 - \alpha)M_2 < +\infty$. Then, $\sup_{x \in C} (\langle \alpha x_{*1} + (1 - \alpha)x_{*2}, x \rangle - f(x)) \leq \alpha M_1 + (1 - \alpha)M_2$. Hence, $\alpha x_{*1} + (1 - \alpha)x_{*2} \in C_{\text{conj}}$ and C_{conj} is convex.

The above implies that $f_{\text{conj}}(\alpha x_{*1} + (1 - \alpha)x_{*2}) \leq \alpha f_{\text{conj}}(x_{*1}) + (1 - \alpha)f_{\text{conj}}(x_{*2})$. Hence, f_{conj} is convex.

By Proposition 8.20, $[f_{\text{conj}}, C_{\text{conj}}]$ is convex. $\forall (s_0, x_{*0}) \in \overline{[f_{\text{conj}}, C_{\text{conj}}]}$, by Proposition 4.13, there exists $((s_k, x_{*k}))_{k=1}^\infty \subseteq [f_{\text{conj}}, C_{\text{conj}}]$ such that $\lim_{k \in \mathbb{N}} (s_k, x_{*k}) = (s_0, x_{*0})$. By Proposition 3.67, we have $\lim_{k \in \mathbb{N}} s_k = s_0$ and $\lim_{k \in \mathbb{N}} x_{*k} = x_{*0}$. $\forall x \in C$, $\forall k \in \mathbb{N}$, since $(s_k, x_{*k}) \in [f_{\text{conj}}, C_{\text{conj}}]$, then, $\langle x_{*k}, x \rangle - f(x) \leq f_{\text{conj}}(x_{*k}) \leq s_k$. By Propositions 7.72 and 3.66, we have $\langle x_{*0}, x \rangle - f(x) = \lim_{k \in \mathbb{N}} (\langle x_{*k}, x \rangle - f(x)) \leq \lim_{k \in \mathbb{N}} s_k = s_0 < +\infty$. Hence, $x_{*0} \in C_{\text{conj}}$ and $f_{\text{conj}}(x_{*0}) \leq s_0$. Then, $(s_0, x_{*0}) \in [f_{\text{conj}}, C_{\text{conj}}]$. This shows that $\overline{[f_{\text{conj}}, C_{\text{conj}}]} \subseteq [f_{\text{conj}}, C_{\text{conj}}]$. By Proposition 3.3, $[f_{\text{conj}}, C_{\text{conj}}]$ is closed.

This completes the proof of the proposition. $\square$

Geometric interpretation of $f_{\text{conj}} : C_{\text{conj}} \rightarrow \mathbb{R}$. Fix $x_* \in C_{\text{conj}}$. $\forall (r, x) \in [f, C]$, we have $f(x) \leq r$. Then, $\langle (-1, x_*), (r, x) \rangle = \langle x_*, x \rangle - r \leq f_{\text{conj}}(x_*)$. It is easy to recognize that $\sup_{(r, x) \in [f, C]} \langle (-1, x_*), (r, x) \rangle = f_{\text{conj}}(x_*)$. Hence, the closed hyperplane $H := \{(r, x) \in \mathbb{R} \times \mathcal{X} \mid \langle (-1, x_*), (r, x) \rangle = f_{\text{conj}}(x_*)\}$ is a supporting hyperplane of $[f, C]$.

Proposition 8.29 *Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be a nonempty convex set, and $f : C \rightarrow \mathbb{R}$ be a convex functional. Assume that $[f, C] =: K \subseteq \mathbb{R} \times \mathcal{X}$ is closed. Then, $\exists x_{*0} \in \mathcal{X}^*$ such that*

$$\sup_{(r, x) \in K} \langle (-1, x_{*0}), (r, x) \rangle < +\infty$$

Then, $x_{*0} \in C_{\text{conj}} \neq \emptyset$.

Proof Fix $x_0 \in C \neq \emptyset$. By Proposition 8.20, K is convex. Clearly, $K \neq \emptyset$. By the assumption of the proposition, K is closed. Let $r_0 := f(x_0) - 1 \in \mathbb{R}$. Then, $(r_0, x_0) \notin K$. By Proposition 8.10 and Example 7.76, $\exists (\bar{s}_0, \bar{x}_{*0}) \in \mathbb{R} \times \mathcal{X}^*$ such that

$$\langle (\bar{s}_0, \bar{x}_{*0}), (r_0, x_0) \rangle < \inf_{(r, x) \in K} \langle (\bar{s}_0, \bar{x}_{*0}), (r, x) \rangle \quad (8.1)$$

Since $(f(x_0), x_0) \in K$, then $\bar{s}_0 r_0 + \langle \bar{x}_{*0}, x_0 \rangle < \bar{s}_0 f(x_0) + \langle \bar{x}_{*0}, x_0 \rangle$, which implies that $\bar{s}_0 > 0$. Let $x_{*0} = -\bar{s}_0^{-1} \bar{x}_{*0} \in \mathcal{X}^*$. Then, (8.1) is equivalent to

$$\langle (-1, x_{*0}), (r_0, x_0) \rangle > \sup_{(r,x) \in K} \langle (-1, x_{*0}), (r, x) \rangle$$

The above implies that $\langle (-1, x_{*0}), (r_0, x_0) \rangle > \sup_{x \in C} (\langle x_{*0}, x \rangle - f(x))$. Hence, $x_{*0} \in C_{\text{conj}} \neq \emptyset$.

This completes the proof of the proposition. $\square$

Proposition 8.30 *Let $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathbb{R} \times \mathcal{X} =: \mathcal{W}$ be a closed convex set.*

Assume that there exists a nonvertical hyperplane such that K is contained in one of the half-spaces associated with the hyperplane, that is $\exists (s_1, x_{*1}) \in \mathcal{W}^* = \mathbb{R} \times \mathcal{X}^*$ with $s_1 \neq 0$ such that $\sup_{(r,x) \in K} \langle (s_1, x_{*1}), (r, x) \rangle < +\infty$.

Then, $\forall (r_0, x_0) \in \mathcal{W} \setminus K$, there exists a nonvertical hyperplane separating (r_0, x_0) and K , that is, $\exists x_{*0} \in \mathcal{X}^*$ such that either

$$\sup_{(r,x) \in K} \langle (-1, x_{*0}), (r, x) \rangle < \langle (-1, x_{*0}), (r_0, x_0) \rangle \quad (8.2a)$$

or

$$\inf_{(r,x) \in K} \langle (-1, x_{*0}), (r, x) \rangle > \langle (-1, x_{*0}), (r_0, x_0) \rangle \quad (8.2b)$$

Proof Fix $(r_0, x_0) \in \mathcal{W} \setminus K$. By Proposition 8.10, $\exists (\bar{s}_0, \bar{x}_{*0}) \in \mathbb{R} \times \mathcal{X}^*$ such that

$$\langle (\bar{s}_0, \bar{x}_{*0}), (r_0, x_0) \rangle < \inf_{(r,x) \in K} \langle (\bar{s}_0, \bar{x}_{*0}), (r, x) \rangle \quad (8.3)$$

We will distinguish three exhaustive and mutually exclusive cases: Case 1: $\bar{s}_0 > 0$; Case 2: $\bar{s}_0 < 0$; Case 3: $\bar{s}_0 = 0$. Case 1: $\bar{s}_0 > 0$. Let $x_{*0} = -\bar{s}_0^{-1} \bar{x}_{*0} \in \mathcal{X}^*$. Then, (8.3) is equivalent to

$$-r_0 + \langle x_{*0}, x_0 \rangle > \sup_{(r,x) \in K} (-r + \langle x_{*0}, x \rangle)$$

Hence, (8.2a) holds.

Case 2: $\bar{s}_0 < 0$. Let $x_{*0} = -\bar{s}_0^{-1} \bar{x}_{*0} \in \mathcal{X}^*$. Then, (8.3) is equivalent to

$$-r_0 + \langle x_{*0}, x_0 \rangle < \inf_{(r,x) \in K} (-r + \langle x_{*0}, x \rangle)$$

Hence, (8.2b) holds.

Case 3: $\bar{s}_0 = 0$. We will further distinguish two exhaustive and mutually exclusive cases: Case 3a: $K = \emptyset$; Case 3b: $K \neq \emptyset$. Case 3a: $K = \emptyset$. Take $x_{*0} = \vartheta \mathcal{X}^*$. Then,

$$-r_0 > -\infty = \sup_{(r,x) \in K} \langle (-1, x_{*0}), (r, x) \rangle$$

Hence, (8.2a) holds.

Case 3b: $K \neq \emptyset$. Define

$$\begin{aligned} M_1 &:= \sup_{(r,x) \in K} \langle (s_1, x_{*1}), (r, x) \rangle - \langle (s_1, x_{*1}), (r_0, x_0) \rangle \in \mathbb{R} \\ M_2 &:= \inf_{(r,x) \in K} \langle (\bar{s}_0, \bar{x}_{*0}), (r, x) \rangle - \langle (\bar{s}_0, \bar{x}_{*0}), (r_0, x_0) \rangle \\ &\in (0, \infty) \subset \mathbb{R} \end{aligned}$$

Let $\delta := M_2/(1+|M_1|) \in (0, \infty) \subset \mathbb{R}$ and $(\tilde{s}_0, \tilde{x}_{*0}) = (\bar{s}_0, \bar{x}_{*0}) - \delta(s_1, x_{*1}) \in \mathbb{R} \times \mathcal{X}^*$. Then, we have

$$\begin{aligned} &\inf_{(r,x) \in K} \langle (\tilde{s}_0, \tilde{x}_{*0}), (r, x) \rangle - \langle (\tilde{s}_0, \tilde{x}_{*0}), (r_0, x_0) \rangle \\ &= \inf_{(r,x) \in K} (\langle (\bar{s}_0, \bar{x}_{*0}), (r, x) \rangle - \delta \langle (s_1, x_{*1}), (r, x) \rangle) \\ &\quad - \langle (\tilde{s}_0, \tilde{x}_{*0}), (r_0, x_0) \rangle \\ &\geq \inf_{(r,x) \in K} \langle (\bar{s}_0, \bar{x}_{*0}), (r, x) \rangle + \inf_{(r,x) \in K} (-\delta \langle (s_1, x_{*1}), (r, x) \rangle) \\ &\quad - \langle (\tilde{s}_0, \tilde{x}_{*0}), (r_0, x_0) \rangle \\ &= \inf_{(r,x) \in K} \langle (\bar{s}_0, \bar{x}_{*0}), (r, x) \rangle - \delta \sup_{(r,x) \in K} \langle (s_1, x_{*1}), (r, x) \rangle \\ &\quad - \langle (\tilde{s}_0, \tilde{x}_{*0}), (r_0, x_0) \rangle \\ &= M_2 - \delta M_1 > 0 \end{aligned}$$

where we have applied Proposition 3.81 in the above. Hence, we have obtained an alternative $(\bar{s}_0, \bar{x}_{*0}) := (\tilde{s}_0, \tilde{x}_{*0})$ such that (8.3) holds. For this alternative pair, $\tilde{s}_0 = -\delta s_1 \neq 0$. Hence, this case can be solved by Case 1 or Case 2 with the alternative pair.

This completes the proof of the proposition. $\square$

Definition 8.31 Let $\mathcal{X}$ be a real normed linear space, $\Gamma \subseteq \mathcal{X}^*$ be a nonempty convex set, and $\varphi : \Gamma \rightarrow \mathbb{R}$ be convex. The pre-conjugate set ${}_{\text{conj}}\Gamma$ is defined as

$${}_{\text{conj}}\Gamma := \{x \in \mathcal{X} \mid \sup_{x_* \in \Gamma} (\langle x_*, x \rangle - \varphi(x_*)) < +\infty\}$$

and the functional ${}_{\text{conj}}\varphi : {}_{\text{conj}}\Gamma \rightarrow \mathbb{R}$ pre-conjugate to φ is defined by

$${}_{\text{conj}}\varphi(x) = \sup_{x_* \in \Gamma} (\langle x_*, x \rangle - \varphi(x_*)), \quad \forall x \in {}_{\text{conj}}\Gamma$$

We will use a compact notation ${}_{\text{conj}}[\varphi, \Gamma]$ for $[{}_{\text{conj}}\varphi, {}_{\text{conj}}\Gamma]$.

In the above definition, ${}_{\text{conj}}\varphi$ takes value in $\mathbb{R}$ since $x \in {}_{\text{conj}}\Gamma$ and $\Gamma \neq \emptyset$.

Next, we state two duality results regarding conjugate convex functionals.

Proposition 8.32 *Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be a nonempty convex set, and $f : C \rightarrow \mathbb{R}$ be a convex functional. Assume that $[f, C]$ is closed. Then, $[f, C] = \text{conj}[[f, C]_{\text{conj}}]$. Therefore, $\forall x_0 \in C$, $f(x_0) = \sup_{x_* \in C_{\text{conj}}} (\langle x_*, x_0 \rangle - f_{\text{conj}}(x_*))$.*

Proof Since $C \neq \emptyset$, then C_{conj} and f_{conj} are well-defined. By Proposition 8.29, $C_{\text{conj}} \neq \emptyset$. By Proposition 8.28, C_{conj} is convex and f_{conj} is convex. Then, $\text{conj}(C_{\text{conj}})$ and $\text{conj}(f_{\text{conj}})$ are well defined. Thus, $\text{conj}[[f, C]_{\text{conj}}]$ is a well-defined set.

We will first show that $[f, C] \subseteq \text{conj}[f_{\text{conj}}, C_{\text{conj}}] = \text{conj}[[f, C]_{\text{conj}}]$. Fix any $(r, x) \in [f, C]$. $\forall x_* \in C_{\text{conj}}$, we have $f_{\text{conj}}(x_*) \geq \langle x_*, x \rangle - f(x)$. Hence, we have $r \geq f(x) \geq \langle x_*, x \rangle - f_{\text{conj}}(x_*)$. Thus, we have

$$r \geq \sup_{x_* \in C_{\text{conj}}} \langle x_*, x \rangle - f_{\text{conj}}(x_*) = \text{conj}(f_{\text{conj}})(x)$$

Hence, $(r, x) \in \text{conj}[f_{\text{conj}}, C_{\text{conj}}]$ and $[f, C] \subseteq \text{conj}[f_{\text{conj}}, C_{\text{conj}}]$.

On the other hand, fix any $(r_0, x_0) \in (\mathbb{R} \times \mathcal{X}) \setminus [f, C]$. By Proposition 8.29, $\exists \bar{x}_{*0} \in \mathcal{X}^*$ such that

$$\sup_{(r, x) \in [f, C]} \langle (-1, \bar{x}_{*0}), (r, x) \rangle < +\infty$$

Note that $[f, C]$ is convex, by Proposition 8.20. By Proposition 8.30, there exists a nonvertical hyperplane separating (r_0, x_0) and $K := [f, C]$, that is, $\exists x_{*0} \in \mathcal{X}^*$ such that either

$$\sup_{(r, x) \in K} \langle (-1, x_{*0}), (r, x) \rangle < \langle (-1, x_{*0}), (r_0, x_0) \rangle \quad (8.4a)$$

or

$$\inf_{(r, x) \in K} \langle (-1, x_{*0}), (r, x) \rangle > \langle (-1, x_{*0}), (r_0, x_0) \rangle \quad (8.4b)$$

Since K is the epigraph of f , then (8.4b) is impossible since the left-hand-side equals to $-\infty$. Therefore, (8.4a) must hold. Then, $c := \sup_{(r, x) \in K} \langle (-1, x_{*0}), (r, x) \rangle = \sup_{x \in C} (\langle x_{*0}, x \rangle - f(x)) = f_{\text{conj}}(x_{*0}) \in \mathbb{R}$ and $x_{*0} \in C_{\text{conj}}$. Then, $(c, x_{*0}) \in [f_{\text{conj}}, C_{\text{conj}}]$. But, $c < -r_0 + \langle x_{*0}, x_0 \rangle$ implies that $r_0 < \langle x_{*0}, x_0 \rangle - f_{\text{conj}}(x_{*0}) \leq \sup_{x_* \in C_{\text{conj}}} \langle x_*, x_0 \rangle - f_{\text{conj}}(x_*)$. Therefore, $(r_0, x_0) \notin \text{conj}[f_{\text{conj}}, C_{\text{conj}}]$. Hence, $\text{conj}[f_{\text{conj}}, C_{\text{conj}}] \subseteq [f, C]$.

This yields $[f, C] = \text{conj}[[f, C]_{\text{conj}}]$, which implies that f and $\text{conj}(f_{\text{conj}})$ admit the same domain of definition and equal to each other on C . Therefore, $\forall x_0 \in C$, $f(x_0) = \text{conj}(f_{\text{conj}})(x_0) = \sup_{x_* \in C_{\text{conj}}} (\langle x_*, x_0 \rangle - f_{\text{conj}}(x_*))$. $\square$

Proposition 8.33 *Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be a nonempty convex set, $f : C \rightarrow \mathbb{R}$ be a convex functional, f be lower semi-continuous at $x_0 \in C$, $C_{\text{conj}} \subseteq \mathcal{X}^*$ be the conjugate set, and $f_{\text{conj}} : C_{\text{conj}} \rightarrow \mathbb{R}$ be the conjugate functional. Then, $f(x_0) = \sup_{x_* \in C_{\text{conj}}} (\langle x_*, x_0 \rangle - f_{\text{conj}}(x_*))$.*

Proof $\forall x_* \in C_{\text{conj}}$, we have $\langle\langle x_*, x_0 \rangle\rangle - f_{\text{conj}}(x_*) = \langle\langle x_*, x_0 \rangle\rangle + \inf_{x \in C} (-\langle\langle x_*, x \rangle\rangle + f(x)) \leq f(x_0)$. Then,

$$f(x_0) \geq \sup_{x_* \in C_{\text{conj}}} (\langle\langle x_*, x_0 \rangle\rangle - f_{\text{conj}}(x_*))$$

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the lower semicontinuity of f at x_0 , $\exists \delta \in (0, \epsilon/2] \subset \mathbb{R}$ such that $f(x) > f(x_0) - \epsilon/2$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta) \cap C$. Consider the real normed linear space $\mathcal{W} := \mathbb{R} \times \mathcal{X}$. Define $K_2 := [f, C]$, which is clearly nonempty. By Proposition 8.20, K_2 is convex. Define $K_1 := \mathcal{B}_{\mathcal{W}}((f(x_0) - \epsilon, x_0), \delta)$. Then, K_1 is convex and $K_1^\circ = K_1 \neq \emptyset$.

Claim 8.33.1 $K_1^\circ \cap K_2 = \emptyset$.

Proof of claim: Suppose $\exists (r, x) \in K_1^\circ \cap K_2$. Then, $x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta) \cap C$, $r < f(x_0) - \epsilon + \delta \leq f(x_0) - \epsilon/2$ and $r \geq f(x)$. This leads to $f(x) < f(x_0) - \epsilon/2$, which is a contradiction. Hence, $K_1^\circ \cap K_2 = \emptyset$. This completes the proof of the claim. $\square$

By Eidelheit Separation Theorem and Example 7.76, $\exists c \in \mathbb{R}$ and $\exists (\bar{s}_0, \bar{x}_{*0}) \in \mathcal{W}^* = \mathbb{R} \times \mathcal{X}^*$ with $(\bar{s}_0, \bar{x}_{*0}) \neq (0, \vartheta_{\mathcal{X}^*})$ such that

$$\sup_{(r, x) \in K_1} \langle\langle (\bar{s}_0, \bar{x}_{*0}), (r, x) \rangle\rangle \leq c \leq \inf_{(r, x) \in K_2} \langle\langle (\bar{s}_0, \bar{x}_{*0}), (r, x) \rangle\rangle \quad (8.5)$$

The second inequality in (8.5) implies that $-\infty < c \leq \inf_{(r, x) \in K_2} (\bar{s}_0 r + \langle\langle \bar{x}_{*0}, x \rangle\rangle)$. Since $K_2 = [f, C]$, then $\bar{s}_0 \geq 0$.

Claim 8.33.2 $\bar{s}_0 > 0$.

Proof of claim: Suppose $\bar{s}_0 = 0$. By the fact that $(\bar{s}_0, \bar{x}_{*0}) \neq (0, \vartheta_{\mathcal{X}^*})$, we have $\bar{x}_{*0} \neq \vartheta_{\mathcal{X}^*}$. Then, (8.5) implies that $\langle\langle \bar{x}_{*0}, x_0 \rangle\rangle \geq c \geq \sup_{(r, x) \in K_1} \langle\langle \bar{x}_{*0}, x \rangle\rangle = \sup_{x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta)} \langle\langle \bar{x}_{*0}, x \rangle\rangle$. This is impossible since $\bar{x}_{*0} \neq \vartheta_{\mathcal{X}^*}$. Therefore, $\bar{s}_0 < 0$. This completes the proof of the claim. $\square$

Let $x_{*0} := -\bar{s}_0^{-1} \bar{x}_{*0} \in \mathcal{X}^*$. Then, (8.5) is equivalent to

$$\begin{aligned} \inf_{(r, x) \in K_1} (\langle\langle x_{*0}, x_0 \rangle\rangle - r) &\geq -\bar{s}_0^{-1} c \geq \sup_{(r, x) \in K_2} (\langle\langle x_{*0}, x \rangle\rangle - r) \\ &= \sup_{x \in C} (\langle\langle x_{*0}, x \rangle\rangle - f(x)) \end{aligned}$$

Hence, $x_{*0} \in C_{\text{conj}}$ and $f_{\text{conj}}(x_{*0}) \leq \langle\langle x_{*0}, x \rangle\rangle - f(x_0) + \epsilon$. Therefore, we have

$$\langle\langle x_{*0}, x_0 \rangle\rangle - f_{\text{conj}}(x_{*0}) \geq f(x_0) - \epsilon$$

By the arbitrariness of ϵ , $f(x_0) = \sup_{x_* \in C_{\text{conj}}} (\langle\langle x_*, x_0 \rangle\rangle - f_{\text{conj}}(x_*))$.

This completes the proof of the proposition. $\square$

Proposition 8.34 Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be a nonempty convex set, $C^\circ \neq \emptyset$, $f : C \rightarrow \mathbb{R}$ be a convex functional, f be continuous at $\bar{x} \in C^\circ$, $C_{\text{conj}} \subseteq \mathcal{X}^*$ be the conjugate set, and $f_{\text{conj}} : C_{\text{conj}} \rightarrow \mathbb{R}$ be the conjugate functional. Then, $\forall x_0 \in C^\circ$, $f(x_0) = \max_{x_* \in C_{\text{conj}}} (\langle\langle x_*, x_0 \rangle\rangle - f_{\text{conj}}(x_*))$.

Proof $\forall x_* \in C_{\text{conj}}$, we have $\langle\langle x_*, x_0 \rangle\rangle - f_{\text{conj}}(x_*) = \langle\langle x_*, x_0 \rangle\rangle + \inf_{x \in C} (-\langle\langle x_*, x \rangle\rangle + f(x)) \leq f(x_0)$. Then,

$$f(x_0) \geq \sup_{x_* \in C_{\text{conj}}} (\langle\langle x_*, x_0 \rangle\rangle - f_{\text{conj}}(x_*))$$

Consider the real normed linear space $\mathcal{W} := \mathbb{R} \times \mathcal{X}$. Let $V = \{(f(x_0), x_0)\} \subseteq \mathcal{W}$. Clearly, V is a linear variety. Define $K := [f, C]$. By Proposition 8.20, K is convex. By Proposition 8.22, K admits relative interior point $(\bar{r}, \bar{x})$ for some $\bar{r} > f(\bar{x})$. By Proposition 8.21, $V([f, C]) = \mathbb{R} \times V(C) = \mathbb{R} \times \mathcal{X} = \mathcal{W}$ since $C^\circ \neq \emptyset$. Hence, $(\bar{r}, \bar{x}) \in K^\circ \neq \emptyset$. Note that $(f(x_0) - \delta, x_0) \notin [f, C]$, $\forall \delta \in (0, \infty) \subset \mathbb{R}$. Then, $(f(x_0), x_0) \notin K^\circ$. Hence, $V \cap K^\circ = \emptyset$. By Mazur's Theorem and Example 7.76, $\exists c \in \mathbb{R}$ and $\exists(\bar{s}_0, \bar{x}_{*0}) \in \mathcal{W}^* = \mathbb{R} \times \mathcal{X}^*$ with $(\bar{s}_0, \bar{x}_{*0}) \neq (0, \vartheta_{\mathcal{X}^*})$ such that

$$\langle\langle (\bar{s}_0, \bar{x}_{*0}), (f(x_0), x_0) \rangle\rangle = c \geq \sup_{(r, x) \in K} \langle\langle (\bar{s}_0, \bar{x}_{*0}), (r, x) \rangle\rangle$$

which is equivalent to

$$\bar{s}_0 f(x_0) + \langle\langle \bar{x}_{*0}, x_0 \rangle\rangle \geq \sup_{(r, x) \in K} (\bar{s}_0 r + \langle\langle \bar{x}_{*0}, x \rangle\rangle) \quad (8.6)$$

Since $K = [f, C]$, then $\bar{s}_0 \leq 0$, otherwise the right-hand-side of (8.6) equals to $+\infty$.

Claim 8.34.1 $\bar{s}_0 < 0$.

Proof of claim: Suppose $\bar{s}_0 = 0$. By the fact that $(\bar{s}_0, \bar{x}_{*0}) \neq (0, \vartheta_{\mathcal{X}^*})$, we have $\bar{x}_{*0} \neq \vartheta_{\mathcal{X}^*}$. Then, (8.6) implies that $\langle\langle \bar{x}_{*0}, x_0 \rangle\rangle \geq \sup_{x \in C} \langle\langle \bar{x}_{*0}, x \rangle\rangle$. Note that $x_0 \in C^\circ$ and $\bar{x}_{*0} \neq \vartheta_{\mathcal{X}^*}$ implies that $\sup_{x \in C} \langle\langle \bar{x}_{*0}, x \rangle\rangle > \langle\langle \bar{x}_{*0}, x_0 \rangle\rangle$. This is a contradiction. Therefore, $\bar{s}_0 < 0$. This completes the proof of the claim. $\square$

Let $x_{*0} := |\bar{s}_0|^{-1} \bar{x}_{*0} \in \mathcal{X}^*$. Then, (8.6) is equivalent to

$$-f(x_0) + \langle\langle x_{*0}, x_0 \rangle\rangle \geq \sup_{(r, x) \in K} (-r + \langle\langle x_{*0}, x \rangle\rangle) = \sup_{x \in C} (\langle\langle x_{*0}, x \rangle\rangle - f(x))$$

Hence, $x_{*0} \in C_{\text{conj}}$ and $-f(x_0) + \langle\langle x_{*0}, x_0 \rangle\rangle \geq f_{\text{conj}}(x_{*0})$.

Therefore, we have

$$f(x_0) \leq \langle\langle x_{*0}, x_0 \rangle\rangle - f_{\text{conj}}(x_{*0}) \leq \sup_{x_* \in C_{\text{conj}}} (\langle\langle x_*, x_0 \rangle\rangle - f_{\text{conj}}(x_*)) \leq f(x_0)$$

Hence, $f(x_0) = \max_{x_* \in C_{\text{conj}}} (\langle\langle x_*, x_0 \rangle\rangle - f_{\text{conj}}(x_*))$, where the maximum is achieved at x_{*0} . $\square$

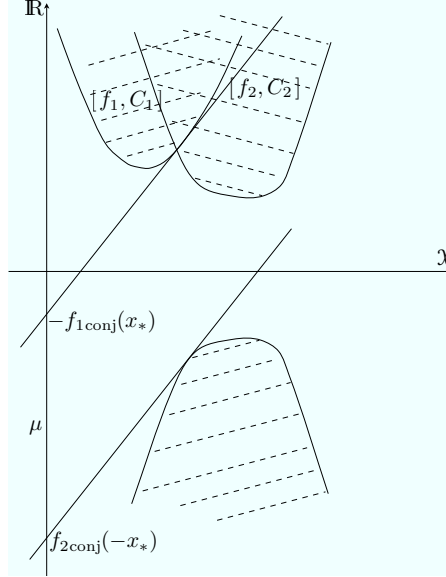


Figure 8.1: Fenchel duality.

8.6 Fenchel Duality Theorem

Let $\mathcal{X}$ be a real normed linear space, $C_1, C_2 \subseteq \mathcal{X}$ be nonempty convex sets, and $f_1 : C_1 \rightarrow \mathbb{R}$ and $f_2 : C_2 \rightarrow \mathbb{R}$ be convex functionals. We consider the problem of

$$\mu := \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x))$$

We assume that $f_{1\text{conj}}$, $C_{1\text{conj}}$, $f_{2\text{conj}}$, $C_{2\text{conj}}$ are easily characterized yet $(f_1 + f_2)_{\text{conj}}$ and $(C_1 \cap C_2)_{\text{conj}}$ are difficult to determine. Then, the above infimum can be equivalently calculated by

$$\mu = \inf_{(r_1, x) \in [f_1, C_1], (r_2, x) \in [f_2, C_2]} r_1 + r_2$$

The idea of Fenchel Duality Theorem can be illustrated in Figure 8.1.

Theorem 8.35 (Fenchel Duality Theorem) *Let $\mathcal{X}$ be a real normed linear space, $C_1, C_2 \subseteq \mathcal{X}$ be nonempty convex sets, $f_1 : C_1 \rightarrow \mathbb{R}$ and $f_2 : C_2 \rightarrow \mathbb{R}$ be convex functionals, and $C_2 \cap C_1^\circ \neq \emptyset$. Assume that f_1 is continuous at $\bar{x} \in C_1^\circ$ and $\mu := \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x))$ is finite. Let $f_{1\text{conj}} : C_{1\text{conj}} \rightarrow \mathbb{R}$ and $f_{2\text{conj}} : C_{2\text{conj}} \rightarrow \mathbb{R}$ be conjugate functionals of f_1 and f_2 , respectively. Then, the following statements hold.*

- (i) $\mu = \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x)) = \max_{x_* \in C_{1\text{conj}} \cap (-C_{2\text{conj}})} (-f_{1\text{conj}}(x_*) - f_{2\text{conj}}(-x_*))$, where the maximum is achieved at some $x_{*0} \in C_{1\text{conj}} \cap$

$(-C_{2\text{conj}})$. If the infimum is achieved by some $x_0 \in C_1 \cap C_2$, then, we have

$$f_{1\text{conj}}(x_{*0}) = \langle x_{*0}, x_0 \rangle - f_1(x_0) \quad (8.7a)$$

$$f_{2\text{conj}}(-x_{*0}) = \langle -x_{*0}, x_0 \rangle - f_2(x_0) \quad (8.7b)$$

(ii) If there exists $x_{*0} \in C_{1\text{conj}} \cap (-C_{2\text{conj}})$ and $x_0 \in C_1 \cap C_2$ such that (8.7) holds, then the infimum is achieved at x_0 and the maximum is achieved at x_{*0} .

Proof $\forall x_* \in C_{1\text{conj}} \cap (-C_{2\text{conj}})$, we have

$$\begin{aligned} -f_{1\text{conj}}(x_*) &= -\sup_{x \in C_1} (\langle x_*, x \rangle - f_1(x)) = \inf_{x \in C_1} (-\langle x_*, x \rangle + f_1(x)) \\ &\leq \inf_{x \in C_1 \cap C_2} (-\langle x_*, x \rangle + f_1(x)) \\ -f_{2\text{conj}}(-x_*) &= -\sup_{x \in C_2} (\langle -x_*, x \rangle - f_2(x)) \\ &= \inf_{x \in C_2} (\langle x_*, x \rangle + f_2(x)) \leq \inf_{x \in C_1 \cap C_2} (\langle x_*, x \rangle + f_2(x)) \end{aligned}$$

Then, we have

$$\begin{aligned} -f_{1\text{conj}}(x_*) - f_{2\text{conj}}(-x_*) &\leq \inf_{x \in C_1 \cap C_2} (-\langle x_*, x \rangle + f_1(x)) \\ &\quad + \inf_{x \in C_1 \cap C_2} (\langle x_*, x \rangle + f_2(x)) \leq \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x)) = \mu \quad (8.8) \end{aligned}$$

(i) Consider the sets $K_1 := [f_1 - \mu, C_1]$ and $K_2 := \{(r, x) \in \mathbb{R} \times \mathcal{X} \mid x \in C_2, r \leq -f_2(x)\} = \{(r, x) \in \mathbb{R} \times \mathcal{X} \mid (-r, x) \in [f_2, C_2]\}$. By Proposition 8.20, K_1 and K_2 are nonempty convex sets. By Proposition 8.22, K_1 admits relative interior point $(\bar{r}, \bar{x})$ for some $\bar{r} \in \mathbb{R}$. By Proposition 8.21, $V([f_1 - \mu, C_1]) = \mathbb{R} \times V(C_1) = \mathbb{R} \times \mathcal{X}$. Then, $(\bar{r}, \bar{x}) \in K_1^\circ \neq \emptyset$.

We will show that $K_1^\circ \cap K_2 = \emptyset$ by an argument of contradiction. Suppose $(r, x) \in K_1^\circ \cap K_2 \neq \emptyset$. Then, we have $(r, x) \in K_1^\circ$, which implies that $x \in C_1$ and $r > f_1(x) - \mu$; and $(r, x) \in K_2$, which implies that $x \in C_2$ and $r \leq -f_2(x)$. Then, $x \in C_1 \cap C_2$ and $f_1(x) + f_2(x) < \mu$. This contradicts with the definition of μ . Hence, $K_1^\circ \cap K_2 = \emptyset$.

By Eidelheit Separation Theorem and Example 7.76, $\exists(\bar{s}_0, \bar{x}_{*0}) \in \mathbb{R} \times \mathcal{X}^*$ with $(\bar{s}_0, \bar{x}_{*0}) \neq \vartheta_{\mathbb{R} \times \mathcal{X}^*}$ such that

$$-\infty < \sup_{(r, x) \in K_1} (\langle \bar{x}_{*0}, x \rangle + r\bar{s}_0) \leq \inf_{(r, x) \in K_2} (\langle \bar{x}_{*0}, x \rangle + r\bar{s}_0) < +\infty \quad (8.9)$$

We will show that $\bar{s}_0 < 0$ by an argument of contradiction. Suppose $\bar{s}_0 \geq 0$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\bar{s}_0 > 0$; Case 2: $\bar{s}_0 = 0$. Case 1: $\bar{s}_0 > 0$. Then, $\sup_{(r, x) \in K_1} (\langle \bar{x}_{*0}, x \rangle + r\bar{s}_0) = +\infty$, which contradicts (8.9). Case 2: $\bar{s}_0 = 0$. Then, $\bar{x}_{*0} \neq \vartheta_{\mathcal{X}^*}$. By (8.9), we have

$$-\infty < \sup_{x \in C_1} \langle \bar{x}_{*0}, x \rangle \leq \inf_{x \in C_2} \langle \bar{x}_{*0}, x \rangle < +\infty$$

Let $\acute{x} \in C_1^\circ \cap C_2 \neq \emptyset$. Then, $\sup_{x \in C_1} \langle \bar{x}_{*0}, x \rangle \leq \langle \bar{x}_{*0}, \acute{x} \rangle$. This is not possible since $\acute{x} \in C_1^\circ$ and $\bar{x}_{*0} \neq \vartheta_{\mathcal{X}^*}$. Hence, we have a contradiction in both cases. Therefore, $\bar{s}_0 < 0$.

Let $x_{*0} = |\bar{s}_0|^{-1} \bar{x}_{*0}$. Then, (8.9) is equivalent to

$$-\infty < \sup_{(r,x) \in K_1} (\langle x_{*0}, x \rangle - r) \leq \inf_{(r,x) \in K_2} (\langle x_{*0}, x \rangle - r) < +\infty \quad (8.10)$$

Note that

$$\begin{aligned} +\infty &> \sup_{(r,x) \in K_1} (\langle x_{*0}, x \rangle - r) = \sup_{x \in C_1} (\langle x_{*0}, x \rangle - f_1(x) + \mu) \\ &\geq \sup_{x \in C_1 \cap C_2} (\langle x_{*0}, x \rangle - f_1(x)) + \mu =: d_1 > -\infty \end{aligned}$$

Hence, $x_{*0} \in C_{1\text{conj}}$. Note also that ,

$$\begin{aligned} -\infty &< \inf_{(r,x) \in K_2} (\langle x_{*0}, x \rangle - r) = \inf_{x \in C_2} (\langle x_{*0}, x \rangle + f_2(x)) \\ &= - \sup_{x \in C_2} (\langle -x_{*0}, x \rangle - f_2(x)) \leq \inf_{x \in C_1 \cap C_2} (\langle x_{*0}, x \rangle + f_2(x)) \\ &=: d_2 < +\infty \end{aligned}$$

Hence, $x_{*0} \in -C_{2\text{conj}}$. By (8.10), $d_1 \leq d_2$. On the other hand, we have

$$\begin{aligned} d_2 - d_1 &= \inf_{x \in C_1 \cap C_2} (\langle x_{*0}, x \rangle + f_2(x)) + \inf_{x \in C_1 \cap C_2} (-\langle x_{*0}, x \rangle + f_1(x)) - \mu \\ &\leq \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x)) - \mu = 0 \end{aligned}$$

Hence, we have

$$d_1 = \sup_{x \in C_1} (\langle x_{*0}, x \rangle - f_1(x) + \mu) = - \sup_{x \in C_2} (\langle -x_{*0}, x \rangle - f_2(x)) = d_2$$

This implies that $d_1 = f_{1\text{conj}}(x_{*0}) + \mu = -f_{2\text{conj}}(-x_{*0}) = d_2$. Therefore, we have $\mu = -f_{1\text{conj}}(x_{*0}) - f_{2\text{conj}}(-x_{*0})$ and $x_{*0} \in C_{1\text{conj}} \cap (-C_{2\text{conj}})$.

This coupled with (8.8) implies

$$\mu = \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x)) = \max_{x_* \in C_{1\text{conj}} \cap (-C_{2\text{conj}})} (-f_{1\text{conj}}(x_*) - f_{2\text{conj}}(-x_*))$$

where the maximum is achieved at some $x_{*0} \in C_{1\text{conj}} \cap (-C_{2\text{conj}})$.

If the infimum is achieved at $x_0 \in C_1 \cap C_2$, then we have

$$\begin{aligned} -f_{1\text{conj}}(x_{*0}) &\leq \inf_{x \in C_1 \cap C_2} (-\langle x_{*0}, x \rangle + f_1(x)) \leq -\langle x_{*0}, x_0 \rangle \\ &\quad + f_1(x_0) \\ -f_{2\text{conj}}(-x_{*0}) &\leq \inf_{x \in C_1 \cap C_2} (\langle x_{*0}, x \rangle + f_2(x)) \leq \langle x_{*0}, x_0 \rangle + f_2(x_0) \end{aligned}$$

$$\mu = -f_{1\text{conj}}(x_{*0}) - f_{2\text{conj}}(-x_{*0}) \leq f_1(x_0) + f_2(x_0) = \mu$$

Hence, (8.7) holds.

(ii) If there exists $x_{*0} \in C_{1\text{conj}} \cap (-C_{2\text{conj}})$ and $x_0 \in C_1 \cap C_2$ such that (8.7) holds, then, by (8.8), we have

$$\begin{aligned} \mu &\geq -f_{1\text{conj}}(x_{*0}) - f_{2\text{conj}}(-x_{*0}) = -\langle x_{*0}, x_0 \rangle + f_1(x_0) \\ &\quad + \langle x_{*0}, x_0 \rangle + f_2(x_0) = f_1(x_0) + f_2(x_0) \geq \mu \end{aligned}$$

Hence, the infimum is achieved at x_0 and the maximum is achieved at x_{*0} .

This completes the proof of the theorem. $\square$

Proposition 8.36 *Let Ω be a set, $J : \Omega \rightarrow \mathbb{R}_e$, Λ be an index set. $A : \Lambda \rightarrow \Omega$. Assume $\bigcup_{\lambda \in \Lambda} A(\lambda) = \Omega$. Then,*

$$\inf_{\omega \in \Omega} J(\omega) = \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega); \quad \sup_{\omega \in \Omega} J(\omega) = \sup_{\lambda \in \Lambda} \sup_{\omega \in A(\lambda)} J(\omega)$$

Proof $\forall \omega_o \in \Omega$, there exists $\lambda_o \in \Lambda$, such that $\omega_o \in A(\lambda_o)$. Then,

$$J(\omega_o) \geq \inf_{\omega \in A(\lambda_o)} J(\omega) \geq \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega)$$

Hence, by Proposition 3.81, we have

$$\inf_{\omega \in \Omega} J(\omega) \geq \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega)$$

On the other hand, $\forall \lambda \in \Lambda$, $A(\lambda) \subseteq \Omega$ implies that

$$\inf_{\omega \in \Omega} J(\omega) \leq \inf_{\omega \in A(\lambda)} J(\omega)$$

Then,

$$\inf_{\omega \in \Omega} J(\omega) \leq \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega)$$

Combining the above arguments, we have

$$\inf_{\omega \in \Omega} J(\omega) = \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega)$$

Substituting $-J$ into this inequality, by Proposition 3.81, yields the other equality.

This completes the proof of the proposition. $\square$

Proposition 8.37 *Let X and Y be sets, $g_1 : X \rightarrow \mathbb{R}_e$, $g_2 : Y \rightarrow \mathbb{R}_e$. Assume that, $\forall x \in X$, $\forall y \in Y$, $g_1(x) + g_2(y) \in \mathbb{R}_e$ is well-defined. Then,*

$$\inf_{(x,y) \in X \times Y} (g_1(x) + g_2(y)) = \inf_{x \in X} g_1(x) + \inf_{y \in Y} g_2(y)$$

when the right-hand-side makes sense.

Proof Define $\mu_l := \inf_{(x,y) \in X \times Y} (g_1(x) + g_2(y))$ and $\mu_r := \inf_{x \in X} g_1(x) + \inf_{y \in Y} g_2(y)$. We will distinguish two exhaustive and mutually exclusive cases:

Case 1: $X \times Y = \emptyset$; Case 2: $X \times Y \neq \emptyset$. Case 1: $X \times Y = \emptyset$. Without loss of generality, assume $X = \emptyset$. Then, $\mu_l = +\infty = \inf_{x \in X} g_1(x)$. Since $\mu_r \in \mathbb{R}_e$ is well-defined, then $\inf_{y \in Y} g_2(y) > -\infty$. Hence, the desired equality holds.

Case 2: $X \times Y \neq \emptyset$. Then, $X \neq \emptyset$ and $Y \neq \emptyset$. This implies that $\inf_{x \in X} g_1(x) = \inf_{(x,y) \in X \times Y} g_1(x)$ and $\inf_{y \in Y} g_2(y) = \inf_{(x,y) \in X \times Y} g_2(y)$. By Proposition 3.81, we have

$$\mu_l \geq \inf_{(x,y) \in X \times Y} g_1(x) + \inf_{(x,y) \in X \times Y} g_2(y) = \mu_r$$

We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: $\mu_r = +\infty$; Case 2b: $\mu_r < +\infty$. Case 2a: $\mu_r = +\infty$. Then, $+\infty \geq \mu_l \geq \mu_r = +\infty$. Hence, the desired equality holds.

Case 2b: $\mu_r < +\infty$. $\forall m \in \mathbb{R}$ with $m > \mu_r$, by Proposition 3.81, $\exists x_0 \in X$ and $\exists y_0 \in Y$ such that $g_1(x_0) + g_2(y_0) < m$. Then, again by Proposition 3.81, we have $\mu_l < m$. This implies that $\mu_l \leq \mu_r$. Hence, the desired equality holds.

This completes the proof of the proposition. $\square$

Proposition 8.38 Let $\mathcal{X}$ be a topological space, Y be a set, and $V : \mathcal{X} \times Y \rightarrow \mathbb{R}$. Assume that V satisfies the following two conditions.

- (i) $\forall x_1 \in \mathcal{X}$, $W(x_1) := \inf_{y \in Y} V(x_1, y) \in \mathbb{R}$. This defines the function $W : \mathcal{X} \rightarrow \mathbb{R}$.
- (ii) $\forall y \in Y$, define the function $f_y : \mathcal{X} \rightarrow \mathbb{R}$ by $f_y(x) = V(x, y)$, $\forall x \in \mathcal{X}$. The collection of functions $\{f_y \mid y \in Y\}$ is equicontinuous.

Then, W is continuous.

Proof $\forall x_0 \in \mathcal{X}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by (ii), $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x_0 \in U$ such that, $\forall \bar{x} \in U$, $\forall y \in Y$, we have $|f_y(\bar{x}) - f_y(x_0)| = |V(\bar{x}, y) - V(x_0, y)| < \epsilon$. By (i), $\exists \bar{y}, y_0 \in Y$ such that $|W(\bar{x}) - V(\bar{x}, \bar{y})| < \epsilon$ and $|W(x_0) - V(x_0, y_0)| < \epsilon$. Note that

$$\begin{aligned} -2\epsilon &< -|W(\bar{x}) - V(\bar{x}, \bar{y})| - |V(\bar{x}, \bar{y}) - V(x_0, \bar{y})| \leq W(\bar{x}) - V(x_0, \bar{y}) \\ &\leq W(\bar{x}) - V(x_0, \bar{y}) + V(x_0, \bar{y}) - W(x_0) = W(\bar{x}) - W(x_0) \\ &= W(\bar{x}) - V(\bar{x}, y_0) + V(\bar{x}, y_0) - W(x_0) \leq V(\bar{x}, y_0) - W(x_0) \\ &\leq |V(\bar{x}, y_0) - V(x_0, y_0)| + |V(x_0, y_0) - W(x_0)| < 2\epsilon \end{aligned}$$

Hence, we have $|W(\bar{x}) - W(x_0)| < 2\epsilon$ and W is continuous at x_0 . By Proposition 3.9, W is continuous.

This completes the proof of the proposition. $\square$

Next, we present a result on game theory.

Proposition 8.39 *Let $\mathcal{X}$ be a reflexive real normed linear space and $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{X}^*$ be nonempty bounded closed convex sets. Then,*

$$\min_{x \in A} \max_{x_* \in B} \langle x_*, x \rangle = \max_{x_* \in B} \min_{x \in A} \langle x_*, x \rangle$$

Proof Let $\mu := \inf_{x \in A} \sup_{x_* \in B} \langle x_*, x \rangle$. Let $h : \mathcal{X} \times B \rightarrow \mathbb{R}$ be given by $h(x, x_*) = \langle x_*, x \rangle$, $\forall x \in \mathcal{X}, \forall x_* \in B$. By Proposition 7.72, h is continuous. Define $f_1 : \mathcal{X} \rightarrow \mathbb{R}_e$ by $f_1(x) = \sup_{x_* \in B} h(x, x_*)$, $\forall x \in \mathcal{X}$. $\forall x \in \mathcal{X}$, define $h_x : \mathcal{X}^* \rightarrow \mathbb{R}$ by $h_x(x_*) = h(x, x_*) = \langle x_*, x \rangle$, $\forall x_* \in \mathcal{X}^*$. Then, h_x is weak* continuous. By Proposition 8.11, B is weak* compact. Then, by Proposition 5.29, $\exists \bar{x}_* \in B$ such that $f_1(x) = h_x(\bar{x}_*) = \max_{x_* \in B} h_x(x_*) = \max_{x_* \in B} h(x, x_*) \in \mathbb{R}$. Hence, $f_1 : \mathcal{X} \rightarrow \mathbb{R}$ takes value in $\mathbb{R}$.

Since B is bounded, then $\exists M_B \in [0, \infty) \subset \mathbb{R}$ such that $\|x_*\| \leq M_B$, $\forall x_* \in B$. $\forall x_* \in B$, define $h_{x_*} : \mathcal{X} \rightarrow \mathbb{R}$ by $h_{x_*}(x) = h(x, x_*)$, $\forall x \in \mathcal{X}$. $\forall x_0 \in \mathcal{X}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta = \epsilon/(1 + M_B) \in (0, \infty) \subset \mathbb{R}$. $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\forall x_* \in B$, we have $|h_{x_*}(x) - h_{x_*}(x_0)| = |h(x, x_*) - h(x_0, x_*)| = |\langle x_*, x - x_0 \rangle| \leq \|x_*\| \|x - x_0\| \leq M_B \delta < \epsilon$, where we have applied Proposition 7.72. Hence, $\{h_{x_*} \mid x_* \in B\}$ is equicontinuous. By Proposition 8.38, $f_1 : \mathcal{X} \rightarrow \mathbb{R}$ is continuous.

$\forall x_1, x_2 \in \mathcal{X}$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we have

$$\begin{aligned} f_1(\alpha x_1 + (1 - \alpha)x_2) &= \max_{x_* \in B} \langle x_*, \alpha x_1 + (1 - \alpha)x_2 \rangle \\ &\leq \sup_{x_* \in B} \alpha \langle x_*, x_1 \rangle + \sup_{x_* \in B} (1 - \alpha) \langle x_*, x_2 \rangle \\ &= \alpha \max_{x_* \in B} \langle x_*, x_1 \rangle + (1 - \alpha) \max_{x_* \in B} \langle x_*, x_2 \rangle \\ &= \alpha f_1(x_1) + (1 - \alpha) f_1(x_2) \end{aligned}$$

where we have applied Proposition 3.81 in the second equality. Hence, f_1 is convex.

Since A is bounded, then, $\exists M_A \in [0, \infty) \subset \mathbb{R}$ such that $\|x\| \leq M_A$, $\forall x \in A$. $\forall x \in A$, $\forall x_* \in B$, by Proposition 7.72, we have $\langle x_*, x \rangle \geq -\|x_*\| \|x\| \geq -M_A M_B$. Then, $f_1(x) \geq -M_A M_B$ since $B \neq \emptyset$. Then, $\mu \geq -M_A M_B$. Since $A \neq \emptyset$, then $\mu < +\infty$. Hence, μ is finite.

Define $f_2 : A \rightarrow \mathbb{R}$ by $f_2(x) = 0$, $\forall x \in A$. Clearly, f_2 is convex. Note $\mu = \inf_{x \in \mathcal{X} \cap A} (f_1(x) + f_2(x))$. Now, it is easy to check that all assumptions for the Fenchel Duality Theorem are satisfied. Then,

$$\mu = \max_{x_* \in \mathcal{X}_{\text{conj}} \cap (-A_{\text{conj}})} (-f_{1\text{conj}}(x_*) - f_{2\text{conj}}(-x_*))$$

Claim 8.39.1 $\mathcal{X}_{\text{conj}} = B$ and $f_{1\text{conj}}(x_*) = 0$, $\forall x_* \in B$.

Proof of claim: $\mathcal{X}_{\text{conj}} = \{x_* \in \mathcal{X}^* \mid \sup_{x \in \mathcal{X}} (\langle x_*, x \rangle - f_1(x)) < +\infty\}$.

$\forall x_{*0} \in B$, $\forall x \in \mathcal{X}$, we have $f_1(x) = \max_{x_* \in B} \langle x_*, x \rangle \geq \langle x_{*0}, x \rangle$. Then, $\langle x_{*0}, x \rangle - f_1(x) \leq 0$. Hence, $x_{*0} \in \mathcal{X}_{\text{conj}}$. Hence, $B \subseteq \mathcal{X}_{\text{conj}}$. The above also implies that $f_{1\text{conj}}(x_{*0}) = \sup_{x \in \mathcal{X}} (\langle x_{*0}, x \rangle - f_1(x)) = 0$.

On the other hand, $\forall x_{*0} \in \mathcal{X}^* \setminus B$, by Proposition 8.10, $\exists x_{**0} \in \mathcal{X}^{**}$ such that $\langle\langle x_{**0}, x_{*0} \rangle\rangle < \inf_{x_* \in B} \langle\langle x_{**0}, x_* \rangle\rangle$. Since $\mathcal{X}$ is reflexive, then $x_{**0} = \phi(x_0)$ for some $x_0 \in \mathcal{X}$, where $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ is the natural mapping. Then, $\langle\langle x_{*0}, x_0 \rangle\rangle < \inf_{x_* \in B} \langle\langle x_*, x_0 \rangle\rangle$ and $\langle\langle x_{*0}, -x_0 \rangle\rangle > \sup_{x_* \in B} \langle\langle x_*, -x_0 \rangle\rangle = f_1(-x_0)$. Then, $\langle\langle x_{*0}, -\alpha x_0 \rangle\rangle - f_1(-\alpha x_0) = \alpha (\langle\langle x_{*0}, -x_0 \rangle\rangle - f_1(-x_0)) > 0$, $\forall \alpha \in (0, \infty) \subset \mathbb{R}$. Hence, $\sup_{x \in \mathcal{X}} (\langle\langle x_{*0}, x \rangle\rangle - f_1(x)) = +\infty$ and $x_{*0} \in \mathcal{X}^* \setminus \mathcal{X}_{\text{conj}}$. This implies that $\mathcal{X}_{\text{conj}} \subseteq B$. Therefore, we have $\mathcal{X}_{\text{conj}} = B$.

This completes the proof of the claim. $\square$

Claim 8.39.2 $A_{\text{conj}} = \mathcal{X}^*$ and $f_{2\text{conj}}(x_*) = \max_{x \in A} \langle\langle x_*, x \rangle\rangle$, $\forall x_* \in \mathcal{X}^*$.

Proof of claim: $A_{\text{conj}} = \{x_* \in \mathcal{X}^* \mid \sup_{x \in A} \langle\langle x_*, x \rangle\rangle < +\infty\}$.

Since $\mathcal{X}$ is reflexive, then, by Definition 7.89, $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ is an isometrical isomorphism. Then, $\phi(A) \subseteq \mathcal{X}^{**}$ is a nonempty bounded closed convex set. By Proposition 7.90, $\mathcal{Y} := \mathcal{X}^*$ is a reflexive real normed linear space. By Proposition 8.11, $\phi(A)$ is weak* compact. $\forall x_* \in \mathcal{X}^*$, $h_{x_*} \circ \phi_{\text{inv}} : \mathcal{X}^{**} \rightarrow \mathbb{R}$ is weak* continuous. By Proposition 5.29, we have

$$\begin{aligned} \sup_{x \in A} \langle\langle x_*, x \rangle\rangle &= \sup_{x_{**} \in \phi(A)} \langle\langle x_{**}, x_* \rangle\rangle = \langle\langle x_{**0}, x_* \rangle\rangle = \langle\langle x_*, \phi_{\text{inv}}(x_{**0}) \rangle\rangle \\ &= \langle\langle x_*, x_0 \rangle\rangle \in \mathbb{R} \end{aligned}$$

for some $x_{**0} \in \phi(A)$ and $x_0 = \phi_{\text{inv}}(x_{**0}) \in A$. Then, $x_* \in A_{\text{conj}}$ and $A_{\text{conj}} = \mathcal{X}^*$. Clearly, $f_{2\text{conj}}(x_*) = \langle\langle x_*, x_0 \rangle\rangle = \max_{x \in A} \langle\langle x_*, x \rangle\rangle$. This completes the proof of the claim. $\square$

Then,

$$\mu = \inf_{x \in A} \max_{x_* \in B} \langle\langle x_*, x \rangle\rangle = \max_{x_* \in B \cap \mathcal{X}^*} (-\max_{x \in A} \langle\langle -x_*, x \rangle\rangle) = \max_{x_* \in B} \min_{x \in A} \langle\langle x_*, x \rangle\rangle$$

To show that the infimum in the above equation is actually achieved, we note that

$$-\mu = \min_{x_* \in B} \max_{x \in (-A)} \langle\langle x_*, x \rangle\rangle = \min_{x_* \in B} \max_{x_{**} \in \phi(-A)} \langle\langle x_{**}, x_* \rangle\rangle$$

By Proposition 7.90, $\mathcal{Y} := \mathcal{X}^*$ is a reflexive real normed linear space and $\phi(\mathcal{X}) = \mathcal{X}^{**}$. Since ϕ is an isometrical isomorphism, then $\phi(-A)$ is a nonempty bounded closed convex set. Applying the result that we have obtained in this proof to the above, we have

$$-\mu = \max_{x_{**} \in \phi(-A)} \min_{x_* \in B} \langle\langle x_{**}, x_* \rangle\rangle = \max_{x \in (-A)} \min_{x_* \in B} \langle\langle x_*, x \rangle\rangle$$

which is equivalent to $\mu = \min_{x \in A} \max_{x_* \in B} \langle\langle x_*, x \rangle\rangle$.

This completes the proof of the proposition. $\square$

8.7 Positive Cones and Convex Mappings

Definition 8.40 Let $\mathcal{X}$ be a normed linear space and $P \subseteq \mathcal{X}$ be a closed convex cone. For $x, y \in \mathcal{X}$, we will write $x \geq y$ (with respect to P) if $x - y \in P$. The cone P defining this relation is called the positive cone in $\mathcal{X}$. The cone $N = -P$ is called the negative cone in $\mathcal{X}$ and we write $y \leq x$ if $y - x \in N$. We will write $x \succ y$ ($y \prec x$) if $x - y \in P^\circ$ ($y - x \in N^\circ = -P^\circ$).

It is easy to check that relations $\leq$ and $\geq$ are reflexive and transitive.

Proposition 8.41 Let $\mathcal{X}$ be a normed linear space with positive cone P . $\forall x_1, x_2, x_3, x_4 \in \mathcal{X}$, $\forall \alpha \in [0, \infty) \subset \mathbb{R}$, we have

- (i) $x_1, x_2 \in P$ implies $x_1 + x_2 \in P$, $\alpha x_1 \in P$, and $-\alpha x_1 \leq \vartheta$;
- (ii) $x_1 \leq x_1$;
- (iii) $x_1 \leq x_2$ and $x_2 \leq x_3$ implies $x_1 \leq x_3$;
- (iv) $x_1 \leq x_2$ and $x_3 \leq x_4$ implies $x_1 + x_3 \leq x_2 + x_4$;
- (v) $x_1 \leq x_2$ implies $\alpha x_1 \leq \alpha x_2$;
- (vi) $x_1 \succ \vartheta$ and $x_2 \geq \vartheta$ implies $x_1 + x_2 \succ \vartheta$;
- (vii) $x_1 \prec x_2$ and $\alpha > 0$ implies $\alpha x_1 \prec \alpha x_2$.

Proof This is straightforward. □

Definition 8.42 Let $\mathcal{X}$ be a real normed linear space and $S \subseteq \mathcal{X}$. The set $S^\oplus := \{x_* \in \mathcal{X}^* \mid \langle x_*, x \rangle \geq 0, \forall x \in S\}$ is called the positive conjugate cone of S . The set $S^\ominus := \{x_* \in \mathcal{X}^* \mid \langle x_*, x \rangle \leq 0, \forall x \in S\}$ is called the negative conjugate cone of S . Clearly, $S^\ominus = -S^\oplus$.

Proposition 8.43 Let $\mathcal{X}$ be a real normed linear space and $S, T \subseteq \mathcal{X}$. Then,

- (i) $S^\oplus \subseteq \mathcal{X}^*$ is a closed convex cone;
- (ii) if $S \subseteq T$, then $T^\oplus \subseteq S^\oplus$.

Proof (i) Clearly, $\vartheta_* \in S^\oplus$. $\forall x_* \in S^\oplus$, $\forall \alpha \in [0, \infty) \subset \mathbb{R}$, $\forall x \in S$, we have $\langle \alpha x_*, x \rangle = \alpha \langle x_*, x \rangle \geq 0$. Hence, $\alpha x_* \in S^\oplus$. Therefore, $S^\oplus$ is a cone with vertex at origin. $\forall x_{*1}, x_{*2} \in S^\oplus$, $\forall x \in S$, we have $\langle x_{*1} + x_{*2}, x \rangle = \langle x_{*1}, x \rangle + \langle x_{*2}, x \rangle \geq 0$. Hence, $x_{*1} + x_{*2} \in S^\oplus$. Therefore, $S^\oplus$ is a convex cone.

$\forall x_* \in \overline{S^\oplus}$, by Proposition 4.13, $\exists (x_{*n})_{n=1}^\infty \subseteq S^\oplus$ such that $\lim_{n \in \mathbb{N}} x_{*n} = x_*$. $\forall x \in S$, by Propositions 7.72 and 3.66, $\langle x_*, x \rangle = \lim_{n \in \mathbb{N}} \langle x_{*n}, x \rangle \geq 0$. Therefore, $x_* \in S^\oplus$ and $\overline{S^\oplus} \subseteq S^\oplus$. By Proposition 3.3, $S^\oplus$ is closed.

(ii) This is straightforward.

This completes the proof of the proposition. □

Proposition 8.44 *Let $\mathcal{X}$ and $\mathcal{Y}$ be real normed linear spaces, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, and $S \subseteq \mathcal{X}$. Then, $(A(S))^\oplus = A'_{\text{inv}}(S^\oplus)$.*

Proof $\forall y_* \in (A(S))^\oplus \subseteq \mathcal{Y}^*$, $\forall s \in S$, we have $\langle\langle A'y_*, s \rangle\rangle = \langle\langle y_*, As \rangle\rangle \geq 0$. Then, $A'y_* \in S^\oplus$ and $y_* \in A'_{\text{inv}}(S^\oplus)$. Hence, $(A(S))^\oplus \subseteq A'_{\text{inv}}(S^\oplus)$.

On the other hand, $\forall y_* \in A'_{\text{inv}}(S^\oplus)$, $A'y_* \in S^\oplus$. $\forall s \in S$, we have $\langle\langle y_*, As \rangle\rangle = \langle\langle A'y_*, s \rangle\rangle \geq 0$. Therefore, we have $y_* \in (A(S))^\oplus$ and $A'_{\text{inv}}(S^\oplus) \subseteq (A(S))^\oplus$.

Hence, $(A(S))^\oplus = A'_{\text{inv}}(S^\oplus)$. This completes the proof of the proposition. $\square$

Definition 8.45 *Let $\mathcal{X}$ be a real normed linear space and $P \subseteq \mathcal{X}$ be the positive cone. We will define $P^\oplus \subseteq \mathcal{X}^*$ to be the positive cone in the dual.*

Proposition 8.46 *Let $\mathcal{X}$ be a real normed linear space with the positive cone P . If $x_0 \in \mathcal{X}$ satisfies that $\langle\langle x_*, x_0 \rangle\rangle \geq 0$, $\forall x_* \in P^\oplus$ (or $x_* \geq \vartheta_*$), then $x_0 \in P$.*

Proof Suppose $x_0 \notin P$, by Proposition 8.10, $\exists x_* \in \mathcal{X}^*$, such that $-\infty < \langle\langle x_*, x_0 \rangle\rangle < \inf_{x \in P} \langle\langle x_*, x \rangle\rangle$. Since P is a cone, then $\inf_{x \in P} \langle\langle x_*, x \rangle\rangle = 0$ (it must be greater than or equal to 0 since otherwise the infimum must be $-\infty$; and it must be less than or equal to 0 since $\vartheta_{\mathcal{X}} \in P$). Hence, $x_* \in P^\oplus$ and $\langle\langle x_*, x_0 \rangle\rangle < 0$. This contradicts with the assumption. Therefore, we must have $x_0 \in P$. This completes the proof of the proposition. $\square$

Proposition 8.47 *Let $\mathcal{X}$ be a real normed linear space with positive cone P . $\forall x_* \in \mathcal{X}^*$, $\forall x \in \mathcal{X}$, we have*

- (i) $x \geq \vartheta$ and $x_* \geq \vartheta_*$ implies $\langle\langle x_*, x \rangle\rangle \geq 0$;
- (ii) $x \leq \vartheta$ and $x_* \geq \vartheta_*$ implies $\langle\langle x_*, x \rangle\rangle \leq 0$;
- (iii) $x > \vartheta$, $x_* \geq \vartheta_*$, and $x_* \neq \vartheta_*$ implies $\langle\langle x_*, x \rangle\rangle > 0$;
- (iv) $x \geq \vartheta$, $x \neq \vartheta$, and $x_* > \vartheta_*$ implies $\langle\langle x_*, x \rangle\rangle > 0$.

Proof This is straightforward. $\square$

Definition 8.48 *Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$, and $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$. A mapping $G : \Omega \rightarrow \mathcal{Z}$ is said to be convex if Ω is convex and $\forall x_1, x_2 \in \Omega$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we have $G(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha G(x_1) + (1 - \alpha)G(x_2)$.*

We note that the convexity of a mapping depends on the definition of the positive cone P .

Proposition 8.49 *Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be convex, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, and $G : \Omega \rightarrow \mathcal{Z}$ be a convex mapping. Then, $\forall z \in \mathcal{Z}$, the set $\Omega_z := \left\{ x \in \Omega \mid G(x) \leq z \right\}$ is convex.*

Proof Fix any $z \in \mathcal{Z}$. $\forall x_1, x_2 \in \Omega_z$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we have $G(x_1) \leq z$ and $G(x_2) \leq z$. Since P is a convex cone, then, $\alpha G(x_1) + (1 - \alpha)G(x_2) \leq \alpha z + (1 - \alpha)z = z$. By the convexity of G , we have $G(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha G(x_1) + (1 - \alpha)G(x_2)$. By Proposition 8.41, we have $G(\alpha x_1 + (1 - \alpha)x_2) \leq z$. Hence, $\alpha x_1 + (1 - \alpha)x_2 \in \Omega_z$. Hence, Ω_z is convex. This completes the proof of the proposition. $\square$

8.8 Lagrange Multipliers

The basic problem to be considered in this section is

$$\mu_0 := \inf_{x \in \Omega, G(x) \leq \vartheta_z} f(x) \quad (8.11)$$

where $\mathcal{X}$ is a real vector space, $\Omega \subseteq \mathcal{X}$ is a nonempty convex set, $f : \Omega \rightarrow \mathbb{R}$ is a convex functional, $\mathcal{Z}$ is a real normed linear space with positive cone $P \subseteq \mathcal{Z}$, and $G : \Omega \rightarrow \mathcal{Z}$ is a convex mapping.

Toward a solution to the above problem, we consider a class of problems:

$$\Gamma := \left\{ z \in \mathcal{Z} \mid \exists x \in \Omega \ni G(x) \leq z \right\} \quad (8.12a)$$

$$\omega(z) := \inf_{x \in \Omega, G(x) \leq z} f(x); \quad \forall z \in \Gamma \quad (8.12b)$$

where $\omega : \Gamma \rightarrow \mathbb{R}_e$ is the *primal functional*. To guarantee that ω is real-valued, we make the following assumption.

Assumption 8.50 $\vartheta_z \in \Gamma$ and (i) $\exists \bar{z} \in {}^\circ\Gamma$ such that $\omega(\bar{z}) \in \mathbb{R}$ or (ii) $\mu := \inf_{x \in \Omega} f(x) > -\infty$ holds.

Fact 8.51 $\Gamma \subseteq \mathcal{Z}$ is convex.

Proof $\forall z_1, z_2 \in \Gamma$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, there exist $x_i \in \Omega$ such that $G(x_i) \leq z_i$, $i = 1, 2$. By the convexity of Ω , we have $\alpha x_1 + (1 - \alpha)x_2 \in \Omega$. Then, by the convexity of G and Proposition 8.41, $G(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha G(x_1) + (1 - \alpha)G(x_2) \leq \alpha z_1 + (1 - \alpha)z_2$. Hence, $\alpha z_1 + (1 - \alpha)z_2 \in \Gamma$. This completes the proof of the fact. $\square$

Fact 8.52 Under Assumption 8.50, $\omega : \Gamma \rightarrow \mathbb{R}$ is real-valued, convex, and nonincreasing, that is, $\forall z_1, z_2 \in \Gamma$ with $z_1 \leq z_2$, we have $\omega(z_1) \geq \omega(z_2)$.

Proof We will first show that $\omega(z) \in \mathbb{R}$, $\forall z \in \Gamma$. Let (i) in Assumption 8.50 hold. Fix any $z \in \Gamma$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $z = \bar{z}$; Case 2: $z \neq \bar{z}$. Case 1: $z = \bar{z}$. Then, $\omega(z) = \omega(\bar{z}) \in \mathbb{R}$.

Case 2: $z \neq \bar{z}$. By the definition of Γ and the fact that f is real-valued, we have $\omega(z) < +\infty$. By (i) of Assumption 8.50, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_z(\bar{z}, \delta) \cap V(\Gamma) \subseteq \Gamma$. Let $\bar{z}_1 := \bar{z} + \frac{\delta}{2\|\bar{z}-z\|}(\bar{z}-z) \in \mathcal{B}_z(\bar{z}, \delta) \cap V(\Gamma) \subseteq \Gamma$ and $\bar{\alpha} = \frac{\|\bar{z}-z\|}{\|\bar{z}-z\|+\delta/2} \in (0, 1) \subset \mathbb{R}$. It is easy to verify that $\bar{z} = \bar{\alpha}\bar{z}_1 + (1-\bar{\alpha})z$. Then, we have

$$\begin{aligned}
-\infty &< \omega(\bar{z}) = \inf_{\bar{x} \in \Omega, G(\bar{x}) \leq \bar{\alpha}\bar{z}_1 + (1-\bar{\alpha})z} f(\bar{x}) \\
&\leq \inf_{\bar{x} = \bar{\alpha}\bar{x}_1 + (1-\bar{\alpha})z, \bar{x}_1 \in \Omega, G(\bar{x}_1) \leq \bar{z}_1, x \in \Omega, G(x) \leq z} f(\bar{x}) \\
&\leq \inf_{\bar{x}_1 \in \Omega, G(\bar{x}_1) \leq \bar{z}_1, x \in \Omega, G(x) \leq z} (\bar{\alpha}f(\bar{x}_1) + (1-\bar{\alpha})f(x)) \\
&= \inf_{\bar{x}_1 \in \Omega, G(\bar{x}_1) \leq \bar{z}_1} \bar{\alpha}f(\bar{x}_1) + \inf_{x \in \Omega, G(x) \leq z} (1-\bar{\alpha})f(x) \\
&= \bar{\alpha} \inf_{\bar{x}_1 \in \Omega, G(\bar{x}_1) \leq \bar{z}_1} f(\bar{x}_1) + (1-\bar{\alpha}) \inf_{x \in \Omega, G(x) \leq z} f(x) \\
&= \bar{\alpha}\omega(\bar{z}_1) + (1-\bar{\alpha})\omega(z)
\end{aligned}$$

where the second equality follows from Proposition 8.37, and the third equality follows from Proposition 3.81. Then, $\omega(z) > -\infty$. Hence, $\omega(z) \in \mathbb{R}$.

In both cases, we have $\omega(z) \in \mathbb{R}$. Therefore, $\omega : \Gamma \rightarrow \mathbb{R}$ is real-valued.

Let (ii) in Assumption 8.50 hold. Fix any $z \in \Gamma$. Then, $\omega(z) \geq \mu > -\infty$. By the definition of Γ , we have $\omega(z) < +\infty$. Hence, $\omega(z) \in \mathbb{R}$.

Thus, under Assumption 8.50, $\omega : \Gamma \rightarrow \mathbb{R}$ is real-valued.

$\forall z_1, z_2 \in \Gamma$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we have

$$\begin{aligned}
\omega(\alpha z_1 + (1-\alpha)z_2) &= \inf_{x \in \Omega, G(x) \leq \alpha z_1 + (1-\alpha)z_2} f(x) \\
&\leq \inf_{x = \alpha x_1 + (1-\alpha)x_2, x_1 \in \Omega, G(x_1) \leq z_1, x_2 \in \Omega, G(x_2) \leq z_2} f(x) \\
&\leq \inf_{x_1 \in \Omega, G(x_1) \leq z_1, x_2 \in \Omega, G(x_2) \leq z_2} (\alpha f(x_1) + (1-\alpha)f(x_2)) \\
&= \inf_{x_1 \in \Omega, G(x_1) \leq z_1} \alpha f(x_1) + \inf_{x_2 \in \Omega, G(x_2) \leq z_2} (1-\alpha)f(x_2) \\
&= \alpha \inf_{x_1 \in \Omega, G(x_1) \leq z_1} f(x_1) + (1-\alpha) \inf_{x_2 \in \Omega, G(x_2) \leq z_2} f(x_2) \\
&= \alpha\omega(z_1) + (1-\alpha)\omega(z_2)
\end{aligned}$$

where the second equality follows from Proposition 8.37, and the third equality follows from Proposition 3.81. Hence, ω is convex.

It is obvious that ω is nonincreasing. This completes the proof of the fact. $\square$

Fact 8.53 *Under Assumption 8.50, we have $\Gamma_{\text{conj}} \subseteq P^\ominus$ and $-\Gamma_{\text{conj}} \subseteq P^\oplus$.*

Proof Since $\vartheta_Z \in \Gamma$ implies that $\exists x_1 \in \Omega$ such that $G(x_1) \leq \vartheta_Z$, then $\Gamma \supseteq G(x_1) + P$. $\forall z_* \in \Gamma_{\text{conj}}$, by Definition 8.27, we have

$$\begin{aligned} +\infty &> \sup_{z \in \Gamma} (\langle\langle z_*, z \rangle\rangle - \omega(z)) \geq \sup_{z \in G(x_1) + P} (\langle\langle z_*, z \rangle\rangle - \omega(z)) \\ &\geq \sup_{z \in G(x_1) + P} (\langle\langle z_*, z \rangle\rangle - \omega(G(x_1))) \\ &= \sup_{\bar{z} \in P} \langle\langle z_*, \bar{z} \rangle\rangle + \langle\langle z_*, G(x_1) \rangle\rangle - \omega(G(x_1)) \end{aligned}$$

where the third inequality follows from Fact 8.52, and the equality follows from Proposition 8.37. Hence, $\sup_{z \geq \vartheta_Z} \langle\langle z_*, z \rangle\rangle < +\infty$. Since P is a cone, then $\sup_{z \geq \vartheta_Z} \langle\langle z_*, z \rangle\rangle = 0$. This implies that $z_* \in P^\ominus$, $\Gamma_{\text{conj}} \subseteq P^\ominus$, and $-\Gamma_{\text{conj}} \subseteq P^\oplus$. This completes the proof of the fact. $\square$

Fact 8.54 *Let Assumption 8.50 hold. Define $\bar{\omega} : P^\ominus \rightarrow \mathbb{R}_e$ by*

$$\bar{\omega}(z_*) = \sup_{z \in \Gamma} (\langle\langle z_*, z \rangle\rangle - \omega(z)); \quad \forall z_* \in P^\ominus$$

Then, $\sup_{z_* \in \Gamma_{\text{conj}}} -\omega_{\text{conj}}(z_*) = \sup_{z_* \in P^\ominus} -\bar{\omega}(z_*)$. $\forall z_* \in P^\oplus$, we have that

$$-\bar{\omega}(-z_*) = \inf_{x \in \Omega} (f(x) + \langle\langle z_*, G(x) \rangle\rangle) =: \varphi(z_*)$$

where $\varphi : P^\oplus \rightarrow \mathbb{R}_e$ is called the dual functional.

Proof Clearly, $\forall z_* \in \Gamma_{\text{conj}}$, $\bar{\omega}(z_*) = \omega_{\text{conj}}(z_*)$ and, by the definition of Γ_{conj} , $\forall z_* \in P^\ominus \setminus \Gamma_{\text{conj}}$, $\bar{\omega}(z_*) = +\infty$. Hence, $\sup_{z_* \in \Gamma_{\text{conj}}} -\omega_{\text{conj}}(z_*) = \sup_{z_* \in P^\ominus} -\bar{\omega}(z_*)$.

$$\forall z_* \in P^\oplus, -z_* \leq \vartheta_{z_*},$$

$$\begin{aligned} -\bar{\omega}(-z_*) &= -\sup_{z \in \Gamma} (\langle\langle -z_*, z \rangle\rangle - \omega(z)) = \inf_{z \in \Gamma} (\langle\langle z_*, z \rangle\rangle + \omega(z)) \\ &= \inf_{z \in \Gamma} (\langle\langle z_*, z \rangle\rangle + \inf_{x \in \Omega, G(x) \leq z} f(x)) = \inf_{z \in \Gamma} \inf_{x \in \Omega, G(x) \leq z} (\langle\langle z_*, z \rangle\rangle + f(x)) \\ &= \inf_{x \in \Omega, z \in \Gamma, G(x) \leq z} (\langle\langle z_*, z \rangle\rangle + f(x)) = \inf_{x \in \Omega} \inf_{z \in \Gamma, z \geq G(x)} (\langle\langle z_*, z \rangle\rangle + f(x)) \\ &= \inf_{x \in \Omega} \inf_{z \geq G(x)} (\langle\langle z_*, z \rangle\rangle + f(x)) = \inf_{x \in \Omega} (f(x) + \inf_{z \geq G(x)} \langle\langle z_*, z \rangle\rangle) \\ &= \inf_{x \in \Omega} (f(x) + \langle\langle z_*, G(x) \rangle\rangle) \end{aligned}$$

where the second equality follows from Proposition 3.81; the fourth equality follows from Proposition 8.37; the fifth and sixth equality follows from

Proposition 8.36; and the eighth equality follows from Proposition 8.37. This completes the proof of the fact. $\square$

The desired theory follows by applying either duality results for convex functionals, which are Propositions 8.33 and 8.34. This leads to two different regularity conditions. To apply Proposition 8.33, we assume that

Assumption 8.55 ω is lower semicontinuous at $\vartheta_{\mathcal{Z}}$.

To apply Proposition 8.34, we assume that

Assumption 8.56 $\exists x_1 \in \Omega$ such that $G(x_1) \leq \vartheta_{\mathcal{Z}}$.

Now, we state the two Lagrange duality results.

Proposition 8.57 Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be nonempty and convex, $f : \Omega \rightarrow \mathbb{R}$ be a convex functional, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, $G : \Omega \rightarrow \mathcal{Z}$ be a convex mapping, and μ_0 be defined as in (8.11). Let Assumptions 8.50 and 8.55 hold. Then,

$$\mu_0 = \sup_{z_* \succeq \vartheta_{\mathcal{Z}}} \inf_{x \in \Omega} (f(x) + \langle \langle z_*, G(x) \rangle \rangle) = \sup_{z_* \succeq \vartheta_{\mathcal{Z}}} \varphi(z_*) \quad (8.13)$$

Furthermore, if the supremum in (8.13) is achieved at $z_{*0} \in P^\oplus$, that is,

$$\mu_0 = \inf_{x \in \Omega} (f(x) + \langle \langle z_{*0}, G(x) \rangle \rangle) \quad (8.14)$$

then the following statement holds: the infimum in (8.11) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \vartheta_{\mathcal{Z}}$ if, and only if, the infimum in (8.14) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \vartheta_{\mathcal{Z}}$ and $\langle \langle z_{*0}, G(x_0) \rangle \rangle = 0$.

Proof By Facts 8.51 and 8.52, Assumption 8.55, and Proposition 8.33, we have

$$\mu_0 = \omega(\vartheta_{\mathcal{Z}}) = \sup_{z_* \in \Gamma_{\text{conj}}} -\omega_{\text{conj}}(z_*)$$

Then, by Facts 8.53 and 8.54, we have

$$\mu_0 = \sup_{z_* \in P^\ominus} -\bar{\omega}(z_*) = \sup_{z_* \succeq \vartheta_{\mathcal{Z}}} -\bar{\omega}(-z_*) = \sup_{z_* \succeq \vartheta_{\mathcal{Z}}} \varphi(z_*)$$

Therefore, (8.13) holds.

Let the supremum in (8.13) be achieved at $z_{*0} \in P^\oplus$, that is, (8.14) holds.

If the infimum in (8.11) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \vartheta_{\mathcal{Z}}$, then, we have $\mu_0 = f(x_0) \geq f(x_0) + \langle \langle z_{*0}, G(x_0) \rangle \rangle \geq \inf_{x \in \Omega} (f(x) + \langle \langle z_{*0}, G(x) \rangle \rangle) = \mu_0$. Hence, the infimum in (8.14) is achieved at x_0 and $\langle \langle z_{*0}, G(x_0) \rangle \rangle = 0$.

On the other hand, if the infimum in (8.14) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \vartheta_{\mathcal{Z}}$ and $\langle \langle z_{*0}, G(x_0) \rangle \rangle = 0$, then $\mu_0 = f(x_0) + \langle \langle z_{*0}, G(x_0) \rangle \rangle = f(x_0)$. Hence, the infimum in (8.11) is achieved at x_0 .

This completes the proof of the proposition. $\square$

Proposition 8.58 *Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be nonempty and convex, $f : \Omega \rightarrow \mathbb{R}$ be a convex functional, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, $P^\circ \neq \emptyset$, $G : \Omega \rightarrow \mathcal{Z}$ be a convex mapping. Let Assumptions 8.50 and 8.56 hold. Then,*

$$\mu_0 = \max_{z_* \succeq \vartheta_{\mathcal{Z}}^*} \inf_{x \in \Omega} (f(x) + \langle \langle z_*, G(x) \rangle \rangle) = \max_{z_* \succeq \vartheta_{\mathcal{Z}}^*} \varphi(z_*) \quad (8.15)$$

where the maximum is achieved at $z_{*0} \in P^\oplus$, that is,

$$\mu_0 = \inf_{x \in \Omega} (f(x) + \langle \langle z_{*0}, G(x) \rangle \rangle) \quad (8.16)$$

Furthermore, the infimum in (8.11) is achieved at $x_0 \in \Omega$ with $G(x_0) \preceq \vartheta_{\mathcal{Z}}$ if, and only if, the infimum in (8.16) is achieved at $x_0 \in \Omega$ with $G(x_0) \preceq \vartheta_{\mathcal{Z}}$ and $\langle \langle z_{*0}, G(x_0) \rangle \rangle = 0$.

Proof By Assumption 8.56, $G(x_1) \prec \vartheta_{\mathcal{Z}}$. Then, $-G(x_1) \in P^\circ$ and $G(x_1) + P \subseteq \Gamma$. By Proposition 7.16, $\Gamma^\circ \supseteq G(x_1) + P^\circ$ and $\vartheta_{\mathcal{Z}} \in \Gamma^\circ$. $\forall z \in G(x_1) + P$, we have $z \succeq G(x_1)$ and, by Fact 8.52, $\omega(z) \leq \omega(G(x_1))$. Since $G(x_1) \prec \vartheta_{\mathcal{Z}}$, then $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{Z}}(\vartheta_{\mathcal{Z}}, \delta) \subseteq G(x_1) + P$. Take $r_0 := \omega(G(x_1)) + \delta \in \mathbb{R}$. It is easy to check that $\mathcal{B}_{\mathbb{R} \times \mathcal{Z}}((r_0, \vartheta_{\mathcal{Z}}), \delta) \subseteq [\omega, \Gamma]$. Hence, $(r_0, \vartheta_{\mathcal{Z}}) \in [\omega, \Gamma]^\circ$. By Proposition 8.22, ω is continuous at $\vartheta_{\mathcal{Z}}$. By Facts 8.51 and 8.52 and Proposition 8.34, we have

$$\mu_0 = \omega(\vartheta_{\mathcal{Z}}) = \max_{z_* \in \Gamma_{\text{conj}}} -\omega_{\text{conj}}(z_*)$$

Then, by Facts 8.53 and 8.54, we have

$$\mu_0 = \max_{z_* \in P^\ominus} -\bar{\omega}(z_*) = \max_{z_* \succeq \vartheta_{\mathcal{Z}}^*} -\bar{\omega}(-z_*) = \max_{z_* \succeq \vartheta_{\mathcal{Z}}^*} \varphi(z_*)$$

Therefore, (8.15) holds and the maximum is achieved at $z_{*0} \in P^\oplus$.

If the infimum in (8.11) is achieved at $x_0 \in \Omega$ with $G(x_0) \preceq \vartheta_{\mathcal{Z}}$, then, we have $\mu_0 = f(x_0) \geq f(x_0) + \langle \langle z_{*0}, G(x_0) \rangle \rangle \geq \inf_{x \in \Omega} (f(x) + \langle \langle z_{*0}, G(x) \rangle \rangle) = \mu_0$. Hence, the infimum in (8.16) is achieved at x_0 and $\langle \langle z_{*0}, G(x_0) \rangle \rangle = 0$.

On the other hand, if the infimum in (8.16) is achieved at $x_0 \in \Omega$ with $G(x_0) \preceq \vartheta_{\mathcal{Z}}$ and $\langle \langle z_{*0}, G(x_0) \rangle \rangle = 0$, then $\mu_0 = f(x_0) + \langle \langle z_{*0}, G(x_0) \rangle \rangle = f(x_0)$. Hence, the infimum in (8.11) is achieved at x_0 .

This completes the proof of the proposition. $\square$

In the Propositions 8.57 and 8.58, z_{*0} is called the *Lagrange multiplier*. Assumption 8.50 guarantees that the primal functional is real-valued and convex. Assumption 8.56 guarantees the existence of a Lagrange multiplier. This assumption is restrictive. On the other hand, Assumption 8.55 guarantees the duality but not the existence of a Lagrange multiplier. This condition is more relaxed.

Corollary 8.59 *Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be nonempty and convex, $f : \Omega \rightarrow \mathbb{R}$ be a convex functional, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, $P^\circ \neq \emptyset$, $G : \Omega \rightarrow \mathcal{Z}$ be a convex mapping. Let Assumptions 8.50 and 8.56 hold and $x_0 \in \Omega$ with $G(x_0) \leq \vartheta_{\mathcal{Z}}$ achieves the infimum in (8.11). Then, there exists $z_{*0} \in \mathcal{Z}^*$ with $z_{*0} \geq \vartheta_{\mathcal{Z}^*}$ such that the Lagrangian $L : \Omega \times P^\oplus \rightarrow \mathbb{R}$ defined by*

$$L(x, z_*) := f(x) + \langle \langle z_*, G(x) \rangle \rangle; \quad \forall x \in \Omega, \forall z_* \in P^\oplus$$

admits a saddle point at (x_0, z_{*0}) , i. e.

$$L(x_0, z_*) \leq L(x_0, z_{*0}) \leq L(x, z_{*0}); \quad \forall x \in \Omega, \forall z_* \in P^\oplus$$

Proof By Proposition 8.58, there exists $z_{*0} \in \mathcal{Z}^*$ with $z_{*0} \geq \vartheta_{\mathcal{Z}^*}$ such that $L(x_0, z_{*0}) \leq L(x, z_{*0})$, $\forall x \in \Omega$ and $\langle \langle z_{*0}, G(x_0) \rangle \rangle = 0$. Then, $\forall z_* \in P^\oplus$, we have $L(x_0, z_*) - L(x_0, z_{*0}) = \langle \langle z_*, G(x_0) \rangle \rangle - \langle \langle z_{*0}, G(x_0) \rangle \rangle = \langle \langle z_*, G(x_0) \rangle \rangle \leq 0$. Hence, the saddle point condition holds. This completes the proof of the corollary. $\square$

Next, we present two sufficiency results on Lagrange multipliers.

Proposition 8.60 *Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be nonempty, $f : \Omega \rightarrow \mathbb{R}$, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, $G : \Omega \rightarrow \mathcal{Z}$.*

Assume that there exist $x_0 \in \Omega$ and $z_{*0} \in \mathcal{Z}^*$ with $z_{*0} \geq \vartheta_{\mathcal{Z}^*}$ such that

$$f(x_0) + \langle \langle z_{*0}, G(x_0) \rangle \rangle \leq f(x) + \langle \langle z_{*0}, G(x) \rangle \rangle; \quad \forall x \in \Omega$$

Then,

$$f(x_0) = \inf_{x \in \Omega, G(x) \leq G(x_0)} f(x)$$

Proof $\forall x \in \Omega$ with $G(x) \leq G(x_0)$, we have $\langle \langle z_{*0}, G(x) \rangle \rangle \leq \langle \langle z_{*0}, G(x_0) \rangle \rangle$, since $z_{*0} \in P^\oplus$. By the assumption of the proposition, $f(x_0) + \langle \langle z_{*0}, G(x_0) \rangle \rangle \leq f(x) + \langle \langle z_{*0}, G(x) \rangle \rangle$. Then, $f(x_0) \leq f(x)$. This completes the proof of the proposition. $\square$

Proposition 8.61 *Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be nonempty, $f : \Omega \rightarrow \mathbb{R}$, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, $G : \Omega \rightarrow \mathcal{Z}$.*

Assume that there exist $x_0 \in \Omega$ and $z_{*0} \in \mathcal{Z}^*$ with $z_{*0} \geq \vartheta_{\mathcal{Z}^*}$ such that the Lagrangian $L : \Omega \times P^\oplus \rightarrow \mathbb{R}$ given by

$$L(x, z_*) = f(x) + \langle \langle z_*, G(x) \rangle \rangle; \quad \forall x \in \Omega, \forall z_* \in P^\oplus$$

admits a saddle point at (x_0, z_{*0}) , i. e.

$$L(x_0, z_*) \leq L(x_0, z_{*0}) \leq L(x, z_{*0}); \quad \forall x \in \Omega, \forall z_* \in P^\oplus$$

Then, $G(x_0) \stackrel{\leq}{=} \vartheta_{\mathcal{Z}}$ and

$$f(x_0) = L(x_0, z_{*0}) = \inf_{x \in \Omega, G(x) \stackrel{\leq}{=} \vartheta_{\mathcal{Z}}} f(x)$$

Proof By the first inequality in the saddle-point condition, we have

$$\langle \langle z_*, G(x_0) \rangle \rangle \leq \langle \langle z_{*0}, G(x_0) \rangle \rangle; \quad \forall z_* \in P^\oplus$$

$\forall z_* \in P^\oplus$, we have $z_* \stackrel{\geq}{=} \vartheta_{\mathcal{Z}^*}$ and $z_* + z_{*0} \geq \vartheta_{\mathcal{Z}^*}$. Then, $\langle \langle z_* + z_{*0}, G(x_0) \rangle \rangle \leq \langle \langle z_{*0}, G(x_0) \rangle \rangle$ and $\langle \langle z_*, G(x_0) \rangle \rangle \leq 0$. By Proposition 8.46, $G(x_0) \in (-P)$ and $G(x_0) \stackrel{\leq}{=} \vartheta_{\mathcal{Z}}$. Furthermore, $0 = \langle \langle \vartheta_{\mathcal{Z}^*}, G(x_0) \rangle \rangle \leq \langle \langle z_{*0}, G(x_0) \rangle \rangle \leq 0$ implies that $\langle \langle z_{*0}, G(x_0) \rangle \rangle = 0$.

$\forall x \in \Omega$ with $G(x) \stackrel{\leq}{=} \vartheta_{\mathcal{Z}}$, we have $f(x) \geq f(x) + \langle \langle z_{*0}, G(x) \rangle \rangle \geq f(x_0) + \langle \langle z_{*0}, G(x_0) \rangle \rangle = f(x_0)$. This completes the proof of the proposition. $\square$

Next, we present a result on the sensitivity of the infimization problem.

Proposition 8.62 *Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be nonempty, $f : \Omega \rightarrow \mathbb{R}$, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, $G : \Omega \rightarrow \mathcal{Z}$.*

Let $z_i \in \mathcal{Z}$, $\mu_i = \inf_{x \in \Omega, G(x) \stackrel{\leq}{=} z_i} f(x)$, $i = 0, 1$. Let $z_{*0} \in P^\oplus \subseteq \mathcal{Z}^*$ be the Lagrange multiplier associated with μ_0 , that is

$$\mu_0 = \inf_{x \in \Omega} (f(x) + \langle \langle z_{*0}, G(x) - z_0 \rangle \rangle)$$

Assume that $\mu_0 \in \mathbb{R}$. Then, we have

$$\mu_1 - \mu_0 \geq \langle \langle -z_{*0}, z_1 - z_0 \rangle \rangle$$

Proof $\forall x \in \Omega$ with $G(x) \stackrel{\leq}{=} z_1$, we have $\mu_0 \leq f(x) + \langle \langle z_{*0}, G(x) - z_0 \rangle \rangle$, which implies that

$$f(x) \geq \mu_0 - \langle \langle z_{*0}, G(x) - z_0 \rangle \rangle \geq \mu_0 - \langle \langle z_{*0}, z_1 - z_0 \rangle \rangle$$

Hence, $\mu_1 \geq \mu_0 + \langle \langle -z_{*0}, z_1 - z_0 \rangle \rangle$. This completes the proof of the proposition. $\square$

Chapter 9

Differentiation in Banach Spaces

In this chapter, we are going to develop the concept of derivative in normed linear spaces.

9.1 Fundamental Notion

Definition 9.1 Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $D \subseteq \mathcal{X}$, and $x_0 \in \overline{D}$. $u \in \mathcal{X}$ is said to be an admissible deviation in D at x_0 if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, we have $\{x_0 + r\bar{u} \mid r \in (0, \epsilon) \subset \mathbb{R}, \bar{u} \in \mathcal{B}(u, \epsilon)\} \cap D \neq \emptyset$. Let $A_D(x_0)$ be the set of admissible deviations in D at x_0 .

Clearly, if $x_0 \in D^\circ$, then $A_D(x_0) = \mathcal{X}$. Another fact is that when $D_1 \subseteq D_2 \subseteq \mathcal{X}$ and $x_0 \in \overline{D_1}$, then $A_{D_1}(x_0) \subseteq A_{D_2}(x_0)$. Yet another fact is that when $x_0 \in \overline{D_1}$, $D_1, D_2 \subseteq \mathcal{X}$, and $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $D_1 \cap \mathcal{B}(x_0, \delta) = D_2 \cap \mathcal{B}(x_0, \delta)$, then $A_{D_1}(x_0) = A_{D_2}(x_0)$.

Proposition 9.2 Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $D \subseteq \mathcal{X}$, and $x_0 \in \overline{D}$. Then, $A_D(x_0) \subseteq \mathcal{X}$ is a closed cone.

Proof Clearly, if $x_0 \in D$, then $\vartheta \in A_D(x_0)$. On the other hand, if $x_0 \in \overline{D} \setminus D$, then, by Proposition 4.13, x_0 is an accumulation point of D . $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \bar{x} \in (\mathcal{B}(x_0, \epsilon^2) \cap D) \setminus \{x_0\}$. Let $v := \bar{x} - x_0$. Then, $\delta := \sqrt{\|v\|} \in (0, \epsilon) \subset \mathbb{R}$. Let $\bar{v} := \delta^{-1}v$. Then, $\bar{v} \in \mathcal{B}(\vartheta, \epsilon)$ and $\bar{x} = x_0 + \delta\bar{v} \in \{x_0 + r\bar{u} \mid r \in (0, \epsilon) \subset \mathbb{R}, \bar{u} \in \mathcal{B}(\vartheta, \epsilon)\} \cap D \neq \emptyset$. Hence, $\vartheta \in A_D(x_0)$. Therefore, $\vartheta \in A_D(x_0)$ if $x_0 \in \overline{D}$.

$\forall u \in A_D(x_0)$, $\forall \alpha \in [0, \infty) \subset \mathbb{R}$, we will show that $\alpha u \in A_D(x_0)$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\alpha = 0$; Case 2: $\alpha > 0$. Case 1: $\alpha = 0$. Then, $\alpha u = \vartheta \in A_D(x_0)$. Case 2: $\alpha > 0$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\bar{\epsilon} = \min\{\alpha\epsilon, \epsilon/\alpha\} \in (0, \infty) \subset \mathbb{R}$.

By $u \in A_D(x_0)$, $\exists \bar{x} \in \{x_0 + r\bar{u} \mid r \in (0, \bar{\epsilon}) \subset \mathbb{R}, \bar{u} \in \mathcal{B}(u, \bar{\epsilon})\} \cap D$. Hence, $\bar{x} \in D$ and $\bar{x} = x_0 + r\bar{u} = x_0 + (r/\alpha)(\alpha\bar{u})$ with $r \in (0, \bar{\epsilon}) \subset \mathbb{R}$ and $\bar{u} \in \mathcal{B}(u, \bar{\epsilon})$. Then, $r/\alpha \in (0, \epsilon) \subset \mathbb{R}$ and $\alpha\bar{u} \in \mathcal{B}(\alpha u, \epsilon)$. Hence, $\bar{x} \in \{x_0 + \tilde{r}\tilde{u} \mid \tilde{r} \in (0, \epsilon) \subset \mathbb{R}, \tilde{u} \in \mathcal{B}(\alpha u, \epsilon)\} \cap D \neq \emptyset$. This implies that $\alpha u \in A_D(x_0)$. Hence, $A_D(x_0)$ is a cone.

$\forall u \in A_D(x_0)$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Proposition 3.3, $\exists u_1 \in A_D(x_0) \cap \mathcal{B}(u, \epsilon/2)$. By $u_1 \in A_D(x_0)$, $\exists \bar{x} \in \{x_0 + r\bar{u} \mid r \in (0, \epsilon/2) \subset \mathbb{R}, \bar{u} \in \mathcal{B}(u_1, \epsilon/2)\} \cap D$. Then, $\bar{x} \in \{x_0 + r\bar{u} \mid r \in (0, \epsilon) \subset \mathbb{R}, \bar{u} \in \mathcal{B}(u, \epsilon)\} \cap D \neq \emptyset$. Hence, $u \in A_D(x_0)$. Then, $A_D(x_0) \subseteq A_D(x_0)$ and $A_D(x_0)$ is closed. This completes the proof of the proposition. $\square$

Definition 9.3 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, and $x_0 \in D$. Assume that $\overline{\text{span}(A_D(x_0))} = \mathcal{X}$. Let $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be such that $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, we have

$$\|f(x) - f(x_0) - L(x - x_0)\| \leq \epsilon \|x - x_0\|$$

Then, L is called the (Fréchet) derivative of f at x_0 and denoted by $f^{(1)}(x_0)$ or $Df(x_0)$. When L exists, we will say that f is (Fréchet) differentiable at x_0 . Df or $f^{(1)}$ will denote the $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ -valued function whose domain of definition is $\text{dom}(f^{(1)}) := \{x \in D \mid Df(x) \in \mathcal{B}(\mathcal{X}, \mathcal{Y}) \text{ exists}\}$. If f is differentiable at x_0 , $\forall x_0 \in D$, we say f is (Fréchet) differentiable. In this case, $Df : D \rightarrow \mathcal{B}(\mathcal{X}, \mathcal{Y})$ or $f^{(1)} : D \rightarrow \mathcal{B}(\mathcal{X}, \mathcal{Y})$.

Clearly, when $\mathcal{X} = \mathcal{Y} = \mathbb{R}$ and $D = [a, b] \subseteq \mathcal{X}$ with $a < b$, then $Df(t)$ is simply the derivative of f at $t \in [a, b]$, as we know before.

Definition 9.4 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, $x_0 \in D$, and $u \in A_D(x_0)$. Let $M_D(x_0) := \{(r, \bar{u}) \in \mathbb{R} \times \mathcal{X} \mid r \in (0, +\infty) \subset \mathbb{R}, \bar{u} \in \mathcal{X}, x_0 + r\bar{u} \in D\}$ and $\bar{g} : M_D(x_0) \rightarrow \mathcal{Y}$ be given by $\bar{g}(r, \bar{u}) = r^{-1}(f(x_0 + r\bar{u}) - f(x_0))$, $\forall (r, \bar{u}) \in M_D(x_0)$. Clearly, $(0, u)$ is an accumulation point of $M_D(x_0)$ since $u \in A_D(x_0)$. The directional derivative of f at x_0 along u , denoted by $Df(x_0; u)$, is the limit $\lim_{(r, \bar{u}) \rightarrow (0, u)} \bar{g}(r, \bar{u})$, when it exists.

Clearly, the directional derivative is unique when it exists, since $\mathcal{Y}$ is Hausdorff.

Proposition 9.5 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, $x_0 \in D$, $\overline{\text{span}(A_D(x_0))} = \mathcal{X}$, $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be the Fréchet derivative of f at x_0 , and $u \in A_D(x_0)$. Then, $Df(x_0; u) = Lu$.

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\bar{\epsilon} = \epsilon/(2 + 2\|u\|) \in (0, \infty) \subset \mathbb{R}$. By $Df(x_0) = L$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\|f(x) - f(x_0) - L(x - x_0)\| \leq \bar{\epsilon} \|x - x_0\|$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta) \cap D$. Let $M_D(x_0)$ and $\bar{g} : M_D(x_0) \rightarrow \mathcal{Y}$ be as defined in Definition 9.4. Let $\bar{\delta} = \min\{\epsilon/(2 + 2\|L\|), \delta/(1 + \|u\|), 1\} \in$

$(0, \infty) \subset \mathbb{R}$. $\forall (r, \bar{u}) \in (M_D(x_0) \cap \mathcal{B}_{\mathbb{R} \times \mathcal{X}}((0, u), \bar{\delta})) \setminus \{(0, u)\}$, we have $\bar{x} := x_0 + r\bar{u} \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$. This implies that

$$\begin{aligned}
\|\bar{g}(r, \bar{u}) - Lu\| &= \|f(x_0 + r\bar{u}) - f(x_0) - rLu\|/r \\
&= \|f(\bar{x}) - f(x_0) - L(\bar{x} - x_0) + rL(\bar{u} - u)\|/r \\
&\leq \|f(\bar{x}) - f(x_0) - L(\bar{x} - x_0)\|/r + \|L(\bar{u} - u)\| \\
&\leq \bar{\epsilon} \|\bar{x} - x_0\|/r + \|L\| \|\bar{u} - u\| = \bar{\epsilon} \|\bar{u}\| + \|L\| \|\bar{u} - u\| \\
&\leq \bar{\epsilon}(\|u\| + \|\bar{u} - u\|) + \|L\| \|\bar{u} - u\| \leq \bar{\epsilon}(\|u\| + \bar{\delta}) + \|L\| \bar{\delta} \\
&< \epsilon
\end{aligned}$$

where the second inequality follows from Proposition 7.64. Hence, we have $Df(x_0; u) = \lim_{(r, \bar{u}) \rightarrow (0, u)} \bar{g}(r, \bar{u}) = Lu$. This completes the proof of the proposition. $\square$

Proposition 9.6 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, $x_0 \in D$, and $\text{span}(\mathcal{A}_D(x_0)) = \mathcal{X}$. Then, f has at most one derivative at x_0 .*

Proof We will prove the result by an argument of contradiction. Suppose, $\exists L_1, L_2 \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ with $L_1 \neq L_2$ such that L_1 and L_2 are derivatives of f at x_0 . By $\text{span}(\mathcal{A}_D(x_0)) = \mathcal{X}$ and Proposition 3.56, $\exists \bar{u} \in \text{span}(\mathcal{A}_D(x_0))$ such that $L_1\bar{u} \neq L_2\bar{u}$. Since L_1 and L_2 are linear, then $\exists u_0 \in \mathcal{A}_D(x_0)$ such that $L_1u_0 \neq L_2u_0$. By Proposition 9.5, $Df(x_0; u_0) = L_1u_0 = L_2u_0$. This is a contradiction. Hence, the result holds. This completes the proof of the proposition. $\square$

Proposition 9.7 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, $x_0 \in D$, and $\text{span}(\mathcal{A}_D(x_0)) = \mathcal{X}$. Assume that $Df(x_0) \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ exists, then f is continuous at x_0 .*

Proof This is straightforward, and is therefore omitted. $\square$

Definition 9.8 *Let X be a set, $\mathcal{Y}$ and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq X \times \mathcal{Y}$, $f : D \rightarrow \mathcal{Z}$, and $(x_0, y_0) \in D$. f is said to be partial (Fréchet) differentiable with respect to y at (x_0, y_0) if $g : D_1 \rightarrow \mathcal{Z}$ given by $g(y) = f(x_0, y)$, $\forall y \in D_1$, where $D_1 := \{y \in \mathcal{Y} \mid (x_0, y) \in D\}$, is differentiable at y_0 . $Dg(y_0) \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ is called the partial derivative of f with respect to y at (x_0, y_0) and is denoted by $\frac{\partial f}{\partial y}(x_0, y_0)$. $\frac{\partial f}{\partial y}$ will denote the $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$ -valued function whose domain of definition is $\text{dom}\left(\frac{\partial f}{\partial y}\right) := \left\{(x, y) \in D \mid \frac{\partial f}{\partial y}(x, y) \in \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \text{ exists}\right\}$. When $\text{dom}\left(\frac{\partial f}{\partial y}\right) = D$, we say that f is partial differentiable with respect to y .*

For notational simplicity, we will adopt the “matrix” notation for our later developments. Let $\mathcal{X}_1, \dots, \mathcal{X}_n$ be normed linear spaces over $\mathbb{K}$, $n \in \mathbb{N}$.

For a vector $(x_1, \dots, x_n) \in \mathcal{X}_1 \times \dots \times \mathcal{X}_n$, we will use the notation of a column vector

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

to alternatively denote it. Let $\mathcal{Y}_1, \dots, \mathcal{Y}_m$ be normed linear spaces over $\mathbb{K}$, $m \in \mathbb{N}$. $\forall L \in B(\mathcal{X}_1 \times \dots \times \mathcal{X}_n, \mathcal{Y}_1 \times \dots \times \mathcal{Y}_m)$ can be equivalently written as

$$\begin{bmatrix} L_{11} & \cdots & L_{1n} \\ \vdots & & \vdots \\ L_{m1} & \cdots & L_{mn} \end{bmatrix}$$

where $L_{ij} \in B(\mathcal{X}_j, \mathcal{Y}_i)$. Then, $(y_1, \dots, y_m) = L(x_1, \dots, x_n)$ can be equivalently expressed as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} L_{11} & \cdots & L_{1n} \\ \vdots & & \vdots \\ L_{m1} & \cdots & L_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n L_{1j}x_j \\ \vdots \\ \sum_{j=1}^n L_{mj}x_j \end{bmatrix}$$

Let $\mathcal{Z}_1, \dots, \mathcal{Z}_p$ be normed linear spaces over $\mathbb{K}$, $p \in \mathbb{N}$, and $\bar{L} \in B(\mathcal{Y}_1 \times \dots \times \mathcal{Y}_m, \mathcal{Z}_1 \times \dots \times \mathcal{Z}_p)$ be given by

$$\begin{bmatrix} \bar{L}_{11} & \cdots & \bar{L}_{1m} \\ \vdots & & \vdots \\ \bar{L}_{p1} & \cdots & \bar{L}_{pm} \end{bmatrix}$$

Then, $\bar{L}L \in B(\mathcal{X}_1 \times \dots \times \mathcal{X}_n, \mathcal{Z}_1 \times \dots \times \mathcal{Z}_p)$ and

$$\begin{aligned} \bar{L}L &= \begin{bmatrix} \bar{L}_{11} & \cdots & \bar{L}_{1m} \\ \vdots & & \vdots \\ \bar{L}_{p1} & \cdots & \bar{L}_{pm} \end{bmatrix} \begin{bmatrix} L_{11} & \cdots & L_{1n} \\ \vdots & & \vdots \\ L_{m1} & \cdots & L_{mn} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{j=1}^m \bar{L}_{1j}L_{j1} & \cdots & \sum_{j=1}^m \bar{L}_{1j}L_{jn} \\ \vdots & & \vdots \\ \sum_{j=1}^m \bar{L}_{pj}L_{j1} & \cdots & \sum_{j=1}^m \bar{L}_{pj}L_{jn} \end{bmatrix} \end{aligned}$$

It is easy to show that the adjoint of L is

$$L' = \begin{bmatrix} L'_{11} & \cdots & L'_{m1} \\ \vdots & & \vdots \\ L'_{1n} & \cdots & L'_{mn} \end{bmatrix}$$

Proposition 9.9 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X} \times \mathcal{Y}$, $f : D \rightarrow \mathcal{Z}$, and $(x_0, y_0) \in D$. Assume $Df(x_0, y_0) \in B(\mathcal{X} \times \mathcal{Y}, \mathcal{Z})$ exists. Let $D_{x_0} := \{y \in \mathcal{Y} \mid (x_0, y) \in D\}$. Assume that $\text{span}(\overline{A_{D_{x_0}}(y_0)}) =$*

$\mathcal{Y}$. Then, f is partial differentiable with respect to y at (x_0, y_0) and $\frac{\partial f}{\partial y}(x_0, y_0) \in B(\mathcal{Y}, \mathcal{Z})$ is given by $\frac{\partial f}{\partial y}(x_0, y_0)(k) = Df(x_0, y_0)(\vartheta_{\mathcal{X}}, k)$, $\forall k \in \mathcal{Y}$.

Let $D_{y_0} := \{x \in \mathcal{X} \mid (x, y_0) \in D\}$. Assume that $\overline{\text{span}(A_{D_{y_0}}(x_0))} = \mathcal{X}$. By symmetry, f is partial differentiable with respect to x at (x_0, y_0) and $\frac{\partial f}{\partial x}(x_0, y_0) \in B(\mathcal{X}, \mathcal{Z})$ is given by $\frac{\partial f}{\partial x}(x_0, y_0)(h) = Df(x_0, y_0)(h, \vartheta_{\mathcal{Y}})$, $\forall h \in \mathcal{X}$.

Hence, we have $Df(x_0, y_0) = \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix}$ in “matrix” notation.

Proof Let $\bar{g} : D_{x_0} \rightarrow \mathcal{Z}$ be given by $\bar{g}(y) = f(x_0, y)$, $\forall y \in D_{x_0}$. Clearly $y_0 \in D_{x_0}$. By the fact that f is differentiable at (x_0, y_0) , then $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta)$, we have $\|f(x, y) - f(x_0, y_0) - Df(x_0, y_0)(x - x_0, y - y_0)\| \leq \epsilon \|(x - x_0, y - y_0)\|$. Then, $\forall y \in D_{x_0} \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta)$, we have $(x_0, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta)$ and $\epsilon \|y - y_0\| \geq \|f(x_0, y) - f(x_0, y_0) - Df(x_0, y_0)(\vartheta_{\mathcal{X}}, y - y_0)\| = \|\bar{g}(y) - \bar{g}(y_0) - Df(x_0, y_0)(\vartheta_{\mathcal{X}}, y - y_0)\|$. Let $L : \mathcal{Y} \rightarrow \mathcal{Z}$ be given by $L(k) = Df(x_0, y_0)(\vartheta_{\mathcal{X}}, k)$, $\forall k \in \mathcal{Y}$. Clearly, L is a linear operator. Note that

$$\begin{aligned} \|L\| &= \sup_{k \in \mathcal{Y}, \|k\| \leq 1} \|Lk\| = \sup_{k \in \mathcal{Y}, \|k\| \leq 1} \|Df(x_0, y_0)(\vartheta_{\mathcal{X}}, k)\| \\ &\leq \sup_{k \in \mathcal{Y}, \|k\| \leq 1} \|Df(x_0, y_0)\| \|(\vartheta_{\mathcal{X}}, k)\| \leq \|Df(x_0, y_0)\| < +\infty \end{aligned}$$

where the first inequality follows from Proposition 7.64 and the last inequality follows from the fact that $Df(x_0, y_0) \in B(\mathcal{X} \times \mathcal{Y}, \mathcal{Z})$. Hence, $L \in B(\mathcal{Y}, \mathcal{Z})$. Then, $\|\bar{g}(y) - \bar{g}(y_0) - L(y - y_0)\| \leq \epsilon \|y - y_0\|$, $\forall y \in D_{x_0} \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta)$. This implies that $D\bar{g}(y_0) = L$. By Definition 9.8, $\frac{\partial f}{\partial y}(x_0, y_0) = L$.

By symmetry, f is partial differentiable with respect to x at (x_0, y_0) and $\frac{\partial f}{\partial x}(x_0, y_0) \in B(\mathcal{X}, \mathcal{Z})$ is given by $\frac{\partial f}{\partial x}(x_0, y_0)(h) = Df(x_0, y_0)(h, \vartheta_{\mathcal{Y}})$, $\forall h \in \mathcal{X}$.

Note that, $\forall h \in \mathcal{X}$ and $\forall k \in \mathcal{Y}$,

$$Df(x_0, y_0)(h, k) = \frac{\partial f}{\partial x}(x_0, y_0)(h) + \frac{\partial f}{\partial y}(x_0, y_0)(k)$$

Then,

$$Df(x_0, y_0) \begin{bmatrix} h \\ k \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}$$

Hence, the desired “matrix” notation follows. This completes the proof of the proposition. $\square$

9.2 The Derivatives of Some Common Functions

Proposition 9.10 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, $x_0 \in D$, and $\overline{\text{span}(A_D(x_0))} = \mathcal{X}$. Assume that $\exists \delta_0 \in (0, \infty) \subset$

$\mathbb{R}$ and $\exists y_0 \in \mathcal{Y}$ such that $f(x) = y_0$, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$. Then, f is Fréchet differentiable at x_0 and $Df(x_0) = \vartheta_{\mathcal{B}(\mathcal{X}, \mathcal{Y})}$.

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.11 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $x_0 \in D_2 \subseteq D_1 \subseteq \mathcal{X}$, and $f : D_1 \rightarrow \mathcal{Y}$. If f is Fréchet differentiable at x_0 and $\text{span}(\mathcal{A}_{D_2}(x_0)) = \mathcal{X}$, then $g := f|_{D_2}$ is Fréchet differentiable at x_0 and $Dg(x_0) = Df(x_0)$. On the other hand, if g is Fréchet differentiable at x_0 and $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) = D_2 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, then f is Fréchet differentiable at x_0 and $Df(x_0) = Dg(x_0)$.

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.12 Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, and $f : \mathcal{X} \rightarrow \mathcal{X}$ be given by $f = \text{id}_{\mathcal{X}}$, that is $f(x) = x$, $\forall x \in \mathcal{X}$. Then, f is Fréchet differentiable and $Df(x)(h) = h$, $\forall x \in \mathcal{X}$ and $\forall h \in \mathcal{X}$. Then, $Df(x) = \text{id}_{\mathcal{X}}$, $\forall x \in \mathcal{X}$.

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.13 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, and $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{X}$ be given by $f = \pi_{\mathcal{X}}$, that is $f(x, y) = x$, $\forall (x, y) \in \mathcal{X} \times \mathcal{Y}$. Then, f is Fréchet differentiable and $Df(x, y)(h, k) = h$, $\forall (x, y) \in \mathcal{X} \times \mathcal{Y}$ and $\forall (h, k) \in \mathcal{X} \times \mathcal{Y}$. In “matrix” notation, we have $Df(x, y) = \begin{bmatrix} \text{id}_{\mathcal{X}} & \vartheta_{\mathcal{B}(\mathcal{Y}, \mathcal{X})} \end{bmatrix}$, $\forall (x, y) \in \mathcal{X} \times \mathcal{Y}$.

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.14 Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $f : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ be given by $f(x_1, x_2) = x_1 + x_2$, $\forall (x_1, x_2) \in \mathcal{X} \times \mathcal{X}$. Then, f is Fréchet differentiable and $Df : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{B}(\mathcal{X} \times \mathcal{X}, \mathcal{X})$ is given by $\forall (x_1, x_2) \in \mathcal{X} \times \mathcal{X}$, $\forall (h_1, h_2) \in \mathcal{X} \times \mathcal{X}$, $Df(x_1, x_2)(h_1, h_2) = h_1 + h_2$. In “matrix” notation, we have $Df(x_1, x_2) = \begin{bmatrix} \text{id}_{\mathcal{X}} & \text{id}_{\mathcal{X}} \end{bmatrix}$.

Proof $\forall (x_1, x_2) \in \mathcal{X} \times \mathcal{X}$, let $L : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ be given by $L(h_1, h_2) = h_1 + h_2$, $\forall (h_1, h_2) \in \mathcal{X} \times \mathcal{X}$. Clearly, L is a linear operator. Note that

$$\begin{aligned} \|L\| &= \sup_{(h_1, h_2) \in \mathcal{X} \times \mathcal{X}, \|(h_1, h_2)\| \leq 1} \|L(h_1, h_2)\| \\ &\leq \sup_{(h_1, h_2) \in \mathcal{X} \times \mathcal{X}, \|(h_1, h_2)\| \leq 1} \|h_1\| + \|h_2\| \\ &\leq \sup_{(h_1, h_2) \in \mathcal{X} \times \mathcal{X}, \|(h_1, h_2)\| \leq 1} \sqrt{2} \left(\|h_1\|^2 + \|h_2\|^2 \right)^{1/2} \leq \sqrt{2} \end{aligned}$$

where the second inequality follows from Cauchy-Schwarz Inequality. Hence, $L \in \mathcal{B}(\mathcal{X} \times \mathcal{X}, \mathcal{X})$.

Clearly, $A_{\mathcal{X} \times \mathcal{X}}(x_1, x_2) = \mathcal{X} \times \mathcal{X}$ since $(x_1, x_2) \in (\mathcal{X} \times \mathcal{X})^\circ$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, set $\delta = 1 \in \mathbb{R}$, $\forall (\bar{x}_1, \bar{x}_2) \in (\mathcal{X} \times \mathcal{X}) \cap \mathcal{B}_{\mathcal{X} \times \mathcal{X}}((x_1, x_2), \delta)$, we have $\|f(\bar{x}_1, \bar{x}_2) - f(x_1, x_2) - L(\bar{x}_1 - x_1, \bar{x}_2 - x_2)\| = 0 \leq \epsilon \|(\bar{x}_1 - x_1, \bar{x}_2 - x_2)\|$. By Definition 9.3, $Df(x_1, x_2) = L$. This completes the proof of the proposition. $\square$

Proposition 9.15 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f_1 : D \rightarrow \mathcal{Y}$, $f_2 : D \rightarrow \mathcal{Y}$, $x_0 \in D$, $\alpha_1, \alpha_2 \in \mathbb{K}$, and $g : D \rightarrow \mathcal{Y}$ be given by $g(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x)$, $\forall x \in D$. Assume that $f_1^{(1)}(x_0)$ and $f_2^{(1)}(x_0)$ exist. Then, g is Fréchet differentiable at x_0 and $g^{(1)}(x_0) = \alpha_1 f_1^{(1)}(x_0) + \alpha_2 f_2^{(1)}(x_0)$.*

Proof Define $L := \alpha_1 f_1^{(1)}(x_0) + \alpha_2 f_2^{(1)}(x_0) \in B(\mathcal{X}, \mathcal{Y})$. By assumption, $\overline{\text{span}(A_D(x_0))} = \mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the differentiability of f_1 at x_0 , $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_1)$, we have $\|f_1(x) - f_1(x_0) - f_1^{(1)}(x_0)(x - x_0)\| \leq \epsilon \|x - x_0\|$. By the differentiability of f_2 at x_0 , $\exists \delta_2 \in (0, \infty) \subset \mathbb{R}$ such that $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_2)$, we have $\|f_2(x) - f_2(x_0) - f_2^{(1)}(x_0)(x - x_0)\| \leq \epsilon \|x - x_0\|$. Let $\delta := \min\{\delta_1, \delta_2\} > 0$. $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, we have

$$\begin{aligned} & \|g(x) - g(x_0) - L(x - x_0)\| \\ &= \|\alpha_1 (f_1(x) - f_1(x_0) - f_1^{(1)}(x_0)(x - x_0)) + \alpha_2 (f_2(x) - f_2(x_0) - f_2^{(1)}(x_0)(x - x_0))\| \\ &\leq (|\alpha_1| + |\alpha_2|)\epsilon \|x - x_0\| \end{aligned}$$

Hence, $g^{(1)}(x_0) = L$. This completes the proof of the proposition. $\square$

Proposition 9.16 *Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $f : \mathbb{K} \times \mathcal{X} \rightarrow \mathcal{X}$ be given by $f(\alpha, x) = \alpha x$, $\forall (\alpha, x) \in \mathbb{K} \times \mathcal{X}$. Then, f is Fréchet differentiable and $Df : \mathbb{K} \times \mathcal{X} \rightarrow B(\mathbb{K} \times \mathcal{X}, \mathcal{X})$ is given by $\forall (\alpha, x) \in \mathbb{K} \times \mathcal{X}$, $\forall (d, h) \in \mathbb{K} \times \mathcal{X}$, $Df(\alpha, x)(d, h) = \alpha h + dx$. Thus, $Df(\alpha, x) = \begin{bmatrix} x & \alpha \text{id}_{\mathcal{X}} \end{bmatrix}$ in “matrix” notation.*

Proof $\forall (\alpha, x) \in \mathbb{K} \times \mathcal{X}$, let $L : \mathbb{K} \times \mathcal{X} \rightarrow \mathcal{X}$ be given by $L(d, h) = \alpha h + dx$, $\forall (d, h) \in \mathbb{K} \times \mathcal{X}$. Clearly, L is a linear operator. Note that

$$\begin{aligned} \|L\| &= \sup_{(d, h) \in \mathbb{K} \times \mathcal{X}, \|(d, h)\| \leq 1} \|L(d, h)\| \\ &\leq \sup_{(d, h) \in \mathbb{K} \times \mathcal{X}, \|(d, h)\| \leq 1} |\alpha| \|h\| + |d| \|x\| \\ &\leq \sup_{(d, h) \in \mathbb{K} \times \mathcal{X}, \|(d, h)\| \leq 1} \left(|\alpha|^2 + \|x\|^2\right)^{1/2} \left(|d|^2 + \|h\|^2\right)^{1/2} \\ &\leq \|(\alpha, x)\| < +\infty \end{aligned}$$

where the second inequality follows from Cauchy-Schwarz Inequality. Hence, $L \in B(\mathbb{K} \times \mathcal{X}, \mathcal{X})$.

Clearly, $A_{\mathbb{K} \times \mathcal{X}}(\alpha, x) = \mathbb{K} \times \mathcal{X}$ since $(\alpha, x) \in (\mathbb{K} \times \mathcal{X})^\circ$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, set $\delta = 2\epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall (\bar{\alpha}, \bar{x}) \in (\mathbb{K} \times \mathcal{X}) \cap \mathcal{B}_{\mathbb{K} \times \mathcal{X}}((\alpha, x), \delta)$, we have

$$\begin{aligned} & \|f(\bar{\alpha}, \bar{x}) - f(\alpha, x) - L(\bar{\alpha} - \alpha, \bar{x} - x)\| \\ &= \|\bar{\alpha}\bar{x} - \alpha x - \alpha(\bar{x} - x) - (\bar{\alpha} - \alpha)x\| = \|(\bar{\alpha} - \alpha)(\bar{x} - x)\| \\ &\leq \frac{1}{2}(\|\bar{\alpha} - \alpha\|^2 + \|\bar{x} - x\|^2) = \frac{1}{2}\|(\bar{\alpha} - \alpha, \bar{x} - x)\|^2 \\ &\leq \epsilon \|(\bar{\alpha} - \alpha, \bar{x} - x)\| \end{aligned}$$

By Definition 9.3, $Df(\alpha, x) = L$. This completes the proof of the proposition. $\square$

In the previous proposition, we have abuse the notation using the “matrix” notation that we identify xd as dx .

To state the next proposition, we will introduce a new notation. Let $A \in B(\mathcal{X}, \mathcal{Y})$ and $B \in B(\mathcal{Y}, \mathcal{Z})$, where $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are normed linear spaces over $\mathbb{K}$. Clearly, $f(A, B) := BA \in B(\mathcal{X}, \mathcal{Z})$. Let $g(A) : B(\mathcal{Y}, \mathcal{Z}) \rightarrow B(\mathcal{X}, \mathcal{Z})$ be given by $g(A)(B) = BA$, $\forall B \in B(\mathcal{Y}, \mathcal{Z})$. Clearly, $g(A)$ is a bounded linear operator. It is easy to see that $g : B(\mathcal{X}, \mathcal{Y}) \rightarrow B(B(\mathcal{Y}, \mathcal{Z}), B(\mathcal{X}, \mathcal{Z}))$ is a bounded linear operator with $\|g\| \leq 1$. This operator g is needed in compact “matrix” notation for many linear operators. We will use the ro to denote g for any normed linear spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$. The meaning of ro is “right operate”. For $x \in \mathcal{X}$, we will identify $\mathcal{X}$ with $B(\mathbb{R}, \mathcal{X})$ and write $ro(x)(A) = Ax$.

This brings us to the next proposition.

Proposition 9.17 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $f : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \rightarrow \mathcal{Y}$ be given by $f(A, x) = Ax$, $\forall (A, x) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$. Then, f is Fréchet differentiable and $Df : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \rightarrow B(B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \mathcal{Y})$ is given by $\forall (A, x) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$, $\forall (\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$, $Df(A, x)(\Delta, h) = Ah + \Delta x$. In “matrix” notation, $Df(A, x) = \begin{bmatrix} ro(x) & A \end{bmatrix}$.*

Proof $\forall (A, x) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$, let $L : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \rightarrow \mathcal{Y}$ be given by $L(\Delta, h) = Ah + \Delta x$, $\forall (\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$. Clearly, L is a linear operator. Note that

$$\begin{aligned} \|L\| &= \sup_{(\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \|(\Delta, h)\| \leq 1} \|L(\Delta, h)\| \\ &\leq \sup_{(\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \|(\Delta, h)\| \leq 1} \|A\| \|h\| + \|\Delta\| \|x\| \\ &\leq \sup_{(\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \|(\Delta, h)\| \leq 1} \left(\|A\|^2 + \|x\|^2 \right)^{1/2} \left(\|\Delta\|^2 + \|h\|^2 \right)^{1/2} \\ &\leq \|(A, x)\| < +\infty \end{aligned}$$

where the first inequality follows from Proposition 7.64, and the second inequality follows from Cauchy-Schwarz Inequality. Hence, $L \in B(B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \mathcal{Y})$.

Clearly, $A_{B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}}(A, x) = B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$ since $(A, x) \in (B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X})^\circ = B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, set $\delta = 2\epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall (\bar{A}, \bar{x}) \in (B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}) \cap \mathcal{B}_{B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}}((A, x), \delta)$, we have

$$\begin{aligned} & \|f(\bar{A}, \bar{x}) - f(A, x) - L(\bar{A} - A, \bar{x} - x)\| \\ &= \|\bar{A}\bar{x} - Ax - A(\bar{x} - x) - (\bar{A} - A)x\| = \|(\bar{A} - A)(\bar{x} - x)\| \\ &\leq \frac{1}{2}(\|\bar{A} - A\|^2 + \|\bar{x} - x\|^2) = \frac{1}{2}\|(\bar{A} - A, \bar{x} - x)\|^2 \\ &\leq \epsilon \|(\bar{A} - A, \bar{x} - x)\| \end{aligned}$$

where the first inequality follows from Proposition 7.64. By Definition 9.3, $Df(A, x) = L$. This completes the proof of the proposition. $\square$

9.3 Chain Rule and Mean Value Theorem

Theorem 9.18 (Chain Rule) *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D_x \subseteq \mathcal{X}$, $D_y \subseteq \mathcal{Y}$, $f : D_x \rightarrow D_y$, $g : D_y \rightarrow \mathcal{Z}$, $x_0 \in D_x$, $y_0 := f(x_0) \in D_y$, and $h := g \circ f : D_x \rightarrow \mathcal{Z}$. Assume that f is Fréchet differentiable at x_0 with $Df(x_0) \in B(\mathcal{X}, \mathcal{Y})$ and g is Fréchet differentiable at y_0 with $Dg(y_0) \in B(\mathcal{Y}, \mathcal{Z})$. Then, h is differentiable at x_0 and $Dh(x_0) \in B(\mathcal{X}, \mathcal{Z})$ is given by*

$$Dh(x_0) = Dg(f(x_0)) \circ Df(x_0) = Dg(y_0)Df(x_0)$$

Proof Define $L : \mathcal{X} \rightarrow \mathcal{Z}$ by, $\forall \bar{x} \in \mathcal{X}$, $L(\bar{x}) = Dg(y_0)Df(x_0)\bar{x} = Dg(y_0)(Df(x_0)(\bar{x}))$. Clearly, L is a linear operator and, by Proposition 7.64, $L \in B(\mathcal{X}, \mathcal{Z})$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\bar{\epsilon} = (\sqrt{4\epsilon + (\|Df(x_0)\| + \|Dg(y_0)\|)^2} - \|Df(x_0)\| - \|Dg(y_0)\|)/2 \in (0, \infty) \subset \mathbb{R}$. By the fact that $Dg(y_0) \in B(\mathcal{Y}, \mathcal{Z})$, $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $\forall y \in D_y \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta_1)$, we have

$$\|g(y) - g(y_0) - Dg(y_0)(y - y_0)\| \leq \bar{\epsilon} \|y - y_0\|$$

By the fact that $Df(x_0) \in B(\mathcal{X}, \mathcal{Y})$, we have $\overline{\text{span}(A_{D_x}(x_0))} = \mathcal{X}$ and $\exists \delta_2 \in (0, \delta_1/(\bar{\epsilon} + \|Df(x_0)\|)) \subset \mathbb{R}$ such that $\forall x \in D_x \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_2)$, we have

$$\|f(x) - f(x_0) - Df(x_0)(x - x_0)\| \leq \bar{\epsilon} \|x - x_0\|$$

Then, by Proposition 7.64, $\|f(x) - y_0\| = \|f(x) - f(x_0)\| \leq \bar{\epsilon} \|x - x_0\| + \|Df(x_0)\| \|x - x_0\| \leq (\bar{\epsilon} + \|Df(x_0)\|) \|x - x_0\| < \delta_1$. This implies that $f(x) \in D_y \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta_1)$. Then, we have

$$\begin{aligned} & \|h(x) - h(x_0) - L(x - x_0)\| = \|g(f(x)) - g(y_0) - L(x - x_0)\| \\ &\leq \|g(f(x)) - g(y_0) - Dg(y_0)(f(x) - y_0)\| \\ &\quad + \|Dg(y_0)(f(x) - f(x_0) - Df(x_0)(x - x_0))\| \end{aligned}$$

$$\begin{aligned}
&\leq \bar{\epsilon} \|f(x) - y_0\| + \|Dg(y_0)\| \|f(x) - f(x_0) - Df(x_0)(x - x_0)\| \\
&\leq \bar{\epsilon} (\bar{\epsilon} + \|Df(x_0)\|) \|x - x_0\| + \bar{\epsilon} \|Dg(y_0)\| \|x - x_0\| \\
&= \bar{\epsilon} (\bar{\epsilon} + \|Df(x_0)\| + \|Dg(y_0)\|) \|x - x_0\| = \epsilon \|x - x_0\|
\end{aligned}$$

where the second inequality follows from Proposition 7.64. By Definition 9.3, $Dh(x_0) = L$. This completes the proof of the proposition. $\square$

Proposition 9.19 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f_1 : D \rightarrow \mathcal{Y}$, $f_2 : D \rightarrow \mathcal{Z}$, $x_0 \in D$, and $g : D \rightarrow \mathcal{Y} \times \mathcal{Z}$ be given by $g(x) = (f_1(x), f_2(x))$, $\forall x \in D$. Then, the following statement holds. g is Fréchet differentiable at x_0 if, and only if, f_1 and f_2 are Fréchet differentiable at x_0 . In this case, $Dg(x_0)(h) = (Df_1(x_0)(h), Df_2(x_0)(h))$, $\forall h \in \mathcal{X}$. In “matrix” notation, $Dg(x_0) = \begin{bmatrix} Df_1(x_0) \\ Df_2(x_0) \end{bmatrix}$.*

Proof “Sufficiency” By the differentiability of f_1 and f_2 at x_0 , we have $\text{span}(A_D(x_0)) = \mathcal{X}$. Define $L : \mathcal{X} \rightarrow \mathcal{Y} \times \mathcal{Z}$ by, $\forall h \in \mathcal{X}$, $L(h) = (Df_1(x_0)(h), Df_2(x_0)(h))$. Clearly, L is a linear operator. Note that

$$\begin{aligned}
\|L\| &= \sup_{h \in \mathcal{X}, \|h\| \leq 1} \|L(h)\| \\
&= \sup_{h \in \mathcal{X}, \|h\| \leq 1} \sqrt{\|Df_1(x_0)(h)\|^2 + \|Df_2(x_0)(h)\|^2} \\
&\leq \sup_{h \in \mathcal{X}, \|h\| \leq 1} \sqrt{\|Df_1(x_0)\|^2 \|h\|^2 + \|Df_2(x_0)\|^2 \|h\|^2} \\
&\leq \sqrt{\|Df_1(x_0)\|^2 + \|Df_2(x_0)\|^2} < +\infty
\end{aligned}$$

where the first inequality follows from Proposition 7.64 and the last inequality follows from the fact $Df_1(x_0) \in B(\mathcal{X}, \mathcal{Y})$ and $Df_2(x_0) \in B(\mathcal{X}, \mathcal{Z})$. Hence, $L \in B(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$.

Since f_1 and f_2 are differentiable at x_0 , then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_1)$, we have $\|f_1(x) - f_1(x_0) - Df_1(x_0)(x - x_0)\| \leq \epsilon/\sqrt{2} \|x - x_0\|$; $\exists \delta_2 \in (0, \infty) \subset \mathbb{R}$ such that $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_2)$, we have $\|f_2(x) - f_2(x_0) - Df_2(x_0)(x - x_0)\| \leq \epsilon/\sqrt{2} \|x - x_0\|$. Let $\delta = \min\{\delta_1, \delta_2\} \in (0, \infty) \subset \mathbb{R}$. $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, we have

$$\begin{aligned}
&\|g(x) - g(x_0) - L(x - x_0)\| \\
&= \|(f_1(x), f_2(x)) - (f_1(x_0), f_2(x_0)) \\
&\quad - (Df_1(x_0)(x - x_0), Df_2(x_0)(x - x_0))\| \\
&= \|(f_1(x) - f_1(x_0) - Df_1(x_0)(x - x_0), f_2(x) \\
&\quad - f_2(x_0) - Df_2(x_0)(x - x_0))\| \\
&= (\|f_1(x) - f_1(x_0) - Df_1(x_0)(x - x_0)\|^2 \\
&\quad + \|f_2(x) - f_2(x_0) - Df_2(x_0)(x - x_0)\|^2)^{1/2}
\end{aligned}$$

$$\leq \epsilon \|x - x_0\|$$

Hence, g is differentiable at x_0 and $Dg(x_0) = L$.

“Necessity” By the differentiability of g at x_0 , we have $\overline{\text{span}(\mathbf{A}_D(x_0))} = \mathcal{X}$. Note that $f_1 = \pi_{\mathcal{Y}} \circ g$. By Chain Rule and Proposition 9.13, f_1 is Fréchet differentiable at x_0 and

$$Df_1(x_0) = \begin{bmatrix} \text{id}_{\mathcal{Y}} & \vartheta_{\mathbf{B}(\mathcal{Z}, \mathcal{Y})} \end{bmatrix} Dg(x_0)$$

By symmetry, f_2 is Fréchet differentiable at x_0 and

$$Df_2(x_0) = \begin{bmatrix} \vartheta_{\mathbf{B}(\mathcal{Y}, \mathcal{Z})} & \text{id}_{\mathcal{Z}} \end{bmatrix} Dg(x_0)$$

$$\text{Then, } Dg(x_0) = \begin{bmatrix} Df_1(x_0) \\ Df_2(x_0) \end{bmatrix}.$$

This completes the proof of the proposition. $\square$

Theorem 9.20 (Mean Value Theorem) *Let $\mathcal{X}$ be a real normed linear space, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathbb{R}$, $x_1, x_2 \in D$, and $\varphi : [0, 1] \rightarrow D$ be given by $\varphi(t) = tx_1 + (1-t)x_2$, $\forall t \in I := [0, 1] \subset \mathbb{R}$. Assume that f is continuous at $\varphi(t)$, $\forall t \in I$ and f is Fréchet differentiable at $\varphi(t)$, $\forall t \in I^\circ$. Then, there exists $t_0 \in I^\circ$ such that*

$$f(x_1) - f(x_2) = Df(\varphi(t_0))(x_1 - x_2)$$

Proof By Propositions 9.19, 9.16, and 9.15 and Chain Rule, φ is Fréchet differentiable and $D\varphi(t)(h) = h(x_1 - x_2)$, $\forall h \in \mathbb{R}, \forall t \in I$. Define $g : I \rightarrow \mathbb{R}$ by $g(t) = f(\varphi(t))$, $\forall t \in I$. By Chain Rule, g is Fréchet differentiable at t , $\forall t \in I^\circ$ and $Dg(t)(h) = Df(\varphi(t))(D\varphi(t)(h)) = hDf(\varphi(t))(x_1 - x_2)$, $\forall h \in \mathbb{R}, \forall t \in I^\circ$. Then, $Dg(t) = Df(\varphi(t))(x_1 - x_2)$, $\forall t \in I^\circ$. By Proposition 3.12, g is continuous. By Mean Value Theorem (Bartle, 1976, see Theorem 27.6), we have

$$\begin{aligned} f(x_1) - f(x_2) &= g(1) - g(0) = Dg(t_0)(1 - 0) \\ &= Df(t_0x_1 + (1 - t_0)x_2)(x_1 - x_2) \end{aligned}$$

for some $t_0 \in I^\circ$. This completes the proof of the theorem. $\square$

Toward a general Mean Value Theorem for vector-valued functions on possibly complex normed linear spaces, we present the following two lemmas.

Lemma 9.21 *Let $D := \{a + i0 \mid a \in I := [0, 1] \subset \mathbb{R}\} \subset \mathbb{C}$, $x_0 := x_r + ix_i \in D$, where $x_r \in I$ and $x_i = 0$, $f : D \rightarrow \mathbb{C}$ be Fréchet differentiable at x_0 , $Df(x_0) = d_r + id_i \in \mathbb{C}$, where $d_r, d_i \in \mathbb{R}$, and $g : I \rightarrow \mathbb{R}$ be given by $g(t) = \text{Re}(f(t + i0))$, $\forall t \in I$. Then, g is Fréchet differentiable at x_r and $Dg(x_r) = d_r = \text{Re}(Df(x_0)) \in \mathbb{R}$.*

Proof Note that $\text{span}(A_I(x_r)) = \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the differentiability of f at x_0 , $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\forall \bar{x} = \bar{x}_r + i\bar{x}_i \in D \cap \mathcal{B}_{\mathbb{C}}(x_0, \delta)$, we have $|f(\bar{x}) - f(x_0) - Df(x_0)(\bar{x} - x_0)| \leq \epsilon |\bar{x} - x_0|$. Note that $\bar{x}_r \in I$ and $\bar{x}_i = 0$. Then, the above implies that $|\text{Re}(f(\bar{x}_r)) - \text{Re}(f(x_r)) - (\bar{x}_r - x_r)d_r + i(\text{Im}(f(\bar{x}_r)) - \text{Im}(f(x_r)) - (\bar{x}_r - x_r)d_i)| \leq \epsilon |\bar{x}_r - x_r|$. This further implies that $|\text{Re}(f(\bar{x}_r)) - \text{Re}(f(x_r)) - (\bar{x}_r - x_r)d_r| \leq \epsilon |\bar{x}_r - x_r|$. Then, $\forall t \in I \cap \mathcal{B}_{\mathbb{R}}(x_r, \delta)$, $\bar{x} := t + i0 \in D \cap \mathcal{B}_{\mathbb{C}}(x_0, \delta)$ and $|g(t) - g(x_r) - (t - x_r)d_r| \leq \epsilon |t - x_r|$. Hence, $Dg(x_r) = d_r$. This completes the proof of the lemma. $\square$

Lemma 9.22 Let $D := \{a + i0 \mid a \in I := [0, 1] \subset \mathbb{R}\} \subset \mathbb{C}$, $f : D \rightarrow \mathbb{C}$ be continuous, f be Fréchet differentiable at $a + i0$, $\forall a \in I^\circ$. Then, $\exists t_0 \in I^\circ$ such that $\text{Re}(f(1) - f(0)) = \text{Re}(Df(t_0))$.

Proof Let $g : I \rightarrow \mathbb{R}$ be given by $g(t) = \text{Re}(f(t + i0))$, $\forall t \in I$. Clearly, g is continuous since f is continuous. By Lemma 9.21, g is Fréchet differentiable at t , $\forall t \in I^\circ$, and $Dg(t) = \text{Re}(Df(t + i0))$, $\forall t \in I^\circ$. By Mean Value Theorem (Bartle, 1976, see Theorem 27.6), $\exists t_0 \in I^\circ$ such that $g(1) - g(0) = Dg(t_0) = \text{Re}(Df(t_0 + i0))$. Then, $\text{Re}(f(1) - f(0)) = g(1) - g(0) = \text{Re}(Df(t_0 + i0)) = \text{Re}(Df(t_0))$. This completes the proof of the lemma. $\square$

Theorem 9.23 (Mean Value Theorem) Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, $x_1, x_2 \in D$, and the line segment connecting x_1 and x_2 be contained in D . Assume that f is continuous at $x = tx_1 + (1 - t)x_2$, $\forall t \in I := [0, 1] \subset \mathbb{R}$ and $Df(x) \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ exists at $x = tx_1 + (1 - t)x_2$, $\forall t \in I^\circ$. Then, $\exists t_0 \in I^\circ$ such that $\|f(x_1) - f(x_2)\| \leq \|Df(t_0x_1 + (1 - t_0)x_2)(x_1 - x_2)\|$.

Proof We will distinguish three exhaustive and mutually exclusive cases: Case 1: $f(x_1) = f(x_2)$; Case 2: $f(x_1) \neq f(x_2)$ and $\mathbb{K} = \mathbb{R}$; Case 3: $f(x_1) \neq f(x_2)$ and $\mathbb{K} = \mathbb{C}$. Case 1: $f(x_1) = f(x_2)$. Take t_0 to be any point in I° . The desired result follows.

Case 2: $f(x_1) \neq f(x_2)$ and $\mathbb{K} = \mathbb{R}$. By Proposition 7.85, $\exists y_* \in \mathcal{Y}^*$ with $\|y_*\| = 1$ such that $\langle y_*, f(x_1) - f(x_2) \rangle = \|f(x_1) - f(x_2)\|$. Define $\varphi : I \rightarrow D$ by $\varphi(t) = tx_1 + (1 - t)x_2$, $\forall t \in I$. By Propositions 9.19, 9.16, and 9.15 and Chain Rule, φ is Fréchet differentiable. Define $g : I \rightarrow \mathbb{R}$ by $g(t) = \langle y_*, f(\varphi(t)) \rangle$, $\forall t \in I$. By Proposition 3.12, g is continuous. By Chain Rule and Propositions 9.17 and 9.19, g is Fréchet differentiable at t , $\forall t \in I^\circ$, and $Dg(t)(d) = \langle y_*, Df(\varphi(t))(D\varphi(t)(d)) \rangle = \langle y_*, Df(\varphi(t))(d(x_1 - x_2)) \rangle = \langle y_*, Df(\varphi(t))(x_1 - x_2) \rangle d$, $\forall d \in \mathbb{R}$ and $\forall t \in I^\circ$. Hence, $Dg(t) = \langle y_*, Df(\varphi(t))(x_1 - x_2) \rangle$, $\forall t \in I^\circ$. By Mean Value Theorem (Bartle, 1976, see Theorem 27.6), there exists $t_0 \in I^\circ$ such that $g(1) - g(0) = Dg(t_0)$. Then, we have $\|f(x_1) - f(x_2)\| = \langle y_*, f(x_1) - f(x_2) \rangle = g(1) - g(0) = Dg(t_0) \leq |Dg(t_0)| = |\langle y_*, Df(\varphi(t_0))(x_1 - x_2) \rangle| \leq \|Df(\varphi(t_0))(x_1 - x_2)\|$, where the last inequality follows from Proposition 7.72. The desired result follows.

Case 3: $f(x_1) \neq f(x_2)$ and $\mathbb{K} = \mathbb{C}$. By Proposition 7.85, $\exists y_* \in \mathcal{Y}^*$ with $\|y_*\| = 1$ such that $\langle\langle y_*, f(x_1) - f(x_2) \rangle\rangle = \|f(x_1) - f(x_2)\|$. Let $\bar{D} := \{a + i0 \mid a \in I\} \subset \mathbb{C}$. Define $\varphi : \bar{D} \rightarrow D$ by $\varphi(t) = tx_1 + (1-t)x_2$, $\forall t \in \bar{D}$. By Propositions 9.19, 9.16, and 9.15 and Chain Rule, φ is Fréchet differentiable and $D\varphi(t)(d) = d(x_1 - x_2)$, $\forall d \in \mathbb{C}$ and $\forall t \in \bar{D}$. Define $g : \bar{D} \rightarrow \mathbb{C}$ by $g(t) = \langle\langle y_*, f(\varphi(t)) \rangle\rangle$, $\forall t \in \bar{D}$. By Proposition 3.12, g is continuous. By Chain Rule and Propositions 9.17 and 9.19, g is Fréchet differentiable at $a + i0$, $\forall a \in I^\circ$, and $Dg(a + i0)(d) = \langle\langle y_*, Df(\varphi(a + i0))(D\varphi(a + i0)(d)) \rangle\rangle = \langle\langle y_*, Df(\varphi(a))(d(x_1 - x_2)) \rangle\rangle = \langle\langle y_*, Df(\varphi(a))(x_1 - x_2) \rangle\rangle d$, $\forall d \in \mathbb{C}$ and $\forall a \in I^\circ$. Hence, $Dg(a + i0) = \langle\langle y_*, Df(\varphi(a + i0))(x_1 - x_2) \rangle\rangle$, $\forall a \in I^\circ$. Note that $g(1) - g(0) = \langle\langle y_*, f(x_1) \rangle\rangle - \langle\langle y_*, f(x_2) \rangle\rangle = \|f(x_1) - f(x_2)\| \in \mathbb{R}$. By Lemma 9.22, there exists $t_0 \in I^\circ$ such that $g(1) - g(0) = \operatorname{Re}(Dg(t_0))$. Then, we have $\|f(x_1) - f(x_2)\| = g(1) - g(0) \leq |Dg(t_0)| = |\langle\langle y_*, Df(\varphi(t_0))(x_1 - x_2) \rangle\rangle| \leq \|Df(\varphi(t_0))(x_1 - x_2)\|$, where the last inequality follows from Proposition 7.72. The desired result follows.

This completes the proof of the theorem. $\square$

Proposition 9.24 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X} \times \mathcal{Y}$, $f : D \rightarrow \mathcal{Z}$, $(x_0, y_0) \in D$. Assume that the following conditions hold.*

- (i) $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$, $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_0)$, we have $(x, ty + (1-t)y_0) \in D$, $\forall t \in I := [0, 1] \subset \mathbb{R}$.
- (ii) f is partial differentiable with respect to x at (x_0, y_0) and $\frac{\partial f}{\partial x}(x_0, y_0) \in \mathcal{B}(\mathcal{X}, \mathcal{Z})$.
- (iii) $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_0)$, f is partial differentiable with respect to y at (x, y) and $\frac{\partial f}{\partial y}$ is continuous at (x_0, y_0) .

Then, f is Fréchet differentiable at (x_0, y_0) and $Df(x_0, y_0) \in \mathcal{B}(\mathcal{X} \times \mathcal{Y}, \mathcal{Z})$ is given by $Df(x_0, y_0)(h, k) = \frac{\partial f}{\partial x}(x_0, y_0)(h) + \frac{\partial f}{\partial y}(x_0, y_0)(k)$, $\forall (h, k) \in \mathcal{X} \times \mathcal{Y}$.

In “matrix” notation, $Df(x_0, y_0) = \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix}$.

Proof We will first show that $\overline{\operatorname{span}(A_D(x_0, y_0))} = \mathcal{X} \times \mathcal{Y}$. Define $D_{x_0} := \{y \in \mathcal{Y} \mid (x_0, y) \in D\}$ and $D_{y_0} := \{x \in \mathcal{X} \mid (x, y_0) \in D\}$. By the partial differentiability of f with respect to x at (x_0, y_0) , we have $\overline{\operatorname{span}(A_{D_{y_0}}(x_0))} = \mathcal{X}$. By the partial differentiability of f with respect to y at (x_0, y_0) , we have $\overline{\operatorname{span}(A_{D_{x_0}}(y_0))} = \mathcal{Y}$. $\forall u \in A_{D_{y_0}}(x_0)$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \bar{x} := x_0 + r\bar{u} \in D_{y_0}$ with $0 < r < \epsilon$ and $\bar{u} \in \mathcal{B}_{\mathcal{X}}(u, \epsilon)$. Then, $(\bar{x}, y_0) = (x_0, y_0) + r(\bar{u}, \vartheta_{\mathcal{Y}}) \in D$ and $(\bar{u}, \vartheta_{\mathcal{Y}}) \in \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((u, \vartheta_{\mathcal{Y}}), \epsilon)$. Hence, $(u, \vartheta_{\mathcal{Y}}) \in A_D(x_0, y_0)$. Then, $A_{D_{y_0}}(x_0) \times \{\vartheta_{\mathcal{Y}}\} \subseteq A_D(x_0, y_0)$, which implies that $\mathcal{X} \times \{\vartheta_{\mathcal{Y}}\} \subseteq \overline{\operatorname{span}(A_D(x_0, y_0))}$. By symmetry, we have $\{\vartheta_{\mathcal{X}}\} \times \mathcal{Y} \subseteq \overline{\operatorname{span}(A_D(x_0, y_0))}$. By Proposition 7.17, $\overline{\operatorname{span}(A_D(x_0, y_0))} = \mathcal{X} \times \mathcal{Y}$.

Define $L : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$ by $L(h, k) = \frac{\partial f}{\partial x}(x_0, y_0)(h) + \frac{\partial f}{\partial y}(x_0, y_0)(k)$, $\forall (h, k) \in \mathcal{X} \times \mathcal{Y}$. Clearly, L is a linear operator. Note that

$$\begin{aligned} \|L\| &= \sup_{(h,k) \in \mathcal{X} \times \mathcal{Y}, \|(h,k)\| \leq 1} \|L(h, k)\| \\ &\leq \sup_{(h,k) \in \mathcal{X} \times \mathcal{Y}, \|(h,k)\| \leq 1} \left(\left\| \frac{\partial f}{\partial x}(x_0, y_0) \right\| \|h\| + \left\| \frac{\partial f}{\partial y}(x_0, y_0) \right\| \|k\| \right) \\ &\leq \sup_{(h,k) \in \mathcal{X} \times \mathcal{Y}, \|(h,k)\| \leq 1} \sqrt{\left\| \frac{\partial f}{\partial x}(x_0, y_0) \right\|^2 + \left\| \frac{\partial f}{\partial y}(x_0, y_0) \right\|^2} \\ &\quad \cdot \sqrt{\|h\|^2 + \|k\|^2} \leq \sqrt{\left\| \frac{\partial f}{\partial x}(x_0, y_0) \right\|^2 + \left\| \frac{\partial f}{\partial y}(x_0, y_0) \right\|^2} < +\infty \end{aligned}$$

where the first inequality follows from Proposition 7.64 and the second inequality follows from Cauchy-Schwarz Inequality. Hence, $L \in \mathcal{B}(\mathcal{X} \times \mathcal{Y}, \mathcal{Z})$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the partial differentiability of f with respect to x at (x_0, y_0) , $\exists \delta_1 \in (0, \delta_0] \subset \mathbb{R}$ such that $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_1)$, we have $\left\| f(x, y) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) \right\| \leq \epsilon/\sqrt{2} \|x - x_0\|$. By the continuity of $\frac{\partial f}{\partial y}$ at (x_0, y_0) , $\exists \delta_2 \in (0, \delta_1] \subset \mathbb{R}$, such that $\forall (\bar{x}, \bar{y}) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_2)$, we have $\left\| \frac{\partial f}{\partial y}(\bar{x}, \bar{y}) - \frac{\partial f}{\partial y}(x_0, y_0) \right\| < \epsilon/\sqrt{2}$.

Claim 9.24.1 $\left\| f(x, y) - f(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \right\| \leq \epsilon/\sqrt{2} \|y - y_0\|$, $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_2)$.

Proof of claim: Fix any $(x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_2)$. Let $\bar{D} = I$ if $\mathbb{K} = \mathbb{R}$; or $\bar{D} = \{a + i0 \mid a \in I\} \subset \mathbb{C}$ if $\mathbb{K} = \mathbb{C}$. Define $\psi : \bar{D} \rightarrow \mathcal{Z}$ by $\psi(t) = f(x, ty + (1 - t)y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(t(y - y_0))$, $\forall t \in \bar{D}$. By Chain Rule, each term in the definition of ψ is Fréchet differentiable. By Proposition 9.15, ψ is Fréchet differentiable. By Mean Value Theorem, $\exists t_0 \in I^\circ$,

$$\begin{aligned} \left\| f(x, y) - f(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \right\| &= \|\psi(1) - \psi(0)\| \leq \|D\psi(t_0)\| \\ &= \left\| \frac{\partial f}{\partial y}(x, t_0 y + (1 - t_0)y_0)(y - y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \right\| \\ &\leq \left\| \frac{\partial f}{\partial y}(x, t_0 y + (1 - t_0)y_0) - \frac{\partial f}{\partial y}(x_0, y_0) \right\| \|y - y_0\| \leq \epsilon/\sqrt{2} \|y - y_0\| \end{aligned}$$

where the second inequality follows from Proposition 7.64. This completes the proof of the claim. $\square$

Therefore, $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_2)$, we have

$$\|f(x, y) - f(x_0, y_0) - L(x - x_0, y - y_0)\|$$

$$\begin{aligned}
&\leq \left\| f(x, y) - f(x, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \right\| \\
&\quad + \left\| f(x, y_0) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) \right\| \\
&\leq \epsilon/\sqrt{2} \|y - y_0\| + \epsilon/\sqrt{2} \|x - x_0\| \leq \epsilon \|(x - x_0, y - y_0)\|
\end{aligned}$$

where the last inequality follows from Cauchy-Schwarz Inequality. Hence, $Df(x_0, y_0) = L$. This completes the proof of the proposition. $\square$

We observe that Conditions (i) of Proposition 9.24 is easily satisfied when $(x_0, y_0) \in D^\circ$.

9.4 Higher Order Derivatives

9.4.1 Basic concept

We introduce the following notation. Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$. Denote $B(\mathcal{X}, \mathcal{Y})$ by $B_1(\mathcal{X}, \mathcal{Y})$. Recursively, denote $B(\mathcal{X}, B_k(\mathcal{X}, \mathcal{Y}))$ by $B_{k+1}(\mathcal{X}, \mathcal{Y})$, $\forall k \in \mathbb{N}$. Note that $B_k(\mathcal{X}, \mathcal{Y})$ is the set of bounded multilinear $\mathcal{Y}$ -valued functions on $\mathcal{X}^k$, $\forall k \in \mathbb{N}$. Define the subset of symmetric functions by $B_{S_k}(\mathcal{X}, \mathcal{Y}) := \{L \in B_k(\mathcal{X}, \mathcal{Y}) \mid L(h_k) \cdots (h_1) = L(v_k) \cdots (v_1), \forall (h_1, \dots, h_k) \in \mathcal{X}^k, \forall (v_1, \dots, v_k) = \text{a permutation of } (h_1, \dots, h_k)\}$. Note that $B_{S_k}(\mathcal{X}, \mathcal{Y})$ is a closed subspace of $B_k(\mathcal{X}, \mathcal{Y})$. Then, by Proposition 7.13, $B_{S_k}(\mathcal{X}, \mathcal{Y})$ is a normed linear space over $\mathbb{K}$. If $\mathcal{Y}$ is a Banach space, then, by Proposition 7.66, $B_k(\mathcal{X}, \mathcal{Y})$ is a Banach space. Then, by Proposition 4.39, $B_{S_k}(\mathcal{X}, \mathcal{Y})$ is a Banach space. For notational consistency, we will denote $B_{S_0}(\mathcal{X}, \mathcal{Y}) := B_0(\mathcal{X}, \mathcal{Y}) := \mathcal{Y}$.

Definition 9.25 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, and $x_0 \in D$. Let $f^{(1)}$ be defined with domain of definition $\text{dom}(f^{(1)})$. We may consider the derivative of $f^{(1)}$. If $f^{(1)}$ is differentiable at $x_0 \in \text{dom}(f^{(1)})$, then f is said to be twice Fréchet differentiable at x_0 . The second order derivative of f at x_0 is $D(Df)(x_0) =: D^2f(x_0) =: f^{(2)}(x_0) \in B(\mathcal{X}, B(\mathcal{X}, \mathcal{Y})) = B_2(\mathcal{X}, \mathcal{Y})$. D^2f or $f^{(2)}$ will denote the $B_2(\mathcal{X}, \mathcal{Y})$ -valued function whose domain of definition is $\text{dom}(f^{(2)}) := \{x \in \text{dom}(f^{(1)}) \mid f^{(2)}(x) \in B_2(\mathcal{X}, \mathcal{Y}) \text{ exists}\}$. Recursively, if $f^{(k)}$ is Fréchet differentiable at $x_0 \in \text{dom}(f^{(k)})$, then f is said to be $(k+1)$ -times Fréchet differentiable at x_0 and the $(k+1)$ st order derivative of f at x_0 is $Df^{(k)}(x_0) =: D^{k+1}f(x_0) =: f^{(k+1)}(x_0) \in B_{k+1}(\mathcal{X}, \mathcal{Y})$, where $k \in \mathbb{N}$. $D^{k+1}f$ or $f^{(k+1)}$ will denote the $B_{k+1}(\mathcal{X}, \mathcal{Y})$ -valued function whose domain of definition is $\text{dom}(f^{(k+1)}) := \{x \in \text{dom}(f^{(k)}) \mid f^{(k+1)}(x) \in B_{k+1}(\mathcal{X}, \mathcal{Y}) \text{ exists}\}$. For notational consistency, we will let $f^{(0)} = f$.

Note that $\text{dom}(f^{(k+1)}) \subseteq \text{dom}(f^{(k)}) \subseteq D$, $\forall k \in \mathbb{N}$.

Definition 9.26 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, and $x_0 \in D$. Assume that $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that f is k -times Fréchet differentiable at x , $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, where $k \in \mathbb{N}$, that is $D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) \subseteq \text{dom}(f^{(k)})$, and $f^{(k)}$ is continuous at x_0 . Then, we say that f is $\mathcal{C}_k$ at x_0 . If f is $\mathcal{C}_k$ at x , $\forall x \in D$, then, we say f is $\mathcal{C}_k$. If f is $\mathcal{C}_k$ at $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, $\forall k \in \mathbb{N}$, then, we say that f is $\mathcal{C}_{\infty}$ at x_0 . If f is $\mathcal{C}_{\infty}$ at x , $\forall x \in D$, then, we say that f is $\mathcal{C}_{\infty}$.

Note that f being $(k+1)$ -times differentiable at $x_0 \in D$ does not imply that f is $\mathcal{C}_k$ at x_0 since $\text{dom}(f^{(k)})$ may not contain $D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\forall \delta \in (0, \infty) \subset \mathbb{R}$. When $\text{dom}(f^{(k)}) \supseteq D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, for some $\delta_0 \in (0, \infty) \subset \mathbb{R}$, and f is $(k+1)$ -times differentiable at x_0 , then f is $\mathcal{C}_k$ at x_0 . In particular, if f is $\mathcal{C}_{k+1}$ at x_0 , then f is $\mathcal{C}_k$ at x , $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, for some $\delta_0 \in (0, \infty) \subset \mathbb{R}$. If f is infinitely many times differentiable at x , $\forall x \in D$, then f is $\mathcal{C}_{\infty}$.

Proposition 9.27 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$ be $\mathcal{C}_{n+m-1}$ at x_0 and $(n+m)$ -times Fréchet differentiable at $x_0 \in D$, where $n, m \in \mathbb{N}$. Fix $(h_1, \dots, h_n) \in \mathcal{X}^n$. Define the function $g : \text{dom}(f^{(n)}) \rightarrow \mathcal{Y}$ by $g(x) = f^{(n)}(x)(h_n) \cdots (h_1)$, $\forall x \in D_n := \text{dom}(f^{(n)})$. Then, the following statements hold.

(i) g is m -times Fréchet differentiable at x_0 and $g^{(m)}(x_0) \in \mathcal{B}_m(\mathcal{X}, \mathcal{Y})$ is given by $g^{(m)}(x_0)(h_{n+m}) \cdots (h_{n+1}) = f^{(n+m)}(x_0)(h_{n+m}) \cdots (h_1)$, $\forall (h_{n+1}, \dots, h_{n+m}) \in \mathcal{X}^m$.

(ii) If f is $\mathcal{C}_{n+m}$ at x_0 , then g is $\mathcal{C}_m$ at x_0 .

Proof We will first prove (i) using mathematical induction on m .
 1° $m = 1$. Since f is $\mathcal{C}_n$ at x_0 , then $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) \subseteq D_n$. By the $(n+1)$ -times differentiability of f at x_0 , we have $\text{span}(\mathcal{A}_{D_n}(x_0)) = \mathcal{X}$. Define $L : \mathcal{X} \rightarrow \mathcal{Y}$ by, $\forall h \in \mathcal{X}$, $L(h) = f^{(n+1)}(x_0)(h)(h_n) \cdots (h_1)$. Clearly, L is a linear operator. Note that

$$\begin{aligned} \|L\| &= \sup_{h \in \mathcal{X}, \|h\| \leq 1} \|L(h)\| = \sup_{h \in \mathcal{X}, \|h\| \leq 1} \|f^{(n+1)}(x_0)(h)(h_n) \cdots (h_1)\| \\ &\leq \sup_{h \in \mathcal{X}, \|h\| \leq 1} \|f^{(n+1)}(x_0)\| \|h\| \|h_n\| \cdots \|h_1\| \\ &\leq \|f^{(n+1)}(x_0)\| \|h_n\| \cdots \|h_1\| < +\infty \end{aligned}$$

where the first inequality follows from Proposition 7.64. Hence, $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the differentiability of $f^{(n)}$ at x_0 , $\exists \delta \in (0, \delta_0) \subset \mathbb{R}$ such that $\forall x \in D_n \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, we have $\|f^{(n)}(x) - f^{(n)}(x_0) - f^{(n+1)}(x_0)(x - x_0)\| \leq \epsilon/(1 + \|h_n\| \cdots \|h_1\|) \|x - x_0\|$. Then, we have

$$\|g(x) - g(x_0) - L(x - x_0)\|$$

$$\begin{aligned}
&= \|f^{(n)}(x)(h_n) \cdots (h_1) - f^{(n)}(x_0)(h_n) \cdots (h_1) \\
&\quad - f^{(n+1)}(x_0)(x - x_0)(h_n) \cdots (h_1)\| \\
&= \|(f^{(n)}(x) - f^{(n)}(x_0) - f^{(n+1)}(x_0)(x - x_0))(h_n) \cdots (h_1)\| \\
&\leq \|f^{(n)}(x) - f^{(n)}(x_0) - f^{(n+1)}(x_0)(x - x_0)\| \|h_n\| \cdots \|h_1\| \\
&\leq \epsilon \|x - x_0\|
\end{aligned}$$

where the first inequality follows from Proposition 7.64. Hence, $Dg(x_0) = L$ and g is Fréchet differentiable at x_0 .

2° Assume that (i) holds for $m \leq k$, $\forall k \in \mathbb{N}$.

3° Consider the case $m = k + 1$. Since f is $\mathcal{C}_{n+k}$ at x_0 , then, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that f is $(n+k)$ -times Fréchet differentiable at $\bar{x}$ and is $\mathcal{C}_{n+k-1}$ at $\bar{x}$, $\forall \bar{x} \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$. By inductive assumption, g is k -times Fréchet differentiable at $\bar{x}$ and $g^{(k)}(\bar{x})(h_{n+k}) \cdots (h_{n+1}) = f^{(n+k)}(\bar{x})(h_{n+k}) \cdots (h_1)$, $\forall (h_{n+1}, \dots, h_{n+k}) \in \mathcal{X}^k$. Hence, $\bar{x} \in \bar{D}_{n+k} := \text{dom}(g^{(k)})$. Then, we have $D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) \subseteq \bar{D}_{n+k} \subseteq D_n \subseteq D$. Note that $\bar{D}_{n+k} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) = D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$. Then, $\overline{\text{span}(A_{\bar{D}_{n+k}}(x_0))} = \overline{\text{span}(A_D(x_0))} = \mathcal{X}$, since f is differentiable at x_0 .

Define $L : \mathcal{X} \rightarrow \mathcal{B}_k(\mathcal{X}, \mathcal{Y})$ by, $\forall h \in \mathcal{X}$, $\forall (h_{n+1}, \dots, h_{n+k}) \in \mathcal{X}^k$, $L(h)(h_{n+k}) \cdots (h_{n+1}) = f^{(n+k+1)}(x_0)(h)(h_{n+k}) \cdots (h_1)$. Clearly, L is a linear operator. Note that

$$\begin{aligned}
\|L\| &= \sup_{h \in \mathcal{X}, \|h\| \leq 1} \|L(h)\| \\
&= \sup_{h \in \mathcal{X}, \|h\| \leq 1, h_{n+i} \in \mathcal{X}, \|h_{n+i}\| \leq 1, i=1, \dots, k} \|L(h)(h_{n+k}) \cdots (h_{n+1})\| \\
&= \sup_{h \in \mathcal{X}, \|h\| \leq 1, h_{n+i} \in \mathcal{X}, \|h_{n+i}\| \leq 1, i=1, \dots, k} \|f^{(n+k+1)}(x_0)(h)(h_{n+k}) \cdots (h_1)\| \\
&\leq \|f^{(n+k+1)}(x_0)\| \|h_n\| \cdots \|h_1\| < +\infty
\end{aligned}$$

where the first inequality follows from Proposition 7.64. Then, $L \in \mathcal{B}_{k+1}(\mathcal{X}, \mathcal{Y})$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the differentiability of $f^{(n+k)}$ at x_0 , $\exists \delta \in (0, \delta_0] \subset \mathbb{R}$ such that $\forall x \in \text{dom}(f^{(n+k)}) \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) = D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, we have $\|f^{(n+k)}(x) - f^{(n+k)}(x_0) - f^{(n+k+1)}(x_0)(x - x_0)\| \leq \epsilon/(1 + \|h_n\| \cdots \|h_1\|) \|x - x_0\|$. Then, $\forall x \in \bar{D}_{n+k} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) = D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$,

$$\begin{aligned}
&\|g^{(k)}(x) - g^{(k)}(x_0) - L(x - x_0)\| \\
&= \sup_{h_{n+i} \in \mathcal{X}, \|h_{n+i}\| \leq 1, i=1, \dots, k} \|(g^{(k)}(x) - g^{(k)}(x_0) \\
&\quad - L(x - x_0))(h_{n+k}) \cdots (h_{n+1})\| \\
&= \sup_{h_{n+i} \in \mathcal{X}, \|h_{n+i}\| \leq 1, i=1, \dots, k} \|f^{(n+k)}(x)(h_{n+k}) \cdots (h_1) \\
&\quad - f^{(n+k)}(x_0)(h_{n+k}) \cdots (h_1) - f^{(n+k+1)}(x_0)(x - x_0)(h_{n+k}) \cdots (h_1)\|
\end{aligned}$$

$$\begin{aligned}
&\leq \|f^{(n+k)}(x) - f^{(n+k)}(x_0) - f^{(n+k+1)}(x_0)(x - x_0)\| \|h_n\| \cdots \|h_1\| \\
&\leq \epsilon \|x - x_0\|
\end{aligned}$$

where the first inequality follows from Proposition 7.64. Hence, $g^{(k+1)}(x_0) = Dg^{(k)}(x_0) = L$. Therefore, g is $(k+1)$ -times differentiable at x_0 .

This completes the induction process and the proof of (i).

For (ii), let f be $\mathcal{C}_{n+m}$ at x_0 . By Definition 9.26, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) \subseteq D_{n+m} := \text{dom}(f^{(n+m)}) \subseteq D$ and $f^{(n+m)}$ is continuous at x_0 . $\forall x \in D_{n+m} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) = D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, f is $\mathcal{C}_{n+m-1}$ at x and $(n+m)$ -times differentiable at x . By (i), g is m -times Fréchet differentiable at x and $\forall (h_{n+1}, \dots, h_{n+m}) \in \mathcal{X}^m$, we have $g^{(m)}(x)(h_{n+m}) \cdots (h_{n+1}) = f^{(n+m)}(x)(h_{n+m}) \cdots (h_1)$. Then, $D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) \subseteq \text{dom}(g^{(m)}) \subseteq D_n \subseteq D$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the continuity of $f^{(n+m)}$ at x_0 , $\exists \delta \in (0, \delta_0] \subset \mathbb{R}$ such that $\|f^{(n+m)}(x) - f^{(n+m)}(x_0)\| < \epsilon/(1 + \|h_n\| \cdots \|h_1\|)$, $\forall x \in D_{n+m} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) = D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$. $\forall \bar{x} \in \text{dom}(g^{(m)}) \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) = D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, we have

$$\begin{aligned}
&\|g^{(m)}(x) - g^{(m)}(x_0)\| \\
&= \sup_{h_{n+i} \in \mathcal{X}, \|h_{n+i}\| \leq 1, i=1, \dots, m} \|(g^{(m)}(x) - g^{(m)}(x_0))(h_{n+m}) \cdots (h_{n+1})\| \\
&= \sup_{h_{n+i} \in \mathcal{X}, \|h_{n+i}\| \leq 1, i=1, \dots, m} \|f^{(n+m)}(x)(h_{n+m}) \cdots (h_1) \\
&\quad - f^{(n+m)}(x_0)(h_{n+m}) \cdots (h_1)\| \\
&\leq \|f^{(n+m)}(x) - f^{(n+m)}(x_0)\| \|h_n\| \cdots \|h_1\| < \epsilon
\end{aligned}$$

where the first inequality follows from Proposition 7.64. Hence, $g^{(m)}$ is continuous at x_0 . Therefore, g is $\mathcal{C}_m$ at x_0 .

This completes the proof of the proposition. $\square$

Proposition 9.28 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $x_0 \in D$, and $f : D \rightarrow \mathcal{Y}$ be $\mathcal{C}_n$ at x_0 , where $n \in \mathbb{N}$. Assume that D is locally convex at x_0 . Then, $f^{(n)}(x_0) \in \mathcal{B}_{S_n}(\mathcal{X}, \mathcal{Y})$.*

Proof By the assumption, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$ is convex. Without loss of generality, assume f is n -times differentiable at x , $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$. We will prove the proposition by mathematical induction on n .

1° $n = 1$. Clearly $f^{(1)}(x_0) \in \mathcal{B}(\mathcal{X}, \mathcal{Y}) = \mathcal{B}_1(\mathcal{X}, \mathcal{Y}) = \mathcal{B}_{S_1}(\mathcal{X}, \mathcal{Y})$.

Next, we consider $n = 2$. We will prove this case using an argument of contradiction. Suppose $f^{(2)}(x_0)$ is not symmetric. Then, $\exists \tilde{h}_0, \tilde{l}_0 \in \mathcal{X}$ such that $f^{(2)}(x_0)(\tilde{h}_0)(\tilde{l}_0) \neq f^{(2)}(x_0)(\tilde{l}_0)(\tilde{h}_0)$. By the differentiability of f at x_0 , we have $\text{span}(A_D(x_0)) = \mathcal{X}$. Then, by $f^{(2)}(x_0) \in \mathcal{B}_2(\mathcal{X}, \mathcal{Y})$ and Proposition 3.56, $\exists \hat{h}_0, \hat{l}_0 \in \text{span}(A_D(x_0))$ such that $f^{(2)}(x_0)(\hat{h}_0)(\hat{l}_0) \neq f^{(2)}(x_0)(\hat{l}_0)(\hat{h}_0)$. Since $f^{(2)}(x_0)$ is multi-linear, then $\exists \tilde{h}_0, \tilde{l}_0 \in A_D(x_0)$ such that $f^{(2)}(x_0)(\tilde{h}_0)(\tilde{l}_0) \neq f^{(2)}(x_0)(\hat{l}_0)(\hat{h}_0)$. By continuity of $f^{(2)}(x_0)$,

$\exists \epsilon_1 \in (0, \infty) \subset \mathbb{R}$ such that $\forall \check{h} \in \mathcal{B}_X(\hat{h}_0, \epsilon_1)$, $\forall \check{l} \in \mathcal{B}_X(\hat{l}_0, \epsilon_1)$, we have $f^{(2)}(x_0)(\check{h})(\check{l}) \neq f^{(2)}(x_0)(\check{l})(\check{h})$. By $\hat{h}_0, \hat{l}_0 \in A_D(x_0)$, $\exists r_h, r_l \in (0, \epsilon_1) \subset \mathbb{R}$, $\exists \check{h}_0 \in \mathcal{B}_X(\hat{h}_0, \epsilon_1)$, $\exists \check{l}_0 \in \mathcal{B}_X(\hat{l}_0, \epsilon_1)$, such that $x_0 + r_h \check{h}_0, x_0 + r_l \check{l}_0 \in D \cap \mathcal{B}_X(x_0, \delta_0)$. Clearly, we have $f^{(2)}(x_0)(r_h \check{h}_0)(r_l \check{l}_0) = r_h r_l f^{(2)}(x_0)(\check{h}_0)(\check{l}_0) \neq r_h r_l f^{(2)}(x_0)(\check{l}_0)(\check{h}_0) = f^{(2)}(x_0)(r_l \check{l}_0)(r_h \check{h}_0)$. Let $h_0 := r_h/2 \check{h}_0$ and $l_0 := r_l/2 \check{l}_0$. Then, by the convexity of the set $D \cap \mathcal{B}_X(x_0, \delta_0)$, we have $x_0, x_0 + h_0, x_0 + l_0, x_0 + h_0 + l_0 \in D \cap \mathcal{B}_X(x_0, \delta_0)$ and $f^{(2)}(x_0)(h_0)(l_0) \neq f^{(2)}(x_0)(l_0)(h_0)$. Clearly, $h_0 \neq \vartheta_X$ and $l_0 \neq \vartheta_X$. Let $\epsilon_0 := \|f^{(2)}(x_0)(h_0)(l_0) - f^{(2)}(x_0)(l_0)(h_0)\| / (\|h_0\| \|l_0\|) \in (0, \infty) \subset \mathbb{R}$. Since f is $\mathcal{C}_2$ at x_0 , then $\exists \delta_1 \in (0, \delta_0] \subset \mathbb{R}$ such that $\forall x \in D \cap \mathcal{B}_X(x_0, \delta_1)$, we have $\|f^{(2)}(x) - f^{(2)}(x_0)\| < \epsilon_0/2$. By proper scaling of h_0 and l_0 , we may assume that $\forall t_1, t_2 \in I := [0, 1] \subset \mathbb{R}$, $x_0 + t_1 h_0 + t_2 l_0 \in D \cap \mathcal{B}_X(x_0, \delta_1)$.

In summary, $\exists h_0, l_0 \in \mathcal{X} \setminus \{\vartheta_X\}$ such that $\epsilon_0 := \|f^{(2)}(x_0)(h_0)(l_0) - f^{(2)}(x_0)(l_0)(h_0)\| / (\|h_0\| \|l_0\|) \in (0, \infty) \subset \mathbb{R}$, $\forall t_1, t_2 \in I$, f is twice differentiable at $x_0 + t_1 h_0 + t_2 l_0 \in D \cap \mathcal{B}_X(x_0, \delta_0)$ and $\|f^{(2)}(x_0 + t_1 h_0 + t_2 l_0) - f^{(2)}(x_0)\| < \epsilon_0/2$.

Let $\bar{D} := I$, if $\mathbb{K} = \mathbb{R}$; or $\bar{D} := \{a + i0 \mid a \in I\} \subset \mathbb{C}$, if $\mathbb{K} = \mathbb{C}$. $\forall t_1 \in I$, define $\psi_{t_1} : \bar{D} \rightarrow \mathcal{B}(\mathcal{X}, \mathcal{Y})$ by $\psi_{t_1}(t_2) = f^{(1)}(x_0 + t_1 h_0 + t_2 l_0) - f^{(1)}(x_0 + t_1 h_0) - f^{(2)}(x_0)(t_2 l_0)$, $\forall t_2 \in \bar{D}$. By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of ψ_{t_1} is Fréchet differentiable. Then, by Proposition 9.15, ψ_{t_1} is Fréchet differentiable at t_2 , $\forall t_2 \in \bar{D}$. By Mean Value Theorem, $\exists \bar{t}_2 \in I^\circ$ such that

$$\begin{aligned} & \|f^{(1)}(x_0 + t_1 h_0 + l_0) - f^{(1)}(x_0 + t_1 h_0) - f^{(2)}(x_0)(l_0)\| \\ &= \|\psi_{t_1}(1) - \psi_{t_1}(0)\| \leq \|D\psi_{t_1}(\bar{t}_2)\| \\ &= \|f^{(2)}(x_0 + t_1 h_0 + \bar{t}_2 l_0)(l_0) - f^{(2)}(x_0)(l_0)\| \\ &\leq \|f^{(2)}(x_0 + t_1 h_0 + \bar{t}_2 l_0) - f^{(2)}(x_0)\| \|l_0\| < \epsilon_0 \|l_0\|/2 \end{aligned}$$

where the second inequality follows from Proposition 7.64.

Define $\gamma : \bar{D} \rightarrow \mathcal{Y}$ by $\gamma(t_1) = f(x_0 + l_0 + t_1 h_0) - f(x_0 + t_1 h_0) - f^{(2)}(x_0)(l_0)(t_1 h_0)$, $\forall t_1 \in \bar{D}$. By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of γ is Fréchet differentiable. Then, by Proposition 9.15, γ is Fréchet differentiable at t_1 , $\forall t_1 \in \bar{D}$. By Mean Value Theorem, $\exists \bar{t}_1 \in I^\circ$ such that

$$\begin{aligned} & \|f(x_0 + l_0 + h_0) - f(x_0 + h_0) - f(x_0 + l_0) + f(x_0) - f^{(2)}(x_0)(l_0)(h_0)\| \\ &= \|\gamma(1) - \gamma(0)\| \leq \|D\gamma(\bar{t}_1)\| = \|f^{(1)}(x_0 + l_0 + \bar{t}_1 h_0)(h_0) \\ &\quad - f^{(1)}(x_0 + \bar{t}_1 h_0)(h_0) - f^{(2)}(x_0)(l_0)(h_0)\| \\ &\leq \|f^{(1)}(x_0 + l_0 + \bar{t}_1 h_0) - f^{(1)}(x_0 + \bar{t}_1 h_0) - f^{(2)}(x_0)(l_0)\| \|h_0\| \\ &= \|\psi_{\bar{t}_1}(1) - \psi_{\bar{t}_1}(0)\| \|h_0\| < \epsilon_0 \|l_0\| \|h_0\|/2 \end{aligned}$$

where the second inequality follows from Proposition 7.64.

By symmetry, we have $\|f(x_0 + h_0 + l_0) - f(x_0 + l_0) - f(x_0 + h_0) + f(x_0) - f^{(2)}(x_0)(h_0)(l_0)\| < \epsilon_0 \|h_0\| \|l_0\|/2$. Then, $\|f^{(2)}(x_0)(h_0)(l_0) - f^{(2)}(x_0)(l_0)(h_0)\| < \epsilon_0 \|h_0\| \|l_0\|$. This leads to the contradiction $\epsilon_0 := \|f^{(2)}(x_0)(h_0)(l_0) - f^{(2)}(x_0)(l_0)(h_0)\| / (\|h_0\| \|l_0\|) < \epsilon_0$.

Hence, $f^{(2)}(x_0)$ must be symmetric and $f^{(2)}(x_0) \in \mathcal{B}_{S^2}(\mathcal{X}, \mathcal{Y})$.

2° Assume that the result holds $\forall n \leq k$, $k \in \{2, 3, \dots\}$.

3° Consider the case $n = k+1$. $\forall (h_1, \dots, h_{k+1}) \in \mathcal{X}^{k+1}$, let $(v_1, \dots, v_{k+1})$ be a permutation of $(h_1, \dots, h_{k+1})$. We need to show that

$$f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_1) = f^{(k+1)}(x_0)(v_{k+1}) \cdots (v_1)$$

Since any permutation can be arrived at in finite number of steps by interchanging two consecutive elements, then, all we need to show is that, $\forall i = 1, \dots, k$,

$$\begin{aligned} & f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_{i+1})(h_i) \cdots (h_1) \\ &= f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_i)(h_{i+1}) \cdots (h_1) \end{aligned}$$

We will distinguish two exhaustive and mutually exclusive cases: Case 1: $1 \leq i < k$; Case 2: $i = k$. Case 1: $1 \leq i < k$. Define $g : \text{dom}(f^{(k)}) \rightarrow \mathcal{Y}$ by $g(x) = f^{(k)}(x)(h_k) \cdots (h_{i+1})(h_i) \cdots (h_1)$, $\forall x \in \text{dom}(f^{(k)})$. $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) = \text{dom}(f^{(k+1)}) \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, f is $(k+1)$ -times differentiable at x , and $\text{dom}(f^{(k)}) \supseteq D \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x)$, where $\delta_x := \delta_0 - \|x - x_0\| \in (0, \infty) \subset \mathbb{R}$. Then, f is $\mathcal{C}_k$ at x . Clearly, $D \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x) = (D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)) \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x)$ is convex. By inductive assumption, $g(x) = f^{(k)}(x)(h_k) \cdots (h_i)(h_{i+1}) \cdots (h_1)$. By Proposition 9.27, g is $\mathcal{C}_1$ at x_0 and $f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_{i+1})(h_i) \cdots (h_1) = g^{(1)}(x_0)(h_{k+1}) = f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_i)(h_{i+1}) \cdots (h_1)$. This case is proved.

Case 2: $i = k$. Define $g : \text{dom}(f^{(k-1)}) \rightarrow \mathcal{Y}$ by, $\forall x \in \text{dom}(f^{(k-1)})$, $g(x) = f^{(k-1)}(x)(h_{k-1}) \cdots (h_1)$. Then, by Proposition 9.27, g is $\mathcal{C}_2$ at x_0 and $g^{(2)}(x_0)(u)(v) = f^{(k+1)}(x_0)(u)(v)(h_{k-1}) \cdots (h_1)$, $\forall u, v \in \mathcal{X}$. Note that $\text{dom}(f^{(k-1)}) \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) = D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$ is convex. By the case of $n = 2$, we have $f^{(k+1)}(x_0)(h_{k+1})(h_k) \cdots (h_1) = g^{(2)}(x_0)(h_{k+1})(h_k) = g^{(2)}(x_0)(h_k)(h_{k+1}) = f^{(k+1)}(x_0)(h_k)(h_{k+1}) \cdots (h_1)$. This case is also proved.

Hence, $f^{(k+1)}(x_0) \in \mathcal{B}_{S^{k+1}}(\mathcal{X}, \mathcal{Y})$.

This completes the induction process and the proof of the proposition. $\square$

Definition 9.29 Let $\mathcal{X}_1, \dots, \mathcal{X}_p$, and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, where $p \in \{2, 3, \dots\}$, $D \subseteq \mathcal{X} := \prod_{i=1}^p \mathcal{X}_i$, $f : D \rightarrow \mathcal{Y}$, and $x_o := (x_{1o}, \dots, x_{po}) \in D$. Assume that $\frac{\partial f}{\partial x_{i_1}} : \text{dom}\left(\frac{\partial f}{\partial x_{i_1}}\right) \rightarrow \mathcal{B}(\mathcal{X}_{i_1}, \mathcal{Y})$ is partial differentiable with respect to x_{i_2} at $x_o \in \text{dom}\left(\frac{\partial f}{\partial x_{i_1}}\right)$, where $i_1, i_2 \in \{1, \dots, p\}$. Then, this partial derivative is denoted by $\frac{\partial^2 f}{\partial x_{i_2} \partial x_{i_1}}(x_o) \in \mathcal{B}(\mathcal{X}_{i_2}, \mathcal{B}(\mathcal{X}_{i_1}, \mathcal{Y}))$, which is one of the second order partial derivatives of

f . We will let $\frac{\partial^2 f}{\partial x_{i_2} \partial x_{i_1}}$ denote the $B(X_{i_2}, B(X_{i_1}, Y))$ -valued function whose domain of definition is $\text{dom}\left(\frac{\partial^2 f}{\partial x_{i_2} \partial x_{i_1}}\right) := \left\{x \in \text{dom}\left(\frac{\partial f}{\partial x_{i_1}}\right) \mid \frac{\partial^2 f}{\partial x_{i_2} \partial x_{i_1}}(x) \in B(X_{i_2}, B(X_{i_1}, Y)) \text{ exists}\right\}$. There are a total of p^2 second order partial derivatives of f . Recursively, we may define k th-order partial derivatives, $k \in \{3, 4, \dots\}$. There are a total of p^k k -order partial derivatives. A typical such derivative is denoted by $\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}$, where $i_1, \dots, i_k \in \{1, \dots, p\}$.

Let X be a normed linear space over $\mathbb{K}$ and $D \subseteq X$. Define $\text{Sm}(D) := \bigcup_{\alpha \in \mathbb{K}} \alpha D$. Clearly, $\text{Sm}(D) \subseteq \text{span}(D)$.

Proposition 9.30 *Let X be a normed linear space over $\mathbb{K}$, $D \subseteq X$, and $x_0 \in \overline{D}$. Then, $\overline{\text{span}(A_D(x_0))} \subseteq \overline{\text{span}((D \cap \mathcal{B}(x_0, \delta)) - x_0)}$ and $A_D(x_0) \subseteq \overline{\text{Sm}((D \cap \mathcal{B}(x_0, \delta)) - x_0)}$, $\forall \delta \in (0, \infty) \subset \mathbb{R}$.*

Proof $\forall u \in A_D(x_0)$, we will show that $u \in \overline{\text{Sm}(D - x_0)}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by $u \in A_D(x_0)$, we have $\exists \bar{x} := x_0 + r\bar{u} \in D$ such that $0 < r < \epsilon$ and $\bar{u} \in \mathcal{B}(u, \epsilon)$. Then, $\bar{u} \in \text{Sm}(D - x_0) \cap \mathcal{B}(u, \epsilon) \neq \emptyset$. Hence, by Proposition 3.3, $u \in \overline{\text{Sm}(D - x_0)}$. By the arbitrariness of u , we have $A_D(x_0) \subseteq \overline{\text{Sm}(D - x_0)} \subseteq \overline{\text{span}(D - x_0)}$. By Proposition 7.17, $\overline{\text{span}(A_D(x_0))} \subseteq \overline{\text{span}(D - x_0)}$. By Definition 3.2, $\overline{\text{span}(A_D(x_0))} \subseteq \overline{\text{span}(D - x_0)}$. $\forall \delta \in (0, \infty) \subset \mathbb{R}$, we have $x_0 \in \overline{D \cap \mathcal{B}(x_0, \delta)}$ and $A_D(x_0) = A_{D \cap \mathcal{B}(x_0, \delta)}(x_0)$. Then, $A_D(x_0) \subseteq \overline{\text{Sm}((D \cap \mathcal{B}(x_0, \delta)) - x_0)}$ and $\overline{\text{span}(A_D(x_0))} \subseteq \overline{\text{span}((D \cap \mathcal{B}(x_0, \delta)) - x_0)}$. This completes the proof of the proposition. $\square$

9.4.2 Interchange order of differentiation

Next, we present two results on the interchangeability of order of differentiation. The first result does not assume the existence of the derivative after the interchange, which then requires a stronger assumption on the set $D = \text{dom}(f)$. The second result assumes the existence of the derivative after the interchange, where the stronger assumption on D can be removed.

Proposition 9.31 *Let X, Y , and Z be normed linear spaces over $\mathbb{K}$, $D \subseteq X \times Y$, $f : D \rightarrow Z$, and $(x_0, y_0) \in D$. Assume that the following conditions are satisfied.*

- (i) D is locally convex at (x_0, y_0) , that is, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap \mathcal{B}_{X \times Y}((x_0, y_0), \delta_0) =: \tilde{D}$ is convex. Furthermore, $\forall (x, y) \in \tilde{D}$, we have $(x, y_0), (x_0, y) \in \tilde{D}$.
- (ii) $\frac{\partial f}{\partial x}(x, y)$, $\frac{\partial f}{\partial y}(x, y)$, and $\frac{\partial^2 f}{\partial x \partial y}(x, y)$ exist, $\forall (x, y) \in \tilde{D}$, and $\frac{\partial^2 f}{\partial x \partial y}$ is continuous at (x_0, y_0) .

Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \delta_0] \subset \mathbb{R}$ such that $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta) =: \tilde{D}_\delta$, we have $\left\| f(x, y) - f(x, y_0) - f(x_0, y) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right\| \leq \epsilon \|x - x_0\| \|y - y_0\|$.

Furthermore, if, in addition, the following condition is satisfied.

(iii) $\exists M \in [0, \infty) \subset \mathbb{R}$, $\forall h \in \mathcal{X}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall (x, y) \in \tilde{D}$, $\forall \bar{\delta} \in (0, \infty) \subset \mathbb{R}$, $\exists h_1, h_2 \in \text{Sm}((D_y \cap \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta})) - x_0)$ such that $\|h - h_1 + h_2\| \leq \epsilon \|h\|$, $\|h_1\| \leq M \|h\|$, and $\|h_2\| \leq M \|h\|$, where $D_y := \{x \in \mathcal{X} \mid (x, y) \in D\}$, $\forall y \in \mathcal{Y}$.

Then, $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$ exists and is given by

$$\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(k)(h) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h)(k); \quad \forall h \in \mathcal{X}, \forall k \in \mathcal{Y}$$

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the continuity of $\frac{\partial^2 f}{\partial x \partial y}$ at (x_0, y_0) , $\exists \delta \in (0, \delta_0] \subset \mathbb{R}$ such that $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta) =: \tilde{D}_\delta \subseteq \tilde{D}$, we have $\left\| \frac{\partial^2 f}{\partial x \partial y}(x, y) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \right\| < \epsilon/3$. Note that, by (i), $\tilde{D}_\delta$ is convex and, $\forall (x, y) \in \tilde{D}_\delta$, we have $(x_0, y), (x, y_0) \in \tilde{D}_\delta$.

Claim 9.31.1 $\forall (x, y) \in \tilde{D}_\delta$, $\left\| f(x, y) - f(x, y_0) - f(x_0, y) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right\| \leq \epsilon/3 \|x - x_0\| \|y - y_0\|$.

Proof of claim: Fix $(x, y) \in \tilde{D}_\delta$. Let $I := [0, 1] \subset \mathbb{R}$. Define $\bar{D} := I$ if $\mathbb{K} = \mathbb{R}$; or $\bar{D} := \{a + i0 \mid a \in I\} \subseteq \mathbb{C}$ if $\mathbb{K} = \mathbb{C}$.

$\forall t_2 \in I$, define $\psi_{t_2} : \bar{D} \rightarrow \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ by $\psi_{t_2}(t_1) = \frac{\partial f}{\partial y}(t_1 x + (1 - t_1)x_0, t_2 y + (1 - t_2)y_0) - \frac{\partial f}{\partial y}(x_0, t_2 y + (1 - t_2)y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(t_1(x - x_0))$, $\forall t_1 \in \bar{D}$. By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of ψ_{t_2} is differentiable. By Proposition 9.15, ψ_{t_2} is differentiable. By Mean Value Theorem, $\exists \bar{t}_1 \in I^\circ$ such that

$$\begin{aligned} & \left\| \frac{\partial f}{\partial y}(x, t_2 y + (1 - t_2)y_0) - \frac{\partial f}{\partial y}(x_0, t_2 y + (1 - t_2)y_0) \right. \\ & \quad \left. - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0) \right\| \\ &= \left\| \psi_{t_2}(1) - \psi_{t_2}(0) \right\| \leq \|D\psi_{t_2}(\bar{t}_1)\| \\ &= \left\| \frac{\partial^2 f}{\partial x \partial y}(\bar{t}_1 x + (1 - \bar{t}_1)x_0, t_2 y + (1 - t_2)y_0)(x - x_0) \right. \\ & \quad \left. - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0) \right\| \leq \epsilon/3 \|x - x_0\| \end{aligned}$$

where the last inequality follows from Proposition 7.64.

Define $\gamma : \bar{D} \rightarrow \mathcal{Z}$ by $\gamma(t_2) = f(x, t_2 y + (1 - t_2)y_0) - f(x_0, t_2 y + (1 - t_2)y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(t_2(y - y_0))$, $\forall t_2 \in \bar{D}$. By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of γ is differentiable. By Proposition 9.15, γ is differentiable. By Mean Value Theorem, $\exists \bar{t}_2 \in I^\circ$ such that

$$\begin{aligned}
& \left\| f(x, y) - f(x, y_0) - f(x_0, y) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right\| \\
&= \left\| \gamma(1) - \gamma(0) \right\| \leq \left\| D\gamma(\bar{t}_2) \right\| \\
&= \left\| \frac{\partial f}{\partial y}(x, \bar{t}_2 y + (1 - \bar{t}_2)y_0)(y - y_0) - \frac{\partial f}{\partial y}(x_0, \bar{t}_2 y + (1 - \bar{t}_2)y_0)(y - y_0) \right. \\
&\quad \left. - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right\| \\
&\leq \left\| \frac{\partial f}{\partial y}(x, \bar{t}_2 y + (1 - \bar{t}_2)y_0) - \frac{\partial f}{\partial y}(x_0, \bar{t}_2 y + (1 - \bar{t}_2)y_0) \right. \\
&\quad \left. - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0) \right\| \|y - y_0\| \\
&= \left\| \psi_{\bar{t}_2}(1) - \psi_{\bar{t}_2}(0) \right\| \|y - y_0\| \leq \epsilon/3 \|x - x_0\| \|y - y_0\|
\end{aligned}$$

where the second inequality follows from Proposition 7.64. This completes the proof of the claim. $\square$

Define $D_{x_0} := \{y \in \mathcal{Y} \mid (x_0, y) \in D\}$; and $\hat{D} := \left\{y \in \mathcal{Y} \mid (x_0, y) \in \overline{\text{dom}\left(\frac{\partial f}{\partial x}\right)}\right\}$. By (ii), we have $\overline{\text{span}(A_{D_{x_0}}(y_0))} = \mathcal{Y}$ and $\overline{\text{span}(A_{D_y}(x))} = \mathcal{X}$, $\forall (x, y) \in \hat{D}$. Clearly, we have $\hat{D} \cap \mathcal{B}_y(y_0, \delta_0) = D_{x_0} \cap \mathcal{B}_y(y_0, \delta_0)$. This implies that $\overline{\text{span}(A_{\hat{D}}(y_0))} = \mathcal{Y}$.

$\forall k \in \mathcal{Y}$, define $L_k : \mathcal{X} \rightarrow \mathcal{Z}$ by $L_k(h) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h)(k)$, $\forall h \in \mathcal{X}$. Clearly, L_k is a linear operator. Note that $\|L_k\| = \sup_{h \in \mathcal{X}, \|h\| \leq 1} \|L_k(h)\| \leq \left\| \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \right\| \|k\| < +\infty$, where the first inequality follows from Proposition 7.64. Hence, $L_k \in \mathcal{B}(\mathcal{X}, \mathcal{Z})$.

Define $L : \mathcal{Y} \rightarrow \mathcal{B}(\mathcal{X}, \mathcal{Z})$ by $L(k) = L_k$, $\forall k \in \mathcal{Y}$. Clearly, L is a linear operator. Note that $\|L\| = \sup_{k \in \mathcal{Y}, \|k\| \leq 1} \|L(k)\| \leq \left\| \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \right\| < +\infty$. Hence, $L \in \mathcal{B}(\mathcal{Y}, \mathcal{B}(\mathcal{X}, \mathcal{Z}))$.

Now, fix any $y \in \hat{D} \cap \mathcal{B}_y(y_0, \delta/\sqrt{2})$, we will show that $\|\Delta_y\| \leq (2M + 1)\epsilon \|y - y_0\|$, where $\Delta_y := \frac{\partial f}{\partial x}(x_0, y) - \frac{\partial f}{\partial x}(x_0, y_0) - L(y - y_0) \in \mathcal{B}(\mathcal{X}, \mathcal{Z})$. This immediately implies that $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = L$ and completes the proof of the proposition.

We will distinguish two exhaustive and mutually exclusive cases: Case 1: $y = y_0$; Case 2: $y \neq y_0$. Case 1: $y = y_0$. The result is immediate. Case 2: $y \neq y_0$. By the existence of $\frac{\partial f}{\partial x}(x_0, y)$, $\exists \delta_{y1} \in (0, \delta/\sqrt{2}] \subset \mathbb{R}$ such that $\forall x \in D_y \cap \mathcal{B}_x(x_0, \delta_{y1})$, we have $\left\| f(x, y) - f(x_0, y) - \frac{\partial f}{\partial x}(x_0, y)(x - x_0) \right\| \leq$

$\epsilon/3 \|y - y_0\| \|x - x_0\|$. By the existence of $\frac{\partial f}{\partial x}(x_0, y_0)$, $\exists \delta_{y2} \in (0, \delta/\sqrt{2}] \subset \mathbb{R}$ such that $\forall x \in D_{y_0} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_{y2})$, we have $\left\| f(x, y_0) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) \right\| \leq \epsilon/3 \|y - y_0\| \|x - x_0\|$. Let $\delta_y = \min\{\delta_{y1}, \delta_{y2}\} \in (0, \delta/\sqrt{2}] \subset \mathbb{R}$. $\forall x \in D_y \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_y)$, we have $(x, y), (x, y_0), (x_0, y) \in \tilde{D}$. Then, $x \in D_{y_0} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_y)$. By Claim 9.31.1 and the preceding argument, we have

$$\begin{aligned}
& \|\Delta_y(x - x_0)\| \\
&= \left\| \frac{\partial f}{\partial x}(x_0, y)(x - x_0) - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) - L(y - y_0)(x - x_0) \right\| \\
&\leq \left\| \frac{\partial f}{\partial x}(x_0, y)(x - x_0) - f(x, y) + f(x_0, y) \right\| + \left\| f(x, y_0) - f(x_0, y_0) \right. \\
&\quad \left. - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) \right\| + \left\| f(x, y) - f(x, y_0) - f(x_0, y) + f(x_0, y_0) \right. \\
&\quad \left. - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right\| \\
&\leq \epsilon \|x - x_0\| \|y - y_0\|
\end{aligned}$$

$\forall h \in \mathcal{X}$, by (iii) and the continuity of Δ_y , $\exists h_1, h_2 \in \text{Sm}((D_y \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_y)) - x_0)$, such that $\|\Delta_y(h - h_1 + h_2)\| \leq \epsilon \|h\| \|y - y_0\|$, $\|h_1\| \leq M \|h\|$, and $\|h_2\| \leq M \|h\|$. Then, $\exists \alpha_1, \alpha_2 \in \mathbb{K}$ and $\exists x_1, x_2 \in D_y \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_y)$ such that $h_i = \alpha_i(x_i - x_0)$, $i = 1, 2$. Then, we have

$$\begin{aligned}
& \|\Delta_y(h)\| \\
&\leq \|\Delta_y(h_1 - h_2)\| + \epsilon \|h\| \|y - y_0\| \\
&= \|\alpha_1 \Delta_y(x_1 - x_0) - \alpha_2 \Delta_y(x_2 - x_0)\| + \epsilon \|h\| \|y - y_0\| \\
&\leq |\alpha_1| \|\Delta_y(x_1 - x_0)\| + |\alpha_2| \|\Delta_y(x_2 - x_0)\| + \epsilon \|h\| \|y - y_0\| \\
&\leq (|\alpha_1| \|x_1 - x_0\| + |\alpha_2| \|x_2 - x_0\| + \|h\|) \epsilon \|y - y_0\| \\
&= (\|h_1\| + \|h_2\| + \|h\|) \epsilon \|y - y_0\| \leq (2M + 1) \epsilon \|y - y_0\| \|h\|
\end{aligned}$$

Then, we have $\|\Delta_y\| \leq (2M + 1) \epsilon \|y - y_0\|$. This case is proved.

This completes the proof of the proposition. $\square$

In the preceding proposition, conditions (i) and (iii) are assumptions on the set D . It is clear that (i) and (iii) are satisfied if $(x_0, y_0) \in D^\circ$.

Proposition 9.32 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X} \times \mathcal{Y}$, $f: D \rightarrow \mathcal{Z}$, and $(x_0, y_0) \in D$. Assume that the following conditions are satisfied.*

- (i) *D is locally convex at (x_0, y_0) , that is, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_0) =: \tilde{D}$ is convex. Furthermore, $\forall (x, y) \in \tilde{D}$, we have $(x, y_0), (x_0, y) \in \tilde{D}$.*
- (ii) *$\frac{\partial f}{\partial x}(x, y)$, $\frac{\partial f}{\partial y}(x, y)$, $\frac{\partial^2 f}{\partial x \partial y}(x, y)$, and $\frac{\partial^2 f}{\partial y \partial x}(x, y)$ exist, $\forall (x, y) \in \tilde{D}$, and $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ are continuous at (x_0, y_0) .*

Then, $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(k)(h) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h)(k)$, $\forall h \in \mathcal{X}$, $\forall k \in \mathcal{Y}$.

Proof Define $D_{x_0} := \{y \in \mathcal{Y} \mid (x_0, y) \in D\}$ and $D_{y_0} := \{x \in \mathcal{X} \mid (x, y_0) \in D\}$. By (ii), we have $\overline{\text{span}(A_{D_{x_0}}(y_0))} = \mathcal{Y}$ and $\overline{\text{span}(A_{D_{y_0}}(x_0))} = \mathcal{X}$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Proposition 9.31, $\exists \delta_1 \in (0, \delta_0] \subset \mathbb{R}$ such that $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_1) =: \tilde{D}_{\delta_1} \subseteq \tilde{D}$, we have $\left\| f(x, y) - f(x, y_0) - f(x_0, y) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right\| \leq \epsilon/2 \|x - x_0\| \|y - y_0\|$.

By symmetry, $\exists \delta_2 \in (0, \delta_0] \subset \mathbb{R}$ such that $\forall (x, y) \in \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_2) \cap D =: \tilde{D}_{\delta_2} \subseteq \tilde{D}$, we have $\left\| f(x, y) - f(x, y_0) - f(x_0, y) + f(x_0, y_0) - \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(y - y_0)(x - x_0) \right\| \leq \epsilon/2 \|x - x_0\| \|y - y_0\|$.

Therefore, let $\delta := \min\{\delta_1, \delta_2\} \in (0, \delta_0] \subset \mathbb{R}$, $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta) =: \tilde{D}_\delta = \tilde{D}_{\delta_1} \cap \tilde{D}_{\delta_2} \subseteq \tilde{D}$, we have $\left\| \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(y - y_0)(x - x_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right\| \leq \epsilon \|x - x_0\| \|y - y_0\|$. By the arbitrariness of ϵ , we have $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(y - y_0)(x - x_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0)$.

Clearly, by (i), $\tilde{D}_\delta$ is convex.

Suppose that the result of the proposition is not true, then $\exists h_0 \in \mathcal{X}$ and $\exists k_0 \in \mathcal{Y}$ such that $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(k_0)(h_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h_0)(k_0)$. By the fact that $\overline{\text{span}(A_{D_{x_0}}(y_0))} = \mathcal{Y}$ and $\overline{\text{span}(A_{D_{y_0}}(x_0))} = \mathcal{X}$ and the continuity of $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$ and $\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$, $\exists \bar{h}_0 \in \overline{\text{span}(A_{D_{y_0}}(x_0))}$ and $\exists \bar{k}_0 \in \overline{\text{span}(A_{D_{x_0}}(y_0))}$ such that $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(\bar{k}_0)(\bar{h}_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(\bar{h}_0)(\bar{k}_0)$. By the multi-linearity of these two second-order partial derivatives, $\exists \hat{h}_0 \in A_{D_{y_0}}(x_0)$ and $\exists \hat{k}_0 \in A_{D_{x_0}}(y_0)$ such that $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(\hat{k}_0)(\hat{h}_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(\hat{h}_0)(\hat{k}_0)$. By Proposition 9.30, we have

$$\begin{aligned} A_{D_{y_0}}(x_0) &\subseteq \overline{\text{Sm}((D_{y_0} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)) - x_0)} \\ A_{D_{x_0}}(y_0) &\subseteq \overline{\text{Sm}((D_{x_0} \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta)) - y_0)} \end{aligned}$$

This implies that, by the continuity of these two second-order partial derivatives, $\exists \check{h}_0 \in \text{Sm}((D_{y_0} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)) - x_0)$ and $\exists \check{k}_0 \in \text{Sm}((D_{x_0} \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta)) - y_0)$ such that $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(\check{k}_0)(\check{h}_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(\check{h}_0)(\check{k}_0)$. Then, $\exists \bar{x} \in D_{y_0} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\exists \bar{y} \in D_{x_0} \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta)$, and $\exists \alpha_h, \alpha_k \in \mathbb{K}$ such that $\bar{h}_0 = \alpha_h(\bar{x} - x_0)$ and $\bar{k}_0 = \alpha_k(\bar{y} - y_0)$. By the multi-linearity of these two second-order partial derivatives, $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(\bar{y} - y_0)(\bar{x} - x_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(\bar{x} - x_0)(\bar{y} - y_0)$. Clearly, $(\bar{x}, y_0), (x_0, \bar{y}) \in \tilde{D}_\delta$. By the convexity of $\tilde{D}_\delta$, we have $(\hat{x}, \hat{y}) := 0.5(\bar{x}, y_0) + 0.5(x_0, \bar{y}) \in \tilde{D}_\delta$. By the multi-linearity of these two second-order partial derivatives, we have

$\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(\hat{y} - y_0)(\hat{x} - x_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(\hat{x} - x_0)(\hat{y} - y_0)$. This is a contradiction. Hence, the result of the proposition must hold.

This completes the proof of the proposition. $\square$

9.4.3 High order derivatives of some common functions

Proposition 9.33 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, and $x_0 \in D$. Assume that $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ and $\exists y_0 \in \mathcal{Y}$ such that $f(x) = y_0$ and $\text{span}(\mathcal{A}_D(x)) = \mathcal{X}$, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$. Then, f is $\mathcal{C}_\infty$ at x_0 and $f^{(i)}(x) = \vartheta_{\mathcal{B}_{\mathcal{S}_i(\mathcal{X}, \mathcal{Y})}}(x)$, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, $\forall i \in \mathbb{N}$.*

Proof $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, let $\delta_x := \delta_0 - \|x - x_0\| > 0$. Then, $\forall \bar{x} \in \overline{D \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x)} \subseteq D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, we have $f(\bar{x}) = y_0$. Note that $\text{span}(\mathcal{A}_D(x)) = \mathcal{X}$. By Proposition 9.10, $f^{(1)}(x) = \vartheta_{\mathcal{B}(\mathcal{X}, \mathcal{Y})}$. Let $D_1 := \text{dom}(f^{(1)})$. Then, $D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) = D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$. Then, $\text{span}(\mathcal{A}_{D_1}(x)) = \mathcal{X}$, $\forall x \in D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$. By recursively applying the above argument and Proposition 9.28, we have $f^{(i)}(x) = \vartheta_{\mathcal{B}_{\mathcal{S}_i(\mathcal{X}, \mathcal{Y})}}(x)$, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, $\forall i \in \mathbb{N}$. This completes the proof of the proposition. $\square$

Proposition 9.34 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D_2 \subseteq D_1 \subseteq \mathcal{X}$, $x_0 \in D_2$, $f : D_1 \rightarrow \mathcal{Y}$, $g := f|_{D_2}$, $k \in \mathbb{N}$, and $n \in \mathbb{N} \cup \{\infty\}$. Then, the following statements holds.*

- (i) *If, $\forall x \in D_2$, $f^{(k)}(x)$ exists and $\overline{\text{span}(\mathcal{A}_{D_2}(x))} = \mathcal{X}$, then g is k -times Fréchet differentiable and $g^{(i)}(x) = f^{(i)}(x)$, $\forall x \in D_2$, $\forall i \in \{1, \dots, k\}$.*
- (ii) *If, $\forall x \in D_2$, $g^{(k)}(x)$ exists and $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$ such that $D_1 \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x) = D_2 \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x)$, then $f^{(k)}(x)$ exists and $f^{(i)}(x) = g^{(i)}(x)$, $\forall x \in D_2$, $\forall i \in \{1, \dots, k\}$.*
- (iii) *If f is $\mathcal{C}_n$ at x_0 and $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\overline{\text{span}(\mathcal{A}_{D_2}(x))} = \mathcal{X}$, $\forall x \in D_2 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, then, g is $\mathcal{C}_n$ at x_0 .*
- (iv) *If g is $\mathcal{C}_n$ at x_0 and $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) = D_2 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, then f is $\mathcal{C}_n$ at x_0 .*

Proof (i) We will use mathematical induction on k to prove this statement.

1° $k = 1$. $\forall x \in D_2$, by Proposition 9.11, we have $g^{(1)}(x)$ exists and $g^{(1)}(x) = f^{(1)}(x)$. Hence, the result holds in this case.

2° Assume that the result holds for $k \leq \bar{k} \in \mathbb{N}$.

3° Consider the case $k = \bar{k} + 1$. By inductive assumption, $g^{(\bar{k})}(x)$ exists and $g^{(i)}(x) = f^{(i)}(x)$, $\forall x \in D_2$, $\forall i \in \{1, \dots, \bar{k}\}$. Then, $\text{dom}(g^{(\bar{k})}) =$

$D_2 \subseteq \text{dom} \left(f^{(\bar{k})} \right) =: \bar{D}_1$. $\forall x \in D_2$, by the assumption, $f^{(\bar{k})}$ is differentiable at x and $\overline{\text{span} (A_{D_2}(x))} = \mathcal{X}$. By Proposition 9.11, $g^{(\bar{k})}$ is Fréchet differentiable at x and $g^{(\bar{k}+1)}(x) = Dg^{(\bar{k})}(x) = Df^{(\bar{k})}(x) = f^{(\bar{k}+1)}(x)$. This completes the induction process.

(ii) We will use mathematical induction on k to prove this statement.

1° $k = 1$. $\forall x \in D_2$, by Proposition 9.11, $f^{(1)}(x)$ exists and $f^{(1)}(x) = g^{(1)}(x)$. Hence, the result holds.

2° Assume that the result holds for $k \leq \bar{k} \in \mathbb{N}$.

3° Consider the case $k = \bar{k} + 1$. By inductive assumption, $f^{(\bar{k})}(x)$ exists and $f^{(i)}(x) = g^{(i)}(x)$, $\forall x \in D_2$, $\forall i \in \{1, \dots, \bar{k}\}$. Then, $\text{dom} \left(g^{(\bar{k})} \right) = D_2 \subseteq \text{dom} \left(f^{(\bar{k})} \right) =: \bar{D}_1 \subseteq D_1$. $\forall x \in D_2$, by the assumption, $g^{(\bar{k})}$ is differentiable at x and $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$ such that $\bar{D}_1 \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x) = D_2 \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x)$. Then, by Proposition 9.11, $f^{(\bar{k})}$ is Fréchet differentiable at x and $f^{(\bar{k}+1)}(x) = Df^{(\bar{k})}(x) = Dg^{(\bar{k})}(x) = g^{(\bar{k}+1)}(x)$. This completes the induction process.

(iii) We will distinguish two exhaustive and mutually exclusive cases:

Case 1: $n \in \mathbb{N}$; Case 2: $n = \infty$.

Case 1: $n \in \mathbb{N}$. Without loss of generality, assume f is n -times differentiable at x , $\forall x \in D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$. Let $\bar{D}_2 := D_2 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$ and $\bar{g} := f|_{\bar{D}_2}$. $\forall x \in \bar{D}_2$, $f^{(n)}(x)$ exists. Let $\delta_x := \delta - \|x - x_0\| \in (0, \infty) \subset \mathbb{R}$. Then, $\bar{D}_2 \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x) = D_2 \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x)$. Hence, $\overline{\text{span} (A_{\bar{D}_2}(x))} = \overline{\text{span} (A_{D_2}(x))} = \mathcal{X}$. Then, by (i), $\bar{g}$ is n -times differentiable and $\bar{g}^{(i)}(x) = f^{(i)}(x)$, $\forall x \in \bar{D}_2$, $\forall i \in \{1, \dots, n\}$. By (ii), $g^{(n)}(x)$ exists and $g^{(i)}(x) = \bar{g}^{(i)}(x) = f^{(i)}(x)$, $\forall x \in \bar{D}_2$, $\forall i \in \{1, \dots, n\}$. By the continuity of $f^{(n)}$ at x_0 , $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \bar{\delta} \in (0, \delta] \subset \mathbb{R}$, $\forall x \in \text{dom} (f^{(n)}) \cap \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta}) = D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta})$, we have $\|f^{(n)}(x) - f^{(n)}(x_0)\| < \epsilon$. $\forall x \in \text{dom} (g^{(n)}) \cap \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta}) = D_2 \cap \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta})$, we have $\|g^{(n)}(x) - g^{(n)}(x_0)\| = \|f^{(n)}(x) - f^{(n)}(x_0)\| < \epsilon$. Hence, $g^{(n)}$ is continuous at x_0 . Therefore, g is $\mathcal{C}_n$ at x_0 .

Case 2: $n = \infty$. By the assumption, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that f is $\mathcal{C}_i$ at x , $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\forall i \in \mathbb{N}$. By Case 1, g is $\mathcal{C}_i$ at x , $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\forall i \in \mathbb{N}$. Hence, g is $\mathcal{C}_{\infty}$ at x_0 .

(iv) We will distinguish two exhaustive and mutually exclusive cases:

Case 1: $n \in \mathbb{N}$; Case 2: $n = \infty$.

Case 1: $n \in \mathbb{N}$. Without loss of generality, assume g is n -times differentiable at x , $\forall x \in D_2 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) =: \bar{D}_2$. Let $\bar{g} := g|_{\bar{D}_2}$. $\forall x \in \bar{D}_2$, $g^{(n)}(x)$ exists. Let $\delta_x := \delta - \|x - x_0\| \in (0, \infty) \subset \mathbb{R}$. Then, $\bar{D}_2 \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x) = D_2 \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x) = D_1 \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x)$. Hence, $\overline{\text{span} (A_{\bar{D}_2}(x))} = \overline{\text{span} (A_{D_2}(x))} = \mathcal{X}$. Then, by (i), $\bar{g}$ is n -times differentiable and $\bar{g}^{(i)}(x) = g^{(i)}(x)$, $\forall x \in \bar{D}_2$, $\forall i \in \{1, \dots, n\}$. By (ii), $f^{(n)}(x)$ exists and $f^{(i)}(x) = \bar{g}^{(i)}(x) = g^{(i)}(x)$, $\forall x \in \bar{D}_2$, $\forall i \in \{1, \dots, n\}$. By the continuity of $g^{(n)}$ at x_0 , $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \bar{\delta} \in (0, \delta] \subset \mathbb{R}$, $\forall x \in \text{dom} (g^{(n)}) \cap \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta}) = D_2 \cap \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta}) = D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta})$, we have

$\|g^{(n)}(x) - g^{(n)}(x_0)\| < \epsilon$. $\forall x \in \text{dom}(f^{(n)}) \cap \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta}) = D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta})$, we have $\|f^{(n)}(x) - f^{(n)}(x_0)\| = \|g^{(n)}(x) - g^{(n)}(x_0)\| < \epsilon$. Hence, $f^{(n)}$ is continuous at x_0 . Therefore, f is $\mathcal{C}_n$ at x_0 .

Case 2: $n = \infty$. By the assumption, g is $\mathcal{C}_i$ at x , $\forall x \in \bar{D}_2$, $\forall i \in \mathbb{N}$. By Case 1, f is $\mathcal{C}_i$ at x , $\forall x \in \bar{D}_2$, $\forall i \in \mathbb{N}$. Hence, f is $\mathcal{C}_\infty$ at x_0 .

This completes the proof of the proposition. $\square$

Proposition 9.35 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$. Assume that there exist open sets $O_1, O_2 \subseteq \mathcal{X}$ such that $D_1 := D \cap O_1$ and $D_2 := D \cap O_2$ satisfy $D_1 \cup D_2 = D$ and $f|_{D_1}$ and $f|_{D_2}$ are $\mathcal{C}_k$, where $k \in \mathbb{N} \cup \{\infty\}$. Then, f is $\mathcal{C}_k$.*

Proof $\forall x \in D$, without loss of generality, assume that $x \in D_1$. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{X}}(x, \delta) \subseteq O_1$, which implies that $D_1 \cap \mathcal{B}_{\mathcal{X}}(x, \delta) = D \cap \mathcal{B}_{\mathcal{X}}(x, \delta)$. Since $f|_{D_1}$ is $\mathcal{C}_k$, then, by Proposition 9.34, $\forall \bar{x} \in D_1$, $f^{(i)}(\bar{x}) = f|_{D_1}^{(i)}(\bar{x})$, $\forall i \in \mathbb{N}$ with $i \leq k$. Then, f is $\mathcal{C}_k$ at x . By the arbitrariness of x , f is $\mathcal{C}_k$. This completes the proof of the proposition. $\square$

Note that D_1 and D_2 are open sets in the subset topology of D . This result should be compared with Theorem 3.11, where continuity on D can be concluded when D_1 and D_2 are relatively open or are relatively closed. For continuous differentiability, D_1 and D_2 must be relatively open for the conclusion to hold.

Proposition 9.36 *Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, and $f : \mathcal{X} \rightarrow \mathcal{X}$ be given by $f = \text{id}_{\mathcal{X}}$, that is $f(x) = x$, $\forall x \in \mathcal{X}$. Then, f is $\mathcal{C}_\infty$, $f^{(1)}(x) = \text{id}_{\mathcal{X}}$, and $f^{(i+1)}(x) = \vartheta_{\text{BS}_{i+1}(\mathcal{X}, \mathcal{X})}$, $\forall x \in \mathcal{X}$, $\forall i \in \mathbb{N}$.*

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.37 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, and $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{X}$ be given by $f = \pi_{\mathcal{X}}$, that is $f(x, y) = x$, $\forall (x, y) \in \mathcal{X} \times \mathcal{Y}$. Then, f is $\mathcal{C}_\infty$, $f^{(1)}(x, y) = \begin{bmatrix} \text{id}_{\mathcal{X}} & \vartheta_{\text{B}(\mathcal{Y}, \mathcal{X})} \end{bmatrix}$ and $f^{(i+1)}(x, y) = \vartheta_{\text{BS}_{i+1}(\mathcal{X} \times \mathcal{Y}, \mathcal{X})}$, $\forall (x, y) \in \mathcal{X} \times \mathcal{Y}$, $\forall i \in \mathbb{N}$.*

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.38 *Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $f : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ be given by $f(x_1, x_2) = x_1 + x_2$, $\forall (x_1, x_2) \in \mathcal{X} \times \mathcal{X}$. Then, f is $\mathcal{C}_\infty$, $f^{(1)}(x_1, x_2) = \begin{bmatrix} \text{id}_{\mathcal{X}} & \text{id}_{\mathcal{X}} \end{bmatrix}$, and $f^{(i+1)}(x_1, x_2) = \vartheta_{\text{BS}_{i+1}(\mathcal{X} \times \mathcal{X}, \mathcal{X})}$, $\forall (x_1, x_2) \in \mathcal{X} \times \mathcal{X}$, $\forall i \in \mathbb{N}$.*

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.39 *Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $f : \mathbb{K} \times \mathcal{X} \rightarrow \mathcal{X}$ be given by $f(\alpha, x) = \alpha x$, $\forall (\alpha, x) \in \mathbb{K} \times \mathcal{X}$. Then, f is $\mathcal{C}_\infty$, $f^{(1)}(\alpha, x) = \begin{bmatrix} x & \alpha \text{id}_{\mathcal{X}} \end{bmatrix}$, $f^{(2)}(\alpha, x)(d_2, h_2)(d_1, h_1) = d_1 h_2 + d_2 h_1$, and $f^{(i+2)}(\alpha, x) = \vartheta_{\text{BS}_{i+2}(\mathbb{K} \times \mathcal{X}, \mathcal{X})}$, $\forall (\alpha, x) \in \mathbb{K} \times \mathcal{X}$, $\forall i \in \mathbb{N}$, $\forall (d_1, h_1) \in \mathbb{K} \times \mathcal{X}$, $\forall (d_2, h_2) \in \mathbb{K} \times \mathcal{X}$.*

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.40 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f_1 : D \rightarrow \mathcal{Y}$, $f_2 : D \rightarrow \mathcal{Y}$, $x_0 \in D$, $\alpha_1, \alpha_2 \in \mathbb{K}$, $n \in \mathbb{N}$, $k \in \mathbb{N} \cup \{\infty\}$, and $g : D \rightarrow \mathcal{Y}$ be given by $g(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x)$, $\forall x \in D$. If f_1 and f_2 are n -times differentiable, then, g is n -times differentiable and $g^{(i)}(x) = \alpha_1 f_1^{(i)}(x) + \alpha_2 f_2^{(i)}(x)$, $\forall x \in D$, $\forall i \in \{1, \dots, n\}$. If f_1 and f_2 are $\mathcal{C}_k$ at x_0 , then g is $\mathcal{C}_k$ at x_0 and $g^{(i)}(x) = \alpha_1 f_1^{(i)}(x) + \alpha_2 f_2^{(i)}(x)$, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, $\forall i \in \mathbb{N}$ with $i \leq k$, for some $\delta_0 \in (0, \infty) \subset \mathbb{R}$.

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.41 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $f : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \rightarrow \mathcal{Y}$ be given by $f(A, x) = Ax$, $\forall (A, x) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$. Then, f is $\mathcal{C}_{\infty}$, $f^{(1)}(A, x) = \begin{bmatrix} \text{ro}(x) & A \end{bmatrix}$, $f^{(2)}(A, x)(\Delta_2, h_2)(\Delta_1, h_1) = \Delta_1 h_2 + \Delta_2 h_1$, and $f^{(i+2)}(A, x) = \vartheta_{B_{S_{i+2}}(B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \mathcal{Y})}$, $\forall (A, x) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$, $\forall i \in \mathbb{N}$, $\forall (\Delta_1, h_1) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$, $\forall (\Delta_2, h_2) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$.

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.42 Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $f : B(\mathcal{Y}, \mathcal{Z}) \times B(\mathcal{X}, \mathcal{Y}) \rightarrow B(\mathcal{X}, \mathcal{Z})$ be given by $f(A_{yz}, A_{xy}) = A_{yz} A_{xy}$, $\forall (A_{yz}, A_{xy}) \in B(\mathcal{Y}, \mathcal{Z}) \times B(\mathcal{X}, \mathcal{Y})$. Then, f is $\mathcal{C}_{\infty}$, $f^{(1)}(A_{yz}, A_{xy}) = \begin{bmatrix} \text{ro}(A_{xy}) & A_{yz} \end{bmatrix}$, $f^{(2)}(A_{yz}, A_{xy})(\Delta_{yz2}, \Delta_{xy2})(\Delta_{yz1}, \Delta_{xy1}) = \Delta_{yz1} \Delta_{xy2} + \Delta_{yz2} \Delta_{xy1}$, and $f^{(i+2)}(A_{yz}, A_{xy}) = \vartheta_{B_{S_{i+2}}(B(\mathcal{Y}, \mathcal{Z}) \times B(\mathcal{X}, \mathcal{Y}), B(\mathcal{X}, \mathcal{Z}))}$, $\forall (A_{yz}, A_{xy}) \in B(\mathcal{Y}, \mathcal{Z}) \times B(\mathcal{X}, \mathcal{Y})$, $\forall i \in \mathbb{N}$, $\forall (\Delta_{yz1}, \Delta_{xy1}) \in B(\mathcal{Y}, \mathcal{Z}) \times B(\mathcal{X}, \mathcal{Y})$, $\forall (\Delta_{yz2}, \Delta_{xy2}) \in B(\mathcal{Y}, \mathcal{Z}) \times B(\mathcal{X}, \mathcal{Y})$.

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 9.43 Let $\mathcal{X}_1, \dots, \mathcal{X}_p$ and $\mathcal{Y}_1, \dots, \mathcal{Y}_m$ be normed linear spaces over $\mathbb{K}$, where $p, m \in \mathbb{N}$, $\mathcal{Z}_{ji} := B(\mathcal{X}_i, \mathcal{Y}_j)$, $i = 1, \dots, p$, $j = 1, \dots, m$, $\mathcal{Z} := \prod_{j=1}^m \prod_{i=1}^p \mathcal{Z}_{ji}$, and $f : \mathcal{Z} \rightarrow B\left(\prod_{i=1}^p \mathcal{X}_i, \prod_{j=1}^m \mathcal{Y}_j\right)$ be given by

$$f(A_{11}, \dots, A_{mp}) = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mp} \end{bmatrix}; \quad \forall (A_{11}, \dots, A_{mp}) \in \mathcal{Z}$$

Then, f is $\mathcal{C}_{\infty}$ and

$$\begin{aligned} f^{(1)}(A_{11}, \dots, A_{mp})(\Delta_{11}, \dots, \Delta_{mp}) &= \begin{bmatrix} \Delta_{11} & \cdots & \Delta_{1p} \\ \vdots & & \vdots \\ \Delta_{m1} & \cdots & \Delta_{mp} \end{bmatrix} \\ f^{(l+1)}(A_{11}, \dots, A_{mp}) &= \vartheta_{B_{S_{l+1}}(\mathcal{Z}, B(\prod_{i=1}^p \mathcal{X}_i, \prod_{j=1}^m \mathcal{Y}_j))} \end{aligned}$$

$\forall (A_{11}, \dots, A_{mp}) \in \mathcal{Z}$, $\forall l \in \mathbb{N}$, $\forall (\Delta_{11}, \dots, \Delta_{mp}) \in \mathcal{Z}$.

Proof This is straightforward, and is therefore omitted. $\square$

9.4.4 Properties of high order derivatives

Proposition 9.44 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f_1 : D \rightarrow \mathcal{Y}$, $f_2 : D \rightarrow \mathcal{Z}$, $x_0 \in D$, $k \in \mathbb{N}$, and $g : D \rightarrow \mathcal{Y} \times \mathcal{Z}$ be given by $g(x) = (f_1(x), f_2(x))$, $\forall x \in D$. Then, the following statements hold.*

- (i) $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $f_1^{(k)}(x)$ and $f_2^{(k)}(x)$ exist, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$ if, and only if, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $g^{(k)}(x)$ exists, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$. In this case $g^{(i)}(x) = \begin{bmatrix} f_1^{(i)}(x) \\ f_2^{(i)}(x) \end{bmatrix}$, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, $\forall i \in \{1, \dots, k\}$.
- (ii) Let $n \in \mathbb{N} \cup \{\infty\}$. Then, f_1 and f_2 are $\mathcal{C}_n$ at x_0 if, and only if, g is $\mathcal{C}_n$ at x_0 .

Proof (i) We will use mathematical induction on k to prove this statement.

1° $k = 1$. The statement holds by Proposition 9.19.

2° Assume that the result holds for $k = \bar{k} \in \mathbb{N}$.

3° Consider the case $k = \bar{k} + 1$. “Necessity” Let $f_1^{(\bar{k}+1)}(x)$ and $f_2^{(\bar{k}+1)}(x)$ exist, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$. By inductive assumption, $g^{(\bar{k})}(x)$ exists and $g^{(i)}(x) = \begin{bmatrix} f_1^{(i)}(x) \\ f_2^{(i)}(x) \end{bmatrix}$, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, $\forall i \in \{1, \dots, \bar{k}\}$. Let

$\bar{D}_1 := \text{dom}(f_1^{(\bar{k})})$, $\bar{D}_2 := \text{dom}(f_2^{(\bar{k})})$ and $\bar{D} := \text{dom}(g^{(\bar{k})})$. Then, $\bar{D}_1 \subseteq D$, $\bar{D}_2 \subseteq D$, $\bar{D} \subseteq D$, $\bar{D} \supseteq D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, and $\bar{D} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) \subseteq \bar{D}_1 \cap \bar{D}_2$. This implies that $\bar{D} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) = D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) =: \hat{D}$ and $\text{span}(\mathcal{A}_{\bar{D}}(x)) = \text{span}(\mathcal{A}_D(x)) = \mathcal{X}$, $\forall x \in \hat{D}$. By the assumption, $f_1^{(\bar{k})}$ and $f_2^{(\bar{k})}$ are differentiable at x , $\forall x \in \hat{D}$. By Propositions 9.19 and 9.11, $\forall x \in \hat{D}$, $g^{(\bar{k})}$ is differentiable at x and $g^{(\bar{k}+1)}(x) = D g^{(\bar{k})}|_{\hat{D}}(x) =$

$$\begin{bmatrix} D f_1^{(\bar{k})}|_{\hat{D}}(x) \\ D f_2^{(\bar{k})}|_{\hat{D}}(x) \end{bmatrix} = \begin{bmatrix} f_1^{(\bar{k}+1)}(x) \\ f_2^{(\bar{k}+1)}(x) \end{bmatrix}.$$

“Sufficiency” Let $g^{(\bar{k}+1)}(x)$ exist, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$. By inductive assumption, $f_1^{(\bar{k})}(x)$ and $f_2^{(\bar{k})}(x)$ exist and $g^{(i)}(x) = \begin{bmatrix} f_1^{(i)}(x) \\ f_2^{(i)}(x) \end{bmatrix}$, $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, $\forall i \in \{1, \dots, \bar{k}\}$. Let $\bar{D}_1 := \text{dom}(f_1^{(\bar{k})})$, $\bar{D}_2 := \text{dom}(f_2^{(\bar{k})})$ and $\bar{D} := \text{dom}(g^{(\bar{k})})$. Then, $\bar{D}_1 \subseteq D$, $\bar{D}_2 \subseteq D$, $\bar{D} \subseteq D$, $\bar{D} \supseteq D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0)$, and $D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) \subseteq \bar{D}_1 \cap \bar{D}_2$. This implies that $\bar{D} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) = D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) = \bar{D}_1 \cap \bar{D}_2 = \bar{D}_2 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) =: \hat{D}$. Then, $\text{span}(\mathcal{A}_{\bar{D}}(x)) = \text{span}(\mathcal{A}_D(x)) = \mathcal{X}$, $\forall x \in \hat{D}$. By Propositions 9.11 and 9.19, we have, $\forall x \in \hat{D}$, $f_1^{(\bar{k})}|_{\hat{D}}$ and $f_2^{(\bar{k})}|_{\hat{D}}$ are differen-

tion 9.11, we have $D f_1^{(\bar{k})} \Big|_{\hat{D}}(x) = D f_1^{(\bar{k})}(x) = f_1^{(\bar{k}+1)}(x)$ and $D f_2^{(\bar{k})} \Big|_{\hat{D}}(x) = D f_2^{(\bar{k})}(x) = f_2^{(\bar{k}+1)}(x)$, $\forall x \in \hat{D}$. Then, $g^{(\bar{k}+1)}(x) = \begin{bmatrix} f_1^{(\bar{k}+1)}(x) \\ f_2^{(\bar{k}+1)}(x) \end{bmatrix}$, $\forall x \in \hat{D}$.

This completes the induction process.

(ii) We will distinguish two exhaustive and mutually exclusive cases:

Case 1: $n \in \mathbb{N}$; Case 2: $n = \infty$.

Case 1: $n \in \mathbb{N}$. “Sufficiency” Let g be $\mathcal{C}_n$ at x_0 . Then, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that g is n -times differentiable at x , $\forall x \in D \cap \mathcal{B}_X(x_0, \delta_0) =: \bar{D}$. By (i), f_1 and f_2 are n -times differentiable at x , $\forall x \in \bar{D}$ and $g^{(n)}(x) = \begin{bmatrix} f_1^{(n)}(x) \\ f_2^{(n)}(x) \end{bmatrix}$, $\forall x \in \bar{D}$. By the continuity of $g^{(n)}$ at x_0 and Proposition 3.32, $f_1^{(n)}$ and $f_2^{(n)}$ are continuous at x_0 . Then, f_1 and f_2 are $\mathcal{C}_n$ at x_0 .

“Necessity” Let f_1 and f_2 be $\mathcal{C}_n$ at x_0 . Then, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that f_1 and f_2 are n -times differentiable at x , $\forall x \in D \cap \mathcal{B}_X(x_0, \delta_0) =: \bar{D}$. By (i), g is n -times differentiable at x , $\forall x \in \bar{D}$ and $g^{(n)}(x) = \begin{bmatrix} f_1^{(n)}(x) \\ f_2^{(n)}(x) \end{bmatrix}$, $\forall x \in \bar{D}$. By the continuity of $f_1^{(n)}$ and $f_2^{(n)}$ at x_0 and Proposition 3.32, $g^{(n)}$ is continuous at x_0 . Then, g is $\mathcal{C}_n$ at x_0 .

Case 2: $n = \infty$. “Sufficiency” Let g be $\mathcal{C}_\infty$ at x_0 . Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that g is $\mathcal{C}_i$ at x , $\forall x \in D \cap \mathcal{B}_X(x_0, \delta)$, $\forall i \in \mathbb{N}$. By Case 1, f_1 and f_2 are $\mathcal{C}_i$ at x , $\forall x \in D \cap \mathcal{B}_X(x_0, \delta)$, $\forall i \in \mathbb{N}$. Then, f_1 and f_2 are $\mathcal{C}_\infty$ at x_0 . “Necessity” Let f_1 and f_2 be $\mathcal{C}_\infty$ at x_0 . Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that f_1 and f_2 are $\mathcal{C}_i$ at x , $\forall x \in D \cap \mathcal{B}_X(x_0, \delta)$, $\forall i \in \mathbb{N}$. By Case 1, g is $\mathcal{C}_i$ at x , $\forall x \in D \cap \mathcal{B}_X(x_0, \delta)$, $\forall i \in \mathbb{N}$. Then, g is $\mathcal{C}_\infty$ at x_0 .

This completes the proof of the proposition. $\square$

Proposition 9.45 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D_1 \subseteq \mathcal{X}$, $D_2 \subseteq \mathcal{Y}$, $f : D_1 \rightarrow D_2$, $g : D_2 \rightarrow \mathcal{Z}$, $x_0 \in D_1$, and $y_0 := f(x_0) \in D_2$. Then, the following statements hold.*

- (i) *Assume that f is $\mathcal{C}_k$ at x_0 and g is $\mathcal{C}_k$ at y_0 , for some $k \in \mathbb{N} \cup \{\infty\}$. Then, $h := g \circ f$ is $\mathcal{C}_k$ at x_0 .*
- (ii) *Let $k \in \mathbb{N}$. Assume that f is k -times differentiable and g is k -times differentiable. Then, h is k -times differentiable.*

Proof (i) We will first use mathematical induction on k to show that the result holds if $k \in \mathbb{N}$.

1° $k = 1$. By g being $\mathcal{C}_1$ at y_0 , then $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $g^{(1)}(y)$ exists, $\forall y \in D_2 \cap \mathcal{B}_Y(y_0, \delta_1)$, and $g^{(1)}$ is continuous at y_0 . By

f being $\mathcal{C}_1$ at x_0 , then, by Proposition 9.7, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $f(x) \in D_2 \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta_1)$ and $f^{(1)}(x)$ exists, $\forall x \in D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, and $f^{(1)}$ is continuous at x_0 . $\forall x \in D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, by Chain Rule, $h^{(1)}(x)$ exists and $h^{(1)}(x) = g^{(1)}(f(x))f^{(1)}(x)$. By Propositions 3.12, 9.7, 3.32, and 9.42, $h^{(1)}$ is continuous at x_0 . Hence, h is $\mathcal{C}_1$ at x_0 .

2° Assume that the result holds for $k \leq \bar{k} \in \mathbb{N}$.

3° Consider the case $k = \bar{k} + 1$. By g being $\mathcal{C}_{\bar{k}+1}$ at y_0 , then $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $g^{(1)}(y)$ exists, $\forall y \in D_2 \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta_1) =: \bar{D}$, and $g^{(1)}$ is $\mathcal{C}_{\bar{k}}$ at y_0 . By f being $\mathcal{C}_{\bar{k}+1}$ at x_0 , then, by Proposition 9.7, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $f(x) \in \bar{D}$ and $f^{(1)}(x)$ exists, $\forall x \in D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) =: \hat{D}$, and $f^{(1)}$ is $\mathcal{C}_{\bar{k}}$ at x_0 . $\forall x \in \hat{D}$, by Chain Rule, $h^{(1)}(x)$ exists and $h^{(1)}(x) = g^{(1)}(f(x))f^{(1)}(x)$. Then, $h^{(1)}|_{\hat{D}} : \hat{D} \rightarrow \mathcal{B}(\mathcal{X}, \mathcal{Z})$ is given by $h^{(1)}|_{\hat{D}}(x) = g^{(1)}|_{\bar{D}}(f|_{\hat{D}}(x))f^{(1)}|_{\hat{D}}(x)$, $\forall x \in \hat{D}$. By Proposition 9.34, $g^{(1)}|_{\bar{D}}$ is $\mathcal{C}_{\bar{k}}$ at y_0 and $f|_{\hat{D}}$ and $f^{(1)}|_{\hat{D}}$ are $\mathcal{C}_{\bar{k}}$ at x_0 . By inductive assumption and Propositions 9.44, 9.42, and 3.32, $h^{(1)}|_{\hat{D}}$ is $\mathcal{C}_{\bar{k}}$ at x_0 . By Proposition 9.34, $h^{(1)}$ is $\mathcal{C}_{\bar{k}}$ at x_0 . Hence, h is $\mathcal{C}_{\bar{k}+1}$ at x_0 .

This completes the induction process.

When $k = \infty$, then $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that g is $\mathcal{C}_i$ at y , $\forall y \in D_2 \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta)$, and f is $\mathcal{C}_i$ at x , $\forall x \in D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\forall i \in \mathbb{N}$, which further implies by the induction conclusion, h is $\mathcal{C}_i$ at x , $\forall x \in D_1 \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\forall i \in \mathbb{N}$. Hence, h is $\mathcal{C}_{\infty}$ at x_0 .

(ii) We will first use mathematical induction on k to show that the result holds.

1° $k = 1$. By Chain Rule, $h^{(1)}(x)$ exists and $h^{(1)}(x) = g^{(1)}(f(x))f^{(1)}(x)$, $\forall x \in D_1$. Hence, the result holds.

2° Assume that the result holds for $k \leq \bar{k} \in \mathbb{N}$.

3° Consider the case $k = \bar{k} + 1$. By Chain Rule, $h^{(1)}(x)$ exists and $h^{(1)}(x) = g^{(1)}(f(x))f^{(1)}(x)$, $\forall x \in D_1$. By inductive assumption and Propositions 9.44 and 9.42, $h^{(1)}$ is $\bar{k}$ -times differentiable. Hence, h is $(\bar{k} + 1)$ -times differentiable. This completes the induction process and the proof of the proposition. $\square$

Proposition 9.46 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X} \times \mathcal{Y}$, $f : D \rightarrow \mathcal{Z}$ be partial differentiable with respect to x , and partial differentiable with respect to y . Then, the following statements hold.*

- (i) *If f is $(n+1)$ -times differentiable, where $n \in \mathbb{N}$, then, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are n -times differentiable.*
- (ii) *If f is $\mathcal{C}_1$ at $(x_0, y_0) \in D$, then $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are continuous at (x_0, y_0) .*
- (iii) *If f is $\mathcal{C}_n$ at $(x_0, y_0) \in D$, where $n \in \{2, 3, \dots\} \cup \{\infty\}$, then $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are $\mathcal{C}_{n-1}$ at (x_0, y_0) .*

Proof By Proposition 9.9, $f^{(1)}(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \end{bmatrix}$, $\forall (x, y) \in D$. Define $g : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{B}(\mathcal{X}, \mathcal{X} \times \mathcal{Y})$ by $g(x, y) = \begin{bmatrix} \text{id}_{\mathcal{X}} \\ \vartheta_{\mathcal{B}(\mathcal{X}, \mathcal{Y})} \end{bmatrix}$. By Proposition 9.33, g is $\mathcal{C}_{\infty}$. It is clear that $\frac{\partial f}{\partial x}(x, y) = f^{(1)}(x, y)g(x, y)$, $\forall (x, y) \in D$.

(i) Since f is $(n+1)$ -times differentiable, then $f^{(1)}$ is n -times differentiable. By Propositions 9.42, 9.44, and 9.45, we have $\frac{\partial f}{\partial x}$ is n -times differentiable. By symmetry, $\frac{\partial f}{\partial y}$ is n -times differentiable. This completes the proof of the proposition.

(ii) Since f is $\mathcal{C}_1$ at (x_0, y_0) , then $f^{(1)}$ is continuous at (x_0, y_0) . By Propositions 9.42, 9.7, 3.32, and 3.12, we have $\frac{\partial f}{\partial x}$ is continuous at (x_0, y_0) . By symmetry, $\frac{\partial f}{\partial y}$ is continuous at (x_0, y_0) .

(iii) Since f is $\mathcal{C}_n$ at (x_0, y_0) , then $f^{(1)}$ is $\mathcal{C}_{n-1}$ at (x_0, y_0) . By Propositions 9.42, 9.44, and 9.45, we have $\frac{\partial f}{\partial x}$ is $\mathcal{C}_{n-1}$ at (x_0, y_0) . By symmetry, $\frac{\partial f}{\partial y}$ is $\mathcal{C}_{n-1}$ at (x_0, y_0) .

This completes the proof of the proposition. $\square$

Proposition 9.47 Let $\mathcal{X}_1, \dots, \mathcal{X}_p$, and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, where $p \in \{2, 3, \dots\}$, $D \subseteq \mathcal{X} := \prod_{i=1}^p \mathcal{X}_i$, $f : D \rightarrow \mathcal{Y}$, $x_0 \in D^\circ$, and $k \in \mathbb{N}$. Assume that $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that all partial derivatives of f up to k th order exist and are continuous at x , $\forall x \in \tilde{D} := \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) \subseteq D$. Then, f is $\mathcal{C}_k$ at x , $\forall x \in \tilde{D}$.

Proof We will prove this using mathematical induction on k .

1° $k = 1$. $\forall x \in \tilde{D}$, by repeated application of Proposition 9.24, we have $f^{(1)}(x)$ exists and $f^{(1)}(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) & \cdots & \frac{\partial f}{\partial x_p}(x) \end{bmatrix}$.

2° Assume that the result holds for $k = \bar{k} \in \mathbb{N}$.

3° Consider the case $k = \bar{k} + 1$. $\forall x \in \tilde{D}$, by the case $k = 1$, $f^{(1)}(x)$ exists and $f^{(1)}(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) & \cdots & \frac{\partial f}{\partial x_p}(x) \end{bmatrix}$. $\forall i \in \{1, \dots, p\}$, by the assumption, all partial derivatives of $\frac{\partial f}{\partial x_i}$ up to $\bar{k}$ th order exist and are continuous at x . Then, by inductive assumption, $\frac{\partial f}{\partial x_i}$ is $\mathcal{C}_{\bar{k}}$ at x . Define the function $g : \hat{D} \rightarrow \prod_{i=1}^p \mathcal{B}(\mathcal{X}_i, \mathcal{Y})$ by $g(x) = (\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_p}(x))$, $\forall x \in \hat{D} := \bigcap_{i=1}^p \text{dom}\left(\frac{\partial f}{\partial x_i}\right)$. Clearly, $\tilde{D} \subseteq \hat{D} \subseteq D$. Then, by Proposition 9.44,

$$g \text{ is } \bar{k}\text{-times differentiable at } x \text{ and } g^{(j)}(x) = \begin{bmatrix} D^j \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ D^j \frac{\partial f}{\partial x_p}(x) \end{bmatrix}, \forall j = 1, \dots, \bar{k}.$$

By Proposition 3.32, g is $\mathcal{C}_{\bar{k}}$ at x . By Propositions 9.45 and 9.43, $f^{(1)}$ is $\mathcal{C}_{\bar{k}}$ at x . Therefore, f is $\mathcal{C}_{\bar{k}+1}$ at x . This completes the induction process and the proof of the proposition. $\square$

Theorem 9.48 (Taylor's Theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, $x_0, x_1 \in D$, and $n \in \mathbb{N}$. Let $\bar{D} := I := [0, 1] \subset \mathbb{R}$ and $\tilde{\bar{D}} := I^\circ$ if $\mathbb{K} = \mathbb{R}$ or $\bar{D} := \{a + i0 \mid a \in I\} \subset \mathbb{C}$ and $\tilde{\bar{D}} := \{a + i0 \mid a \in I^\circ\} \subset \mathbb{C}$ if $\mathbb{K} = \mathbb{C}$. Let $\varphi : \bar{D} \rightarrow D$ be given by $\varphi(t) = tx_1 + (1-t)x_0$, $\forall t \in \bar{D}$. Assume that $\text{dom}(f^{(n)}) \supseteq \varphi(\bar{D})$, $\text{dom}(f^{(n+1)}) \supseteq \varphi(\tilde{\bar{D}})$, and $f^{(n)}$ is continuous at $x = \varphi(t)$, $\forall t \in \bar{D}$. Let $R_n \in \mathcal{Y}$ be given by*

$$R_n := f(x_1) - \left(f(x_0) + \frac{1}{1!} f^{(1)}(x_0)(x_1 - x_0) + \cdots + \frac{1}{n!} f^{(n)}(x_0) \underbrace{(x_1 - x_0) \cdots (x_1 - x_0)}_{n\text{-times}} \right)$$

Then, the following statements hold.

(i) If $\mathcal{Y} = \mathbb{R}$ and $\mathbb{K} = \mathbb{R}$, then $\exists \bar{t}_0 \in I^\circ$ such that

$$R_n = \frac{1}{(n+1)!} f^{(n+1)}(\varphi(\bar{t}_0)) \underbrace{(x_1 - x_0) \cdots (x_1 - x_0)}_{(n+1)\text{-times}}$$

(ii) $\exists \bar{t}_0 \in I^\circ$ such that

$$\|R_n\| \leq \frac{1}{(n+1)!} \|f^{(n+1)}(\varphi(\bar{t}_0))\| \|x_1 - x_0\|^{n+1}$$

Proof (i) Let $\mathcal{Y} = \mathbb{R}$ and $\mathbb{K} = \mathbb{R}$. Define $F : I \rightarrow \mathbb{R}$ by

$$F(t) = f(\varphi(1)) - \left(f(\varphi(1-t)) + \frac{t}{1!} f^{(1)}(\varphi(1-t))(x_1 - x_0) + \cdots + \frac{t^n}{n!} f^{(n)}(\varphi(1-t)) \underbrace{(x_1 - x_0) \cdots (x_1 - x_0)}_{n\text{-times}} + R_n t^{n+1} \right); \quad \forall t \in I$$

By Propositions 3.12, 3.32, 9.7, 7.23, and 7.65, F is continuous. Clearly, φ is differentiable. By Chain Rule and Propositions 9.10, 9.15–9.17 and 9.19, F is differentiable at t , $\forall t \in I^\circ$. Clearly, $F(0) = F(1) = 0$. By Mean Value Theorem 9.20, $\exists t_0 \in I^\circ$ such that $0 = F(1) - F(0) = DF(t_0)$. Then, we have

$$\begin{aligned} 0 &= - \left(-f^{(1)}(\varphi(1-t_0))(x_1 - x_0) + f^{(1)}(\varphi(1-t_0))(x_1 - x_0) \right. \\ &\quad \left. - \frac{t_0}{1!} f^{(2)}(\varphi(1-t_0))(x_1 - x_0)(x_1 - x_0) + \cdots \right. \\ &\quad \left. + \frac{t_0^{n-1}}{(n-1)!} f^{(n)}(\varphi(1-t_0)) \underbrace{(x_1 - x_0) \cdots (x_1 - x_0)}_{n\text{-times}} \right) \end{aligned}$$

$$\begin{aligned}
& -\frac{t_0^n}{n!} f^{(n+1)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(n+1)\text{-times}} + (n+1)R_n t_0^n \Big) \\
& = \frac{t_0^n}{n!} f^{(n+1)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(n+1)\text{-times}} - (n+1)R_n t_0^n
\end{aligned}$$

Hence,

$$R_n = \frac{1}{(n+1)!} f^{(n+1)}(\varphi(\bar{t}_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(n+1)\text{-times}}$$

where $\bar{t}_0 = 1 - t_0 \in I^\circ$.

(ii) By Proposition 7.85, $\exists y_* \in \mathcal{Y}^*$ with $\|y_*\| \leq 1$ such that $\|R_n\| = \langle\langle y_*, R_n \rangle\rangle$. Define $G : \bar{D} \rightarrow \mathbb{K}$ by

$$\begin{aligned}
G(t) &= \langle\langle y_*, f(x_1) - f(\varphi(1-t)) \\
&\quad - \sum_{i=1}^n \frac{t^i}{i!} f^{(i)}(\varphi(1-t)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{i\text{-times}} \rangle\rangle - \|R_n\| t^{n+1}
\end{aligned}$$

$\forall t \in \bar{D}$. By Propositions 3.12, 3.32, 9.7, 7.23, and 7.65, G is continuous. Clearly, φ is differentiable. By Chain Rule and Propositions 9.10, 9.15–9.17 and 9.19, G is differentiable at t , $\forall t \in \bar{D}$. Clearly, $G(0) = G(1) = 0$.

We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathbb{K} = \mathbb{R}$; Case 2: $\mathbb{K} = \mathbb{C}$.

Case 1: $\mathbb{K} = \mathbb{R}$. By Mean Value Theorem 9.20, $\exists t_0 \in I^\circ$ such that $0 = G(1) - G(0) = DG(t_0)$. Then, we have

$$\begin{aligned}
0 &= \langle\langle y_*, f^{(1)}(\varphi(1-t_0))(x_1-x_0) \\
&\quad - \sum_{i=1}^n \frac{t_0^{i-1}}{(i-1)!} f^{(i)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{i\text{-times}} \\
&\quad + \sum_{i=1}^n \frac{t_0^i}{i!} f^{(i+1)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(i+1)\text{-times}} \rangle\rangle \\
&\quad - (n+1) \|R_n\| t_0^n \\
&= \langle\langle y_*, \frac{t_0^n}{n!} f^{(n+1)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(n+1)\text{-times}} \rangle\rangle \\
&\quad - (n+1) \|R_n\| t_0^n
\end{aligned}$$

Then, we have

$$\|R_n\| = \frac{1}{(n+1)!} \langle\langle y_*, f^{(n+1)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(n+1)\text{-times}} \rangle\rangle$$

$$\begin{aligned}
&\leq \frac{1}{(n+1)!} \|y_*\| \|f^{(n+1)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(n+1)\text{-times}}\| \\
&\leq \frac{1}{(n+1)!} \|f^{(n+1)}(\varphi(\bar{t}_0))\| \|x_1-x_0\|^{n+1}
\end{aligned}$$

where the last two inequalities follows from Proposition 7.64 and $\bar{t}_0 = 1-t_0 \in I^\circ$.

Case 2: $\mathbb{K} = \mathbb{C}$. By Lemma 9.22, $\exists t_0 \in I^\circ$ such that $\operatorname{Re}(G(1)-G(0)) = \operatorname{Re}(DG(t_0))$. Then, we have

$$\begin{aligned}
0 &= \operatorname{Re}(\langle\langle y_*, f^{(1)}(\varphi(1-t_0))(x_1-x_0) \rangle\rangle) \\
&\quad - \operatorname{Re}(\langle\langle y_*, \sum_{i=1}^n \frac{t_0^{i-1}}{(i-1)!} f^{(i)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{i\text{-times}} \rangle\rangle) \\
&\quad + \operatorname{Re}(\langle\langle y_*, \sum_{i=1}^n \frac{t_0^i}{i!} f^{(i+1)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(i+1)\text{-times}} \rangle\rangle) \\
&\quad - (n+1) \|R_n\| t_0^n \\
&= \operatorname{Re}(\langle\langle y_*, \frac{t_0^n}{n!} f^{(n+1)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(n+1)\text{-times}} \rangle\rangle) \\
&\quad - (n+1) \|R_n\| t_0^n
\end{aligned}$$

Hence,

$$\begin{aligned}
\|R_n\| &= \frac{1}{(n+1)!} \operatorname{Re}(\langle\langle y_*, f^{(n+1)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(n+1)\text{-times}} \rangle\rangle) \\
&\leq \frac{1}{(n+1)!} \|y_*\| \|f^{(n+1)}(\varphi(1-t_0)) \underbrace{(x_1-x_0) \cdots (x_1-x_0)}_{(n+1)\text{-times}}\| \\
&\leq \frac{1}{(n+1)!} \|f^{(n+1)}(\varphi(\bar{t}_0))\| \|x_1-x_0\|^{n+1}
\end{aligned}$$

where the last two inequalities follows from Proposition 7.64 and $\bar{t}_0 = 1-t_0 \in I^\circ$.

This completes the proof of the theorem. $\square$

9.5 Mapping Theorems

Definition 9.49 Let $\mathcal{X} := (X, \rho)$ be a metric space, $S \subseteq X$, and $T : S \rightarrow X$. T is said to be a contraction mapping on S if $T(S) \subseteq S$ and $\exists \alpha \in [0, 1) \subset \mathbb{R}$ such that $\forall x_1, x_2 \in S$, we have $\rho(T(x_1), T(x_2)) \leq \alpha \rho(x_1, x_2)$. Then, α is called an contraction index for T .

Theorem 9.50 (Contraction Mapping Theorem) *Let $S \neq \emptyset$ be a closed subset of a complete metric space $\mathcal{X} := (X, \rho)$ and T be a contraction mapping on S with contraction index $\alpha \in [0, 1) \subset \mathbb{R}$. Then, the following statements hold.*

- (i) $\exists! x_0 \in S$ such that $x_0 = T(x_0)$.
- (ii) $\forall x_1 \in S$, recursively define $x_{n+1} = T(x_n)$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} x_n = x_0$.
- (iii) $\rho(x_n, x_0) \leq \frac{\alpha^{n-1}}{1-\alpha} \rho(x_2, x_1)$, $\rho(x_n, x_0) \leq \frac{\alpha}{1-\alpha} \rho(x_n, x_{n-1})$, and $\rho(x_n, x_0) \leq \alpha \rho(x_{n-1}, x_0)$, $\forall n \in \{2, 3, \dots\}$.

Proof Fix any $x_1 \in S \neq \emptyset$. Recursively define $x_{n+1} = T(x_n) \in S$, $\forall n \in \mathbb{N}$. Then, $\rho(x_n, x_{n-1}) = \rho(T(x_{n-1}), T(x_{n-2})) \leq \alpha \rho(x_{n-1}, x_{n-2}) \leq \dots \leq \alpha^{n-2} \rho(x_2, x_1)$, $\forall n \in \{3, 4, \dots\}$. Therefore, $(x_n)_{n=1}^\infty \subseteq S$ is a Cauchy sequence and converges to $x_0 \in S$ by $\mathcal{X}$ being complete, S being closed, and Proposition 4.39. Clearly, T is continuous. Then, by Proposition 3.66, $x_0 = T(x_0)$. $\forall \bar{x} \in S$ such that $\bar{x} = T(\bar{x})$. Then, $\rho(x_0, \bar{x}) \leq \alpha \rho(x_0, \bar{x})$ and $\rho(x_0, \bar{x}) = 0$. Hence, $\bar{x} = x_0$. Hence the statements (i) and (ii) are true.

Note that, by Propositions 3.66 and 4.30, $\forall n \in \{2, 3, \dots\}$,

$$\begin{aligned} \rho(x_n, x_0) &= \lim_{m \in \mathbb{N}} \rho(x_n, x_m) \leq \sum_{i=n}^{\infty} \rho(x_{i+1}, x_i) \leq \sum_{i=n+1}^{\infty} \alpha^{i-2} \rho(x_2, x_1) \\ &= \frac{\alpha^{n-1}}{1-\alpha} \rho(x_2, x_1) \\ \rho(x_n, x_0) &\leq \sum_{i=n}^{\infty} \rho(x_{i+1}, x_i) \leq \sum_{i=1}^{\infty} \alpha^i \rho(x_n, x_{n-1}) = \frac{\alpha}{1-\alpha} \rho(x_n, x_{n-1}) \\ \rho(x_n, x_0) &= \rho(T(x_{n-1}), T(x_0)) \leq \alpha \rho(x_{n-1}, x_0) \end{aligned}$$

This completes the proof of the theorem. $\square$

Lemma 9.51 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $x_0 \in D$, $f : D \rightarrow \mathcal{Y}$ be $\mathcal{C}_1$ at x_0 . Assume that D is locally convex at x_0 . Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that, $\forall x_1, x_2 \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, we have $\|f(x_1) - f(x_2) - f^{(1)}(x_0)(x_1 - x_2)\| \leq \epsilon \|x_1 - x_2\|$.*

Proof By the assumption, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) =: \tilde{D}$ is convex. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by f being $\mathcal{C}_1$ at x_0 , $\exists \delta \in (0, \delta_0] \subset \mathbb{R}$ such that $\forall x \in D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) =: \bar{D}$, we have $\|f^{(1)}(x) - f^{(1)}(x_0)\| < \epsilon$. Define $\gamma : D \rightarrow \mathcal{Y}$ by $\gamma(x) = f(x) - f^{(1)}(x_0)(x - x_0)$, $\forall x \in D$. By Propositions 9.45, 9.38, 9.41, and 9.44, γ is $\mathcal{C}_1$ at x_0 and $\gamma^{(1)}(x) = f^{(1)}(x) - f^{(1)}(x_0)$, $\forall x \in \bar{D}$. Clearly, $\bar{D}$ is convex. Then, $\forall x_1, x_2 \in \bar{D}$, by Mean Value Theorem 9.23 and Proposition 7.64, we have $\|f(x_1) - f(x_2) - f^{(1)}(x_0)(x_1 - x_2)\| = \|\gamma(x_1) - \gamma(x_2)\| \leq \|\gamma^{(1)}(t_0 x_1 + (1 - t_0)x_2)(x_1 - x_2)\| \leq \epsilon \|x_1 - x_2\|$, for some $t_0 \in (0, 1) \subset \mathbb{R}$. This completes the proof of the lemma. $\square$

Theorem 9.52 (Injective Mapping Theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $F : D \rightarrow \mathcal{Y}$ be $\mathcal{C}_1$ at $x_0 \in D^\circ$. Assume that $F^{(1)}(x_0) \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is injective and $M := \mathcal{R}(F^{(1)}(x_0)) \subseteq \mathcal{Y}$ is closed. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ with $U := \mathcal{B}_{\mathcal{X}}(x_0, \delta) \subseteq D$ such that $F|_U : U \rightarrow F(U)$ is bijective and admits a continuous inverse $F_i : F(U) \rightarrow U$.*

Proof By Proposition 7.13 and Proposition 4.39, $M \subseteq \mathcal{Y}$ is a Banach space. Then, $F^{(1)}(x_0) : \mathcal{X} \rightarrow M$ is bijective. By Open Mapping Theorem 7.103, the inverse A of $F^{(1)}(x_0) : \mathcal{X} \rightarrow M$ belongs to $\mathcal{B}(M, \mathcal{X})$. $\forall h \in \mathcal{X}$, we have $\|h\| = \|AF^{(1)}(x_0)h\| \leq \|A\| \|F^{(1)}(x_0)h\|$. Then, $\exists r \in (0, \infty) \subset \mathbb{R}$ such that $r\|A\| \leq 1$ and $r\|h\| \leq \|F^{(1)}(x_0)h\|$, $\forall h \in \mathcal{X}$.

By Lemma 9.51, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ with $U := \mathcal{B}_{\mathcal{X}}(x_0, \delta) \subseteq D$ such that $\forall x_1, x_2 \in U$, we have $\|F(x_1) - F(x_2) - F^{(1)}(x_0)(x_1 - x_2)\| \leq r\|x_1 - x_2\|/2$. $\forall x_1, x_2 \in U$, $r\|x_1 - x_2\| \leq \|F^{(1)}(x_0)(x_1 - x_2)\| \leq \|F(x_1) - F(x_2)\| + \|F(x_1) - F(x_2) - F^{(1)}(x_0)(x_1 - x_2)\| \leq \|F(x_1) - F(x_2)\| + r\|x_1 - x_2\|/2$. This implies that $\|F(x_1) - F(x_2)\| \geq r\|x_1 - x_2\|/2$. Hence, $F|_U : U \rightarrow F(U)$ is injective and surjective. Then, $F|_U : U \rightarrow F(U)$ is bijective and admits an inverse $F_i : F(U) \rightarrow U$. $\forall y_1, y_2 \in F(U)$, we have $\|F_i(y_1) - F_i(y_2)\| \leq 2/r\|F(F_i(y_1)) - F(F_i(y_2))\| = 2/r\|y_1 - y_2\|$. Hence, F_i is uniformly continuous. This completes the proof of the theorem $\square$

Theorem 9.53 (Surjective Mapping Theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $F : D \rightarrow \mathcal{Y}$, $x_0 \in D^\circ$, and $y_0 := F(x_0) \in \mathcal{Y}$. Assume that F is $\mathcal{C}_1$ at x_0 , and $F^{(1)}(x_0) \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is surjective. Then, $\exists r \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, and $\exists c_1 \in [0, \infty) \subset \mathbb{R}$ with $c_1\delta \leq r$ such that $\forall \bar{y} \in \mathcal{B}_{\mathcal{Y}}(y_0, \delta/2)$, $\forall \bar{x} \in \mathcal{B}_{\mathcal{X}}(x_0, r/2)$ with $\bar{y} = F(\bar{x})$, $\forall y \in \mathcal{B}_{\mathcal{Y}}(\bar{y}, \delta/2)$, $\exists x \in \mathcal{B}_{\mathcal{X}}(x_0, r) \subseteq D$ with $\|x - \bar{x}\| \leq c_1\|y - \bar{y}\|$, we have $y = F(x)$.*

Proof Let $M := \mathcal{N}(F^{(1)}(x_0))$, which is a closed subspace by Proposition 7.68. By Proposition 7.45, the quotient space $\mathcal{X}/M$ is a Banach space. By Proposition 7.70, $F^{(1)}(x_0) = A \circ \phi$, where $\phi : \mathcal{X} \rightarrow \mathcal{X}/M$ is the natural homomorphism, and $A \in \mathcal{B}(\mathcal{X}/M, \mathcal{Y})$ is injective. Since $F^{(1)}(x_0)$ is surjective, then A is bijective and, by Open Mapping Theorem 7.103, $A^{-1} \in \mathcal{B}(\mathcal{Y}, \mathcal{X}/M)$. Let $c_1 := 4\|A^{-1}\| \in [0, \infty) \subset \mathbb{R}$.

Define $\gamma : D \rightarrow \mathcal{X}/M$ by $\gamma(x) = A^{-1}(F(x) - F^{(1)}(x_0)(x - x_0))$, $\forall x \in D$. By Propositions 9.45, 9.38, 9.41, 9.34, and 9.44, γ is $\mathcal{C}_1$ at x_0 and $\gamma^{(1)}(x) = A^{-1}(F^{(1)}(x) - F^{(1)}(x_0))$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, r_0) \subseteq D$, for some $r_0 \in (0, \infty) \subset \mathbb{R}$. Clearly, $\gamma^{(1)}(x_0) = \vartheta_{\mathcal{B}(\mathcal{X}, \mathcal{X}/M)}$. Then, by Lemma 9.51, $\exists r \in (0, r_0] \subset \mathbb{R}$ such that $\forall x_1, x_2 \in \mathcal{B}_{\mathcal{X}}(x_0, r) \subseteq D$, we have $\|\gamma(x_1) - \gamma(x_2)\| \leq \|x_1 - x_2\|/4$.

Let $\delta \in (0, \infty) \subset \mathbb{R}$ be such that $c_1\delta \leq r$. Fix any $\bar{y} \in \mathcal{B}_{\mathcal{Y}}(y_0, \delta/2)$ and any $\bar{x} \in \mathcal{B}_{\mathcal{X}}(x_0, r/2)$ with $\bar{y} = F(\bar{x})$. Fix any $y \in \mathcal{B}_{\mathcal{Y}}(\bar{y}, \delta/2)$. Recursively define $x_1 := \bar{x}$, $[x_{k+1}] = [x_k] + A^{-1}(y - F(x_k))$ and select $x_{k+1} \in [x_{k+1}]$ such that $\|x_{k+1} - x_k\| \leq 2\|[x_{k+1}] - [x_k]\|$, $\forall k \in \mathbb{N}$, where $[x_k] = \phi(x_k) = x_k + M$ is the coset containing x_k . Clearly, $x_1 \in \mathcal{B}_{\mathcal{X}}(x_0, r)$. Note that

$\|x_2 - x_1\| \leq 2\|[x_2] - [x_1]\| = 2\|A^{-1}(y - \bar{y})\| \leq 2\|A^{-1}\|\|y - \bar{y}\| = c_1\|y - \bar{y}\|/2 < r/4$, where the second inequality follows from Proposition 7.64. Then, $x_2 \in \mathcal{B}_{\mathcal{X}}(x_0, r)$. Assume that $x_1, \dots, x_k \in \mathcal{B}_{\mathcal{X}}(x_0, r)$ for some $k \in \{2, 3, \dots\}$. Note that $\forall i \in \{2, \dots, k\}$, $\|x_{i+1} - x_i\| \leq 2\|[x_{i+1}] - [x_i]\| = 2\|A^{-1}(y - F(x_i) + A\phi(x_i)) - A^{-1}(y - F(x_{i-1}) + A\phi(x_{i-1}))\| = 2\|\gamma(x_i) - \gamma(x_{i-1})\| \leq \|x_i - x_{i-1}\|/2$. Then, $\|x_{i+1} - x_i\| \leq \|x_2 - x_1\|/2^{i-1}$. Hence, $\|x_{k+1} - \bar{x}\| \leq \sum_{i=1}^k \|x_{i+1} - x_i\| \leq \sum_{i=1}^k \|x_2 - x_1\|/2^{i-1} = (2 - 1/2^{k-1})\|x_2 - x_1\| < r/2$. This implies that $x_{k+1} \in \mathcal{B}_{\mathcal{X}}(x_0, r)$. Inductively, we have $(x_k)_{k=0}^{\infty} \subseteq \mathcal{B}_{\mathcal{X}}(x_0, r)$ and $\|x_{k+1} - x_k\| \leq \|x_2 - x_1\|/2^{k-1}$, $\forall k \in \mathbb{N}$. Hence, $(x_k)_{k=0}^{\infty}$ is a Cauchy sequence. By the completeness of $\mathcal{X}$, we have $\lim_{k \in \mathbb{N}} x_k = x \in \mathcal{X}$. Note that, by Propositions 3.66 and 4.30, $\|x - \bar{x}\| = \lim_{k \in \mathbb{N}} \|x_k - x_1\| \leq \lim_{k \in \mathbb{N}} (2 - 1/2^{k-2})\|x_2 - x_1\| = 2\|x_2 - x_1\| < r/2$. Hence, $x \in \mathcal{B}_{\mathcal{X}}(x_0, r)$. By the differentiability of F and Propositions 9.7, 3.66, and 7.69, we have $\phi(x) = \lim_{k \in \mathbb{N}} \phi(x_{k+1}) = \lim_{k \in \mathbb{N}} (A^{-1}(y - F(x_k)) + \phi(x_k)) = \phi(x) + A^{-1}(y - F(x))$. This implies that $y = F(x)$. Note that $\|x - \bar{x}\| \leq 2\|x_2 - x_1\| \leq c_1\|y - \bar{y}\|$. This completes the proof of the theorem. $\square$

Theorem 9.54 (Open Mapping Theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$ be open, and $F : D \rightarrow \mathcal{Y}$ be $\mathcal{C}_1$. Assume that $F^{(1)}(x) \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is surjective, $\forall x \in D$. Then, F is an open mapping.*

Proof Fix any open subset $U \subseteq D$, where U is open in the subset topology of D . Since D is open, then U is open in $\mathcal{X}$. We will show that $F(U)$ is an open set in $\mathcal{Y}$. Fix any $y_0 \in F(U)$, there exists $x_0 \in U$ such that $y_0 = F(x_0)$. Then, $\exists r \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{X}}(x_0, r) \subseteq U$. It is easy to check that all assumptions of Surjective Mapping Theorem are satisfied at x_0 . Then, there exist an open set $V \subseteq \mathcal{Y}$ with $y_0 \in V$ and $c_1 \in [0, \infty) \subset \mathbb{R}$ such that $\forall \bar{y} \in V$, $\exists \bar{x} \in D$ with $\|\bar{x} - x_0\| \leq c_1\|\bar{y} - y_0\|$, we have $\bar{y} = F(\bar{x})$. Take $\delta \in (0, \infty) \subset \mathbb{R}$ such that $c_1\delta \leq r$ and $\mathcal{B}_{\mathcal{Y}}(y_0, \delta) \subseteq V$. Then, $\forall \bar{y} \in \mathcal{B}_{\mathcal{Y}}(y_0, \delta)$, $\|\bar{x} - x_0\| \leq c_1\|\bar{y} - y_0\| < r$. Then, $\bar{x} \in \mathcal{B}_{\mathcal{X}}(x_0, r) \subseteq U$ and $\bar{y} \in F(U)$. Hence, $\mathcal{B}_{\mathcal{Y}}(y_0, \delta) \subseteq F(U)$. Therefore, $y_0 \in (F(U))^{\circ}$. By the arbitrariness of y_0 , we have $F(U)$ is open in $\mathcal{Y}$. By the arbitrariness of U , F is an open mapping. This completes the proof of the theorem. $\square$

Proposition 9.55 *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$ and $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be bijective. Then, $\forall T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ with $\|T - A\|\|A^{-1}\| < 1$, we have T is bijective and $\|T^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2\|T - A\|}{1 - \|A^{-1}\|\|T - A\|}$.*

Proof By Open Mapping Theorem 7.103, $A^{-1} \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$. We will first prove the result for the special case $\mathcal{Y} = \mathcal{X}$ and $A = \text{id}_{\mathcal{X}}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X}$ is a singleton set; Case 2: $\exists \bar{x} \in \mathcal{X}$ such that $\bar{x} \neq \vartheta_{\mathcal{X}}$.

Case 1: $\mathcal{X}$ is a singleton set. Then, $\|A^{-1}\| = \|A\| = 0$. $\forall T \in \mathcal{B}(\mathcal{X}, \mathcal{X})$, we have $T = \text{id}_{\mathcal{X}}$. Then, T is bijective and $\|T^{-1} - \text{id}_{\mathcal{X}}\| = 0$. The result holds for this case.

Case 2: $\exists \bar{x} \in \mathcal{X}$ such that $\bar{x} \neq \vartheta_{\mathcal{X}}$. Then, $\|A\| = \|A^{-1}\| = 1$. $\forall T \in B(\mathcal{X}, \mathcal{X})$ with $\|T - \text{id}_{\mathcal{X}}\| < 1$, let $\Delta := T - \text{id}_{\mathcal{X}}$. We will show that T is bijective. $\forall x_1, x_2 \in \mathcal{X}$ with $T(x_1) = T(x_2)$, we have $x_1 + \Delta(x_1) = x_2 + \Delta(x_2)$, which implies that $\Delta(x_1 - x_2) = x_2 - x_1$. By Proposition 7.64, we have $\|\Delta\| \|x_1 - x_2\| \geq \|x_1 - x_2\|$. Since $\|\Delta\| < 1$ then $\|x_1 - x_2\| = 0$ and $x_1 = x_2$. Therefore, T is injective. $\forall x_0 \in \mathcal{X}$, define $\phi : \mathcal{X} \rightarrow \mathcal{X}$ by $\phi(x) = x_0 - \Delta(x)$, $\forall x \in \mathcal{X}$. Clearly, ϕ is a contraction mapping on $\mathcal{X}$ with contraction index $\|\Delta\|$. By Contraction Mapping Theorem, there exists a unique $\bar{x} \in \mathcal{X}$ such that $\bar{x} = \phi(\bar{x})$. Then, $x_0 = \bar{x} + \Delta(\bar{x}) = T(\bar{x})$. Hence, T is surjective. Then, T is bijective. By Open Mapping Theorem 7.103, $T^{-1} \in B(\mathcal{X}, \mathcal{X})$.

$\forall y \in \mathcal{Y}$, let $x = T^{-1}y$. Then, $y = Tx = x + \Delta x$ and $x = y - \Delta x$. By Proposition 7.64, we have $\|x\| \leq \|y\| + \|\Delta\| \|x\|$ and $\|x\| \leq \frac{\|y\|}{1 - \|\Delta\|}$. By the arbitrariness of y , we have $\|T^{-1}\| \leq 1/(1 - \|\Delta\|)$. This further implies that $\|T^{-1} - \text{id}_{\mathcal{X}}\| = \|T^{-1}(\text{id}_{\mathcal{X}} - T)\| \leq \|T^{-1}\| \|\Delta\| \leq \|\Delta\|/(1 - \|\Delta\|)$. The result holds in this case.

Hence, the result holds for the special case $\mathcal{Y} = \mathcal{X}$ and $A = \text{id}_{\mathcal{X}}$. Now consider the general case. $\forall T \in B(\mathcal{X}, \mathcal{Y})$ with $\|T - A\| \|A^{-1}\| < 1$, we have $\bar{T} := A^{-1}T \in B(\mathcal{X}, \mathcal{X})$. Note that $\|\bar{T} - \text{id}_{\mathcal{X}}\| = \|A^{-1}(T - A)\| \leq \|A^{-1}\| \|T - A\| < 1$. Then, $\|\bar{T} - \text{id}_{\mathcal{X}}\| \|\text{id}_{\mathcal{X}}\| < 1$. By the special case, we have $\bar{T}$ is bijective and $\|\bar{T}^{-1} - \text{id}_{\mathcal{X}}\| \leq \frac{\|\text{id}_{\mathcal{X}}\|^2 \|\bar{T} - \text{id}_{\mathcal{X}}\|}{1 - \|\text{id}_{\mathcal{X}}\| \|\bar{T} - \text{id}_{\mathcal{X}}\|}$. Then, T is bijective and, by Proposition 7.64,

$$\begin{aligned} \|\bar{T} - \text{id}_{\mathcal{X}}\| &= \|A^{-1}(T - A)\| \leq \|A^{-1}\| \|T - A\| \\ \|A^{-1}\| \|\text{id}_{\mathcal{X}}\| &= \|A^{-1}\| \\ \|T^{-1} - A^{-1}\| &= \|(\bar{T}^{-1} - \text{id}_{\mathcal{X}})A^{-1}\| \leq \|A^{-1}\| \frac{\|\text{id}_{\mathcal{X}}\|^2 \|\bar{T} - \text{id}_{\mathcal{X}}\|}{1 - \|\text{id}_{\mathcal{X}}\| \|\bar{T} - \text{id}_{\mathcal{X}}\|} \\ &\leq \frac{\|A^{-1}\|^2 \|T - A\|}{1 - \|A^{-1}\| \|T - A\|} \end{aligned}$$

This completes the proof of the proposition. $\square$

Proposition 9.56 *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D := \{L \in B(\mathcal{X}, \mathcal{Y}) \mid L \text{ is bijective}\}$, and $f : D \rightarrow B(\mathcal{Y}, \mathcal{X})$ be given by $f(A) = A^{-1}$, $\forall A \in D$. Then, D is open in $B(\mathcal{X}, \mathcal{Y})$, f is $\mathcal{C}_{\infty}$, and $f^{(1)}(A)(\Delta) = -A^{-1}\Delta A^{-1}$, $\forall A \in D$, $\forall \Delta \in B(\mathcal{X}, \mathcal{Y})$.*

Proof By Proposition 9.55 and Open Mapping Theorem 7.103, D is open and f is continuous. $\forall A \in D$, $\text{span}(A_D(A)) = B(\mathcal{X}, \mathcal{Y})$ since D is open and $A \in D^{\circ}$. Define $L : B(\mathcal{X}, \mathcal{Y}) \rightarrow B(\mathcal{Y}, \mathcal{X})$ by $L(\Delta) = -A^{-1}\Delta A^{-1}$, $\forall \Delta \in B(\mathcal{X}, \mathcal{Y})$. Clearly, L is a linear operator. Note that

$$\|L\| = \sup_{\Delta \in B(\mathcal{X}, \mathcal{Y}), \|\Delta\| \leq 1} \|L(\Delta)\| \leq \|A^{-1}\|^2$$

where the inequality follows from Proposition 7.64. Hence, L is a bounded linear operator.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the continuity of f , $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\|f(\bar{A}) - f(A)\| < \epsilon$, $\forall \bar{A} \in \mathcal{B}_{\mathcal{B}(\mathcal{X}, \mathcal{Y})}(A, \delta)$. Then, $\forall \bar{A} \in \mathcal{B}_{\mathcal{B}(\mathcal{X}, \mathcal{Y})}(A, \delta)$, we have

$$\begin{aligned} \|f(\bar{A}) - f(A) - L(\bar{A} - A)\| &= \|\bar{A}^{-1} - A^{-1} + A^{-1}(\bar{A} - A)A^{-1}\| \\ &= \|-A^{-1}(\bar{A} - A)\bar{A}^{-1} + A^{-1}(\bar{A} - A)A^{-1}\| \\ &\leq \|A^{-1}\| \|\bar{A} - A\| \|f(\bar{A}) - f(A)\| \leq \epsilon \|A^{-1}\| \|\bar{A} - A\| \end{aligned}$$

where the first inequality follows from Proposition 7.64. Hence, we have $f^{(1)}(A) = L$. Then, f is differentiable. Note that $f^{(1)}(A) = -f(A)\text{ro}(f(A))$, $\forall A \in D$. By Propositions 9.42, 9.7, 3.12, and 3.32, $f^{(1)}$ is continuous. Hence, f is $\mathcal{C}_1$.

Assume that f is $\mathcal{C}_k$, for some $k \in \mathbb{N}$. We will show that f is $\mathcal{C}_{k+1}$. Then, f is $\mathcal{C}_\infty$. Note that $f^{(1)}(A) = -f(A)\text{ro}(f(A))$, $\forall A \in D$. By Propositions 9.45, 9.42, and 9.44, $f^{(1)}$ is $\mathcal{C}_k$. Then, f is $\mathcal{C}_{k+1}$. This completes the proof of the proposition. $\square$

Theorem 9.57 (Inverse Function Theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $F : D \rightarrow \mathcal{Y}$ be $\mathcal{C}_1$ at $x_0 \in D^\circ$. Assume that $F^{(1)}(x_0) \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is bijective. Then, $\exists$ open set $U \subseteq D$ with $x_0 \in U$ and $\exists$ open set $V \subseteq \mathcal{Y}$ with $y_0 := F(x_0) \in V$ such that*

- (i) $F|_U : U \rightarrow V$ is bijective;
- (ii) the inverse mapping $F_i : V \rightarrow U$ of $F|_U$ is differentiable, $F_i^{(1)} : V \rightarrow \mathcal{B}(\mathcal{Y}, \mathcal{X})$ is given by $F_i^{(1)}(y) = (F^{(1)}(F_i(y)))^{-1}$, $\forall y \in V$, and $F_i^{(1)}$ is continuous at y_0 ;
- (iii) if F is k -times differentiable for some $k \in \mathbb{N}$, then F_i is k -times differentiable;
- (iv) if F is $\mathcal{C}_k$ at x_0 for some $k \in \mathbb{N} \cup \{\infty\}$, then F_i is $\mathcal{C}_k$ at y_0 .

Proof By Open Mapping Theorem 7.103, $(F^{(1)}(x_0))^{-1} \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$. By F being $\mathcal{C}_1$ at $x_0 \in D^\circ$, $\exists \bar{r} \in (0, \infty) \subset \mathbb{R}$ such that $F^{(1)}(x)$ exists, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \bar{r}) \subseteq D$. Define $T : \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, \bar{r}) \rightarrow \mathcal{X}$ by $T(x) = (F^{(1)}(x_0))^{-1}(F(x + x_0) - y_0)$, $\forall x \in \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, \bar{r}) \subseteq D - x_0$. Clearly, $T(\vartheta_{\mathcal{X}}) = \vartheta_{\mathcal{X}}$. By Propositions 9.45, 9.38, 9.34, and 9.44, T is $\mathcal{C}_1$ at $\vartheta_{\mathcal{X}}$, T is differentiable, and $T^{(1)}(x) = (F^{(1)}(x_0))^{-1}F^{(1)}(x + x_0)$, $\forall x \in \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, \bar{r})$. Clearly, $T^{(1)}(\vartheta_{\mathcal{X}}) = \text{id}_{\mathcal{X}}$.

Define $\psi : \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, \bar{r}) \rightarrow \mathcal{X}$ by $\psi(x) = T(x) - x$. Then, by Propositions 9.38, 9.45, and 9.44, ψ is differentiable, ψ is $\mathcal{C}_1$ at $\vartheta_{\mathcal{X}}$, and $\psi^{(1)}(x) = T^{(1)}(x) - \text{id}_{\mathcal{X}}$, $\forall x \in \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, \bar{r})$. Clearly, $\psi(\vartheta_{\mathcal{X}}) = \vartheta_{\mathcal{X}}$ and $\psi^{(1)}(\vartheta_{\mathcal{X}}) = \vartheta_{\mathcal{B}(\mathcal{X}, \mathcal{X})}$. Fix any $\alpha \in (0, 1) \subset \mathbb{R}$. Then, $\exists r_1 \in (0, \bar{r}) \subset \mathbb{R}$

such that $\overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1) \subseteq D - x_0$ and $\|\psi^{(1)}(x)\| \leq \alpha$, $\forall x \in \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1)$. $\forall x_1, x_2 \in \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1)$, by Mean Value Theorem, $\|\psi(x_1) - \psi(x_2)\| \leq \sup_{t_0 \in (0,1) \subset \mathbb{R}} \|\psi^{(1)}(t_0 x_1 + (1-t_0)x_2)(x_1 - x_2)\| \leq \alpha \|x_1 - x_2\|$, where the last inequality follows from Proposition 7.64.

$\forall \bar{x} \in \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, (1-\alpha)r_1)$, define $\phi : \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1) \rightarrow \mathcal{X}$ by $\phi(x) = \bar{x} - \psi(x)$, $\forall x \in \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1)$. $\forall x \in \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1)$, $\|\phi(x)\| \leq \|\bar{x}\| + \|\psi(x)\| < (1-\alpha)r_1 + \|\psi(x) - \psi(\vartheta_{\mathcal{X}})\| \leq (1-\alpha)r_1 + \alpha\|x\| \leq r_1$. Hence, $\phi : \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1) \rightarrow \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1) \subseteq \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1)$. It is easy to see that ϕ is a contraction mapping with contraction index α . By Contraction Mapping Theorem, $\exists! \hat{x} \in \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1)$ such that $\hat{x} = \phi(\hat{x}) \in \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1)$, which is equivalent to $\bar{x} = T(\hat{x})$. Let $\bar{U} := T_{\text{inv}}(\mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, (1-\alpha)r_1)) \cap \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1)$ and $\bar{V} := \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, (1-\alpha)r_1)$. Note that $\bar{U}$ and $\bar{V}$ are open sets in $\mathcal{X}$ since T is continuous by Proposition 9.7. Then, $T|_{\bar{U}} : \bar{U} \rightarrow \bar{V}$ is bijective. Since $T(\vartheta_{\mathcal{X}}) = \vartheta_{\mathcal{X}}$, then $\vartheta_{\mathcal{X}} \in \bar{U}$.

Hence, there exists an inverse mapping $T_i : \bar{V} \rightarrow \bar{U}$ of $T|_{\bar{U}}$. Let $U := \bar{U} + x_0$ and $V := F^{(1)}(x_0)(\bar{V}) + y_0$. Clearly, U and V are open sets in $\mathcal{X}$ and $\mathcal{Y}$, respectively. Note that $F|_U(x) = F^{(1)}(x_0) T|_{\bar{U}}(x - x_0) + y_0$, $\forall x \in U$. Then, $F|_U : U \rightarrow V$ is bijective, whose inverse function is $F_i : V \rightarrow U$. The inverse function F_i is given by $F_i(y) = T_i((F^{(1)}(x_0))^{-1}(y - y_0)) + x_0$, $\forall y \in V$. Hence, the statement (i) holds.

Next, we will show that $F_i : V \rightarrow U$ is differentiable. Note that $\forall x \in \bar{V}$, $x = T(T_i(x)) = T_i(x) + \psi(T_i(x))$ and $T_i(x) = x - \psi(T_i(x))$. $\forall x_1, x_2 \in \bar{V}$, we have $T_i(x_1), T_i(x_2) \in \bar{U} \subseteq \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1)$ and

$$\begin{aligned} \|T_i(x_1) - T_i(x_2)\| &= \|x_1 - x_2 - \psi(T_i(x_1)) + \psi(T_i(x_2))\| \\ &\leq \|x_1 - x_2\| + \|\psi(T_i(x_1)) - \psi(T_i(x_2))\| \\ &\leq \|x_1 - x_2\| + \alpha \|T_i(x_1) - T_i(x_2)\| \end{aligned}$$

which further implies that $\|T_i(x_1) - T_i(x_2)\| \leq \|x_1 - x_2\|/(1-\alpha)$. Therefore, T_i is continuous. By Propositions 3.12, 7.23, and 3.32, F_i is continuous. We need the following intermediate result.

Claim 9.57.1 $\forall x \in U$, let $y = F(x)$. Then, F_i is differentiable at $F(x)$ and $F_i^{(1)}(y) = (F^{(1)}(x))^{-1}$.

Proof of claim: Fix any $x \in U$ and let $y = F(x)$. Then, $x - x_0 \in \bar{U} \subseteq \mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r_1)$ and $\|\psi^{(1)}(x - x_0)\| \leq \alpha$. Note that $T = \text{id}_{\mathcal{X}} + \psi$ and $T^{(1)}(x - x_0) = \text{id}_{\mathcal{X}} + \psi^{(1)}(x - x_0)$. By Proposition 9.55, $T^{(1)}(x - x_0)$ is bijective and $\|(T^{(1)}(x - x_0))^{-1}\| < \infty$. Note that $F^{(1)}(x) = F^{(1)}(x_0)T^{(1)}(x - x_0)$. Then, $F^{(1)}(x)$ is bijective and $c_1 := \|(F^{(1)}(x))^{-1}\| < \infty$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$ with $\epsilon c_1 < 1$, by the differentiability of F at x , $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $\forall h \in \mathcal{X}$ with $\|h\| < \delta_1$, we have $\|F(x + h) - F(x) - F^{(1)}(x)(h)\| \leq \epsilon \|h\|$. By the continuity of F_i , $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\forall u \in \mathcal{Y}$ with $\|u\| < \delta$, we have $y + u \in V$ and $\|F_i(y + u) - F_i(y)\| < \delta_1$.

$\forall u \in \mathcal{Y}$ with $\|u\| < \delta$, let $\beta := \|F_i(y+u) - F_i(y) - (F^{(1)}(x))^{-1}u\| \geq 0$. Let $h := F_i(y+u) - x \in \mathcal{X}$. Then, $\|h\| = \|F_i(y+u) - F_i(y)\| < \delta_1$. Note that

$$\begin{aligned} \beta &= \|x + h - x - (F^{(1)}(x))^{-1}u\| = \|(F^{(1)}(x))^{-1}(u - F^{(1)}(x)h)\| \\ &\leq \|(F^{(1)}(x))^{-1}\| \|u - F^{(1)}(x)h\| \\ &= c_1 \|F(x+h) - F(x) - F^{(1)}(x)h\| \\ &\leq c_1 \epsilon \|h\| = \epsilon c_1 \|F_i(y+u) - F_i(y)\| \\ &\leq \epsilon c_1 (\|F_i(y+u) - F_i(y) - (F^{(1)}(x))^{-1}u\| + \|(F^{(1)}(x))^{-1}u\|) \\ &\leq \epsilon c_1 (\beta + c_1 \|u\|) \end{aligned}$$

Then, $\beta \leq \frac{\epsilon c_1^2}{1 - \epsilon c_1} \|u\|$. Hence, F_i is differentiable at y and $F_i^{(1)}(y) = (F^{(1)}(x))^{-1}$. This completes the proof of the claim. $\square$

Then, $\forall y \in V$, $F_i^{(1)}(y) = (F^{(1)}(F_i(y)))^{-1}$. By Propositions 9.56 and 3.12, the continuity of F_i , and the continuity of $F^{(1)}$ at x_0 , we have $F_i^{(1)}$ is continuous at y_0 . Then, the statement (ii) holds.

For (iii), we will use mathematical induction on k .

1° $k = 1$. The result holds by (ii).

2° Assume that the result holds for $k = \bar{k} \in \mathbb{N}$.

3° Consider the case $k = \bar{k} + 1$. By inductive assumption, F_i is $\bar{k}$ -times differentiable. Clearly, $F^{(1)}$ is $\bar{k}$ -times differentiable. By (ii) and Propositions 9.45 and 9.56, $F_i^{(1)}$ is $\bar{k}$ -times differentiable. Then, F_i is $(\bar{k}+1)$ -times differentiable. This completes the induction process and the proof of the statement (iii).

For (iv), we will use mathematical induction on k to show that the result holds when $k \in \mathbb{N}$.

1° $k = 1$. The result holds by (ii).

2° Assume that the result holds for $k = \bar{k} \in \mathbb{N}$.

3° Consider the case $k = \bar{k} + 1$. Clearly, $F^{(1)}$ is $\mathcal{C}_{\bar{k}}$ at x_0 . By the inductive assumption, F_i is $\mathcal{C}_{\bar{k}}$ at y_0 . By (ii) and Propositions 9.45 and 9.56, $F_i^{(1)}$ is $\mathcal{C}_{\bar{k}}$ at y_0 . Hence, F_i is $\mathcal{C}_{\bar{k}+1}$ at y_0 . This completes the induction process.

When $k = \infty$. $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that F is $\mathcal{C}_i$ at x , $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta) \subseteq D$, $\forall i \in \mathbb{N}$. By taking $r_1 < \delta$ in the proof of (i), we have $U \subseteq \mathcal{B}_{\mathcal{X}}(x_0, \delta)$. Then, by the induction conclusion, F_i is $\mathcal{C}_i$ at y , $\forall y \in F(U) = V$, $\forall i \in \mathbb{N}$. Hence, F_i is $\mathcal{C}_{\infty}$ at y_0 .

This completes the proof of the theorem. $\square$

Theorem 9.58 (Implicit Function Theorem) Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $\mathcal{Y}$ and $\mathcal{Z}$ be Banach spaces over $\mathbb{K}$, $D \subseteq \mathcal{X} \times \mathcal{Y}$, $F : D \rightarrow \mathcal{Z}$ be continuous. Assume that F is partial differentiable with respect to y and $\frac{\partial F}{\partial y}$ is continuous at $(x_0, y_0) \in D^\circ$, $F(x_0, y_0) = \vartheta_{\mathcal{Z}}$, and $\frac{\partial F}{\partial y}(x_0, y_0) \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ is bijective. Then, the following statements hold.

- (i) There exist an open set $U_0 \in \mathcal{O}$ with $x_0 \in U_0$ and $r_1 \in (0, \infty) \subset \mathbb{R}$ such that $U_0 \times \mathcal{B}_{\mathcal{Y}}(y_0, r_1) \subseteq D$ and $\forall x \in U_0$, $\exists! y \in \mathcal{B}_{\mathcal{Y}}(y_0, r_1)$

satisfying $F(x, y) = \vartheta_Z$. This defines a function $\phi : U_0 \rightarrow \mathcal{B}_Y(y_0, r_1)$ by $\phi(x) = y, \forall x \in U_0$. Then, ϕ is continuous.

(ii) $\forall (x, y) \in U_0 \times \mathcal{B}_Y(y_0, r_1)$, $\frac{\partial F}{\partial y}(x, y)$ is bijective and $\left\| \left(\frac{\partial F}{\partial y}(x, y) \right)^{-1} \right\| < +\infty$.

Proof By Open Mapping Theorem 7.103, $\left(\frac{\partial F}{\partial y}(x_0, y_0) \right)^{-1} \in \mathcal{B}(Z, Y)$. Define a mapping $\psi : D \rightarrow Y$ by

$$\psi(x, y) = y - \left(\frac{\partial F}{\partial y}(x_0, y_0) \right)^{-1} F(x, y); \quad \forall (x, y) \in D$$

Note that $\psi(x_0, y_0) = y_0$. Then, by Propositions 7.23, 3.12, 3.27, and 3.32, ψ is continuous. By the partial differentiability of F with respect to y , Chain Rule, and Propositions 9.41, 9.15, and 9.19, ψ is partial differentiable with respect to y and

$$\frac{\partial \psi}{\partial y}(x, y) = \text{id}_Y - \left(\frac{\partial F}{\partial y}(x_0, y_0) \right)^{-1} \frac{\partial F}{\partial y}(x, y); \quad \forall (x, y) \in D$$

By the continuity of $\frac{\partial F}{\partial y}$ at (x_0, y_0) , then $\frac{\partial \psi}{\partial y}$ is continuous at $(x_0, y_0) \in D^\circ$ and $\frac{\partial \psi}{\partial y}(x_0, y_0) = \vartheta_{\mathcal{B}(Y, Y)}$. Fix any $\alpha \in (0, 1) \subset \mathbb{R}$. Then, $\exists U_1 \in \mathcal{O}$ with $x_0 \in U_1$ and $\exists r_1 \in (0, \infty) \subset \mathbb{R}$ such that $U_1 \times \overline{\mathcal{B}}_Y(y_0, r_1) \subseteq D$ and $\left\| \frac{\partial \psi}{\partial y}(x, y) \right\| \leq \alpha, \forall (x, y) \in U_1 \times \overline{\mathcal{B}}_Y(y_0, r_1)$. By the continuity of ψ , $\exists U_0 \in \mathcal{O}$ with $x_0 \in U_0 \subseteq U_1$ such that $\|\psi(x, y_0) - y_0\| = \|\psi(x, y_0) - \psi(x_0, y_0)\| < (1 - \alpha)r_1, \forall x \in U_0$.

Fix any $x \in U_0$, define mapping $\gamma_x : \overline{\mathcal{B}}_Y(y_0, r_1) \rightarrow Y$ by $\gamma_x(y) = \psi(x, y)$, $\forall y \in \overline{\mathcal{B}}_Y(y_0, r_1)$. We will show that γ_x is a contraction mapping with contraction index α . $\forall y \in \overline{\mathcal{B}}_Y(y_0, r_1)$, we have

$$\begin{aligned} \|\gamma_x(y) - y_0\| &= \|\psi(x, y) - \psi(x_0, y_0)\| \\ &\leq \|\psi(x, y) - \psi(x, y_0)\| + \|\psi(x, y_0) - \psi(x_0, y_0)\| \\ &< \sup_{t \in (0, 1) \subset \mathbb{R}} \left\| \frac{\partial \psi}{\partial y}(x, ty + (1 - t)y_0)(y - y_0) \right\| + (1 - \alpha)r_1 \\ &\leq \sup_{t \in (0, 1) \subset \mathbb{R}} \left\| \frac{\partial \psi}{\partial y}(x, ty + (1 - t)y_0) \right\| \|y - y_0\| + (1 - \alpha)r_1 \\ &\leq \alpha r_1 + (1 - \alpha)r_1 = r_1 \end{aligned}$$

where the second inequality follows from Mean Value Theorem and the third inequality follows from Proposition 7.64. Then, $\gamma_x : \overline{\mathcal{B}}_Y(y_0, r_1) \rightarrow \mathcal{B}_Y(y_0, r_1) \subseteq \overline{\mathcal{B}}_Y(y_0, r_1)$. $\forall y_1, y_2 \in \overline{\mathcal{B}}_Y(y_0, r_1)$, we have

$$\|\gamma_x(y_1) - \gamma_x(y_2)\| = \|\psi(x, y_1) - \psi(x, y_2)\|$$

$$\begin{aligned}
&\leq \sup_{t \in (0,1) \subset \mathbb{R}} \left\| \frac{\partial \psi}{\partial y}(x, ty_1 + (1-t)y_2)(y_1 - y_2) \right\| \\
&\leq \sup_{t \in (0,1) \subset \mathbb{R}} \left\| \frac{\partial \psi}{\partial y}(x, ty_1 + (1-t)y_2) \right\| \|y_1 - y_2\| \\
&\leq \alpha \|y_1 - y_2\|
\end{aligned}$$

where the first inequality follows from Mean Value Theorem and the second inequality follows from Proposition 7.64. Hence, γ_x is a contraction mapping with contraction index α .

By Contraction Mapping Theorem, $\exists! y \in \overline{\mathcal{B}}_{\mathcal{Y}}(y_0, r_1)$ such that $y = \gamma_x(y) \in \mathcal{B}_{\mathcal{Y}}(y_0, r_1)$, $y = \lim_{n \in \mathbb{N}} \gamma_{x,n}(y_0)$, where $\gamma_{x,n}(y_0)$ is recursively defined by $\gamma_{x,1}(y_0) = y_0$ and $\gamma_{x,k+1}(y_0) = \gamma_x(\gamma_{x,k}(y_0))$, $\forall k \in \mathbb{N}$, and $\|\gamma_{x,n}(y_0) - y\| \leq \frac{\alpha^{n-1}}{1-\alpha} \|\gamma_x(y_0) - y_0\| < r_1 \alpha^{n-1}$, $\forall n \in \{2, 3, \dots\}$. By $\gamma_x(y) = y$, we conclude that $F(x, y) = \vartheta_{\mathcal{Z}}$. Hence, $\forall x \in U_0$, $\exists! y \in \mathcal{B}_{\mathcal{Y}}(y_0, r_1)$ such that $F(x, y) = \vartheta_{\mathcal{Z}}$, since $F(x, y) = \vartheta_{\mathcal{Z}} \Leftrightarrow y = \gamma_x(y)$. Then, we may define $\phi : U_0 \rightarrow \mathcal{B}_{\mathcal{Y}}(y_0, r_1)$ by $\phi(x) = y = \lim_{n \in \mathbb{N}} \gamma_{x,n}(y_0)$, $\forall x \in U_0$. Hence, $F(x, \phi(x)) = \vartheta_{\mathcal{Z}}$, $\forall x \in U_0$.

Next, we show that ϕ is continuous. Fix any $\bar{x} \in U_0$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ with $n_0 > 1$ such that $\alpha^{n_0-1} r_1 < \epsilon/3$. By the continuity of ψ , we have that $\gamma_{x,n_0}(y_0)$ is continuous with respect to x , that is, $\exists \bar{U} \in \mathcal{O}$ with $\bar{x} \in \bar{U} \subseteq U_0$ such that $\forall x_1 \in \bar{U}$, we have $\|\gamma_{x_1,n_0}(y_0) - \gamma_{\bar{x},n_0}(y_0)\| < \epsilon/3$. Then,

$$\begin{aligned}
\|\phi(x_1) - \phi(\bar{x})\| &\leq \|\phi(x_1) - \gamma_{x_1,n_0}(y_0)\| + \|\gamma_{x_1,n_0}(y_0) - \gamma_{\bar{x},n_0}(y_0)\| \\
&\quad + \|\gamma_{\bar{x},n_0}(y_0) - \phi(\bar{x})\| \leq r_1 \alpha^{n_0-1} + \epsilon/3 + r_1 \alpha^{n_0-1} < \epsilon
\end{aligned}$$

Hence, ϕ is continuous. Thus, the statement (i) is proved.

(ii) Note that, $\forall (x, y) \in U_1 \times \overline{\mathcal{B}}_{\mathcal{Y}}(y_0, r_1)$, we have $\left\| \frac{\partial \psi}{\partial y}(x, y) \right\| \leq \alpha$. By Proposition 9.55, $\text{id}_{\mathcal{Y}} - \frac{\partial \psi}{\partial y}(x, y)$ is bijective with continuous inverse. Note that $\text{id}_{\mathcal{Y}} - \frac{\partial \psi}{\partial y}(x, y) = \left(\frac{\partial F}{\partial y}(x_0, y_0) \right)^{-1} \frac{\partial F}{\partial y}(x, y)$. Therefore, $\frac{\partial F}{\partial y}(x, y)$ is bijective with continuous inverse.

This completes the proof of the theorem. $\square$

Theorem 9.59 (Implicit Function Theorem) *Let $\mathcal{X}$ be normed linear space over $\mathbb{K}$, $\mathcal{Y}$ and $\mathcal{Z}$ be Banach spaces over $\mathbb{K}$, $D \subseteq \mathcal{X} \times \mathcal{Y}$, $F : D \rightarrow \mathcal{Z}$ be continuous. Assume that F is partial differentiable with respect to y and $\frac{\partial F}{\partial y}$ is continuous at $(x_0, y_0) \in D^\circ$, $F(x_0, y_0) = \vartheta_{\mathcal{Z}}$, and $\frac{\partial F}{\partial y}(x_0, y_0) \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ is bijective. Then, the following statements hold.*

- (i) *There exist $r_0, r_1 \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{X}}(x_0, r_0) \times \mathcal{B}_{\mathcal{Y}}(y_0, r_1) \subseteq D$ and $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, r_0)$, $\exists! y \in \mathcal{B}_{\mathcal{Y}}(y_0, r_1)$ satisfying $F(x, y) = \vartheta_{\mathcal{Z}}$. This defines a function $\phi : \mathcal{B}_{\mathcal{X}}(x_0, r_0) \rightarrow \mathcal{B}_{\mathcal{Y}}(y_0, r_1)$ by $\phi(x) = y$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, r_0)$. Then, ϕ is continuous.*

- (ii) If F is Fréchet differentiable at $(x, \phi(x)) \in \mathcal{B}_X(x_0, r_0) \times \mathcal{B}_Y(y_0, r_1)$ for some $x \in \mathcal{B}_X(x_0, r_0)$, then ϕ is Fréchet differentiable at x and

$$\phi^{(1)}(x) = - \left(\frac{\partial F}{\partial y}(x, \phi(x)) \right)^{-1} \frac{\partial F}{\partial x}(x, \phi(x))$$

- (iii) Let $n \in \mathbb{N}$. If F is n -times Fréchet differentiable, then ϕ is n -times Fréchet differentiable.

- (iv) Let $n \in \mathbb{N} \cup \{\infty\}$ and $\bar{x} \in \mathcal{B}_X(x_0, r_0)$. If F is $\mathcal{C}_n$ at $(\bar{x}, \phi(\bar{x}))$, then ϕ is $\mathcal{C}_n$ at $\bar{x}$.

Proof By Implicit Function Theorem 9.58, the statement (i) holds. Furthermore, $\forall (x, y) \in \mathcal{B}_X(x_0, r_0) \times \mathcal{B}_Y(y_0, r_1)$, $\frac{\partial F}{\partial y}(x, y)$ is bijective and

$$\left\| \left(\frac{\partial F}{\partial y}(x, y) \right)^{-1} \right\| < +\infty.$$

(ii) Fix some $x \in \mathcal{B}_X(x_0, r_0)$ such that F is differentiable at $(x, \phi(x))$. Let $y := \phi(x) \in \mathcal{B}_Y(y_0, r_1)$. By Proposition 9.9, $\begin{bmatrix} \frac{\partial F}{\partial x}(x, y) & \frac{\partial F}{\partial y}(x, y) \end{bmatrix} = F^{(1)}(x, y)$. Let $L := - \left(\frac{\partial F}{\partial y}(x, \phi(x)) \right)^{-1} \frac{\partial F}{\partial x}(x, \phi(x)) \in \mathcal{B}(X, Y)$ and $c_1 := \left\| \left(\frac{\partial F}{\partial y}(x, y) \right)^{-1} \right\| < +\infty$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$ with $\epsilon c_1 < 1$, by the differentiability of F at (x, y) , $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $\forall (h, k) \in \mathcal{B}_{X \times Y}((\vartheta_X, \vartheta_Y), \delta_1)$, we have $(x + h, y + k) \in D$ and $\left\| F(x + h, y + k) - F(x, y) - \frac{\partial F}{\partial x}(x, y)h - \frac{\partial F}{\partial y}(x, y)k \right\| \leq \epsilon \|(h, k)\|$. By the continuity of ϕ at x , $\exists \delta \in (0, \min\{r_0 - \|x - x_0\|, \delta_1/\sqrt{2}\}) \subset \mathbb{R}$ such that $\forall h \in \mathcal{B}_X(\vartheta_X, \delta)$, we have $\|\phi(x + h) - \phi(x)\| = \|\phi(x + h) - y\| < \delta_1/\sqrt{2}$.

$\forall h \in \mathcal{B}_X(\vartheta_X, \delta)$, let $\beta := \|\phi(x + h) - \phi(x) - Lh\| \geq 0$. Then, $\|(h, \phi(x + h) - y)\| < \delta_1$. Note that, by Proposition 7.64,

$$\begin{aligned} \beta &= \left\| \left(\frac{\partial F}{\partial y}(x, y) \right)^{-1} \left(\frac{\partial F}{\partial y}(x, y)(\phi(x + h) - \phi(x)) + \frac{\partial F}{\partial x}(x, y)h \right) \right\| \\ &\leq c_1 \left\| F(x + h, \phi(x + h)) - F(x, y) - \frac{\partial F}{\partial x}(x, y)h \right. \\ &\quad \left. - \frac{\partial F}{\partial y}(x, y)(\phi(x + h) - y) \right\| \\ &\leq c_1 \epsilon \|(h, \phi(x + h) - y)\| \leq c_1 \epsilon (\|h\| + \|\phi(x + h) - y\|) \\ &\leq c_1 \epsilon (\|h\| + \|\phi(x + h) - \phi(x) - Lh\| + \|Lh\|) \\ &\leq c_1 \epsilon (\beta + (1 + \|L\|) \|h\|) \end{aligned}$$

Then, we have $\beta \leq \frac{c_1 \epsilon (1 + \|L\|)}{1 - c_1 \epsilon} \|h\|$. Hence, $\phi^{(1)}(x) = L$. Then, the statement (ii) holds.

For (iii), we will use mathematical induction on n to prove this result.

1° $n = 1$. The result follows from (ii).

2° Assume that the result holds for $n = \bar{n} \in \mathbb{N}$.

3° Consider the case $n = \bar{n} + 1$. By (ii), we have ϕ is Fréchet differentiable and $\phi^{(1)}(x) = -\left(\frac{\partial F}{\partial y}(x, \phi(x))\right)^{-1} \frac{\partial F}{\partial x}(x, \phi(x))$, $\forall x \in \mathcal{B}_X(x_0, r_0)$. By inductive assumption, ϕ is $\bar{n}$ -times differentiable. By Proposition 9.46, $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$ are $\bar{n}$ -times differentiable. By Propositions 9.45, 9.56, 9.44, and 9.42, $\phi^{(1)}$ is $\bar{n}$ -times differentiable. Then, ϕ is $(\bar{n} + 1)$ -times differentiable. This completes the induction process.

For (iv), let $\bar{y} = \phi(\bar{x})$. We will first use mathematical induction on n to prove the result for $n \in \mathbb{N}$.

1° $n = 1$. By F being $\mathcal{C}_1$ at $(\bar{x}, \bar{y})$, then $\exists \bar{r} \in (0, \infty) \subset \mathbb{R}$ such that F is differentiable at (x, y) , $\forall (x, y) \in \bar{D} := \mathcal{B}_{X \times Y}((\bar{x}, \bar{y}), \bar{r}) \subseteq D$. By Propositions 9.34 and 9.9, $F|_{\bar{D}}^{(1)}(x, y) = \begin{bmatrix} \frac{\partial F}{\partial x}(x, y) & \frac{\partial F}{\partial y}(x, y) \end{bmatrix}$, $\forall (x, y) \in \bar{D}$. By Proposition 9.46, $\frac{\partial F}{\partial x}|_{\bar{D}}$ and $\frac{\partial F}{\partial y}|_{\bar{D}}$ are continuous at $(\bar{x}, \bar{y})$. Since $\bar{D}$ is open, then $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$ are continuous at $(\bar{x}, \bar{y})$. By the continuity of ϕ , $\exists \bar{\delta} \in (0, \min\{r_0 - \|\bar{x} - x_0\|, \bar{r}\}) \subset \mathbb{R}$ such that F is differentiable at $(x, \phi(x))$, $\forall x \in \mathcal{B}_X(\bar{x}, \bar{\delta})$. By (ii), ϕ is differentiable at x and $\phi^{(1)}(x) = -\left(\frac{\partial F}{\partial y}(x, \phi(x))\right)^{-1} \frac{\partial F}{\partial x}(x, \phi(x))$, $\forall x \in \mathcal{B}_X(\bar{x}, \bar{\delta})$. By Propositions 3.12, 3.32, 9.56, 9.42, and 9.7, $\phi^{(1)}$ is continuous at $\bar{x}$. Hence, ϕ is $\mathcal{C}_1$ at $\bar{x}$.

2° Assume that the result holds for $n = \bar{n} \in \mathbb{N}$.

3° Consider the case $n = \bar{n} + 1$. By F being $\mathcal{C}_{\bar{n}+1}$ at $(\bar{x}, \bar{y})$, then $\exists \bar{r} \in (0, \infty) \subset \mathbb{R}$ such that F is differentiable at (x, y) , $\forall (x, y) \in \bar{D} := \mathcal{B}_{X \times Y}((\bar{x}, \bar{y}), \bar{r}) \subseteq D$. By Propositions 9.34 and 9.9, $F|_{\bar{D}}^{(1)}(x, y) = \begin{bmatrix} \frac{\partial F}{\partial x}(x, y) & \frac{\partial F}{\partial y}(x, y) \end{bmatrix}$, $\forall (x, y) \in \bar{D}$. By Proposition 9.46, $\frac{\partial F}{\partial x}|_{\bar{D}}$ and $\frac{\partial F}{\partial y}|_{\bar{D}}$ are $\mathcal{C}_{\bar{n}}$ at $(\bar{x}, \bar{y})$. Since $\bar{D}$ is open, then, by Proposition 9.34, $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$ are $\mathcal{C}_{\bar{n}}$ at $(\bar{x}, \bar{y})$. By the inductive assumption, ϕ is $\mathcal{C}_{\bar{n}}$ at $\bar{x}$. By the continuity of ϕ , $\exists \bar{\delta} \in (0, \min\{r_0 - \|\bar{x} - x_0\|, \bar{r}\}) \subset \mathbb{R}$ such that F is differentiable at $(x, \phi(x))$, $\forall x \in \mathcal{B}_X(\bar{x}, \bar{\delta})$. By (ii), ϕ is differentiable at x and $\phi^{(1)}(x) = -\left(\frac{\partial F}{\partial y}(x, \phi(x))\right)^{-1} \frac{\partial F}{\partial x}(x, \phi(x))$, $\forall x \in \mathcal{B}_X(\bar{x}, \bar{\delta})$. By Propositions 9.45, 9.44, 9.56, and 9.42, $\phi^{(1)}$ is $\mathcal{C}_{\bar{n}}$ at $\bar{x}$. Hence, ϕ is $\mathcal{C}_{\bar{n}+1}$ at $\bar{x}$.

This completes the induction process. Hence, (iv) holds when $n \in \mathbb{N}$.

When $n = \infty$. There exists $\delta \in (0, \infty) \subset \mathbb{R}$, such that F is $\mathcal{C}_i$ at $(\hat{x}, \hat{y})$, $\forall (\hat{x}, \hat{y}) \in \mathcal{B}_{X \times Y}((\bar{x}, \bar{y}), \delta) \subseteq \mathcal{B}_X(x_0, r_0) \times \mathcal{B}_Y(y_0, r_0)$, $\forall i \in \mathbb{N}$. By the induction result, we have ϕ is $\mathcal{C}_i$ at $\hat{x}$, $\forall \hat{x} \in \mathcal{B}_X(\bar{x}, \delta/\sqrt{2}) \cap \phi_{\text{inv}}(\mathcal{B}_Y(\bar{y}, \delta/\sqrt{2}))$. Hence, ϕ is $\mathcal{C}_{\infty}$ at $\bar{x}$. Hence, (iv) holds.

This completes the proof of the proposition. $\square$

Proposition 9.60 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $\mathcal{Y}$ and $\mathcal{Z}$ be

normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X} \times \mathcal{Y}$, $f : D \rightarrow \mathcal{Z}$ be partial differentiable with respect to y and $\frac{\partial f}{\partial y}(x, y) = \vartheta_{B(\mathcal{Y}, \mathcal{Z})}$, $\forall (x, y) \in D$. Let $\bar{D} := \pi_{\mathcal{X}}(D)$, where $\pi_{\mathcal{X}}$ is the projection function of $\mathcal{X} \times \mathcal{Y}$ to $\mathcal{X}$. Assume that $\forall x \in \bar{D}$, the set $D_x := \{y \in \mathcal{Y} \mid (x, y) \in D\} \subseteq \mathcal{Y}$ is convex. Then, there exists a function $\phi : \bar{D} \rightarrow \mathcal{Z}$ such that $f(x, y) = \phi(x)$, $\forall (x, y) \in D$. Furthermore, the following statements hold.

- (i) If f is continuous at $(x_0, y_0) \in D^\circ$, then ϕ is continuous at x_0 .
- (ii) If $\mathcal{X}$ is a normed linear space $\mathcal{X}$ over $\mathbb{K}$ and f is $\mathcal{C}_k$ at $(x_0, y_0) \in D^\circ$, where $k \in \mathbb{N} \cup \{\infty\}$, then ϕ is $\mathcal{C}_k$ at x_0 .

Proof $\forall x \in \bar{D} = \pi_{\mathcal{X}}(D)$, $D_x \neq \emptyset$. By Axiom of Choice, $\exists g : \bar{D} \rightarrow \mathcal{Y}$ such that $g(x) \in D_x$, $\forall x \in \bar{D}$. Define $\phi : \bar{D} \rightarrow \mathcal{Z}$ by $\phi(x) = f(x, g(x))$, $\forall x \in \bar{D}$. $\forall (x, y) \in D$, we have $x \in \bar{D}$ and $y, g(x) \in D_x$. By the convexity of D_x , the line segment connecting y and $g(x)$ is contained in D_x . By Mean Value Theorem 9.23, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ such that $\|f(x, y) - \phi(x)\| = \|f(x, y) - f(x, g(x))\| \leq \left\| \frac{\partial f}{\partial y}(x, t_0 y + (1 - t_0)g(x))(y - g(x)) \right\| = 0$. Hence, $f(x, y) = \phi(x)$.

(i) Let f be continuous at $(x_0, y_0) \in D^\circ$. Then, $\exists U \in \mathcal{O}$ with $x_0 \in U$ and $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $U \times B_{\mathcal{Y}}(y_0, \delta) \subseteq D$. $\forall x \in U$, we have $\phi(x) = f(x, y_0)$. Hence, ϕ is continuous at x_0 .

(ii) Let $\mathcal{X}$ be a normed linear space $\mathcal{X}$ over $\mathbb{K}$ and f be $\mathcal{C}_k$ at $(x_0, y_0) \in D^\circ$, where $k \in \mathbb{N} \cup \{\infty\}$. Then, $\exists \delta_x, \delta_y \in (0, \infty) \subset \mathbb{R}$ such that $B_{\mathcal{X}}(x_0, \delta_x) \times B_{\mathcal{Y}}(y_0, \delta_y) \subseteq D$. $\forall x \in B_{\mathcal{X}}(x_0, \delta_x)$, we have $\phi(x) = f(x, y_0)$. By Proposition 9.45, $\phi|_{B_{\mathcal{X}}(x_0, \delta_x)}$ is $\mathcal{C}_k$ at x_0 . By Proposition 9.34, ϕ is $\mathcal{C}_k$ at x_0 .

This completes the proof of the proposition. $\square$

9.6 Global Inverse Function Theorem

Definition 9.61 Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be topological spaces, $F : \mathcal{X} \rightarrow \mathcal{Y}$ and $\sigma : \mathcal{Z} \rightarrow \mathcal{Y}$. We will say $\theta : \mathcal{Z} \rightarrow \mathcal{X}$ inverts F along σ if $\sigma = F \circ \theta$.

Lemma 9.62 Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces, $F : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous and countably proper, $x_0 \in \mathcal{X}$, and $y_0 := F(x_0) \in \mathcal{Y}$. Assume that $\forall x \in \mathcal{X}$, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x \in U$ and $\exists V \in \mathcal{O}_{\mathcal{Y}}$ with $F(x) \in V$ such that $F|_U : U \rightarrow V$ is a homeomorphism. Then, given any continuous mapping $\sigma : [a, b] \rightarrow \mathcal{Y}$ with $\sigma(t_0) = y_0$, where $a, t_0, b \in \mathbb{R}$ and $a \leq t_0 \leq b$, there exists a unique continuous mapping $\theta : [a, b] \rightarrow \mathcal{X}$ with $\theta(t_0) = x_0$ that inverts F along σ .

Proof We will distinguish three exhaustive and mutually exclusive cases: Case 1: $a = b = t_0$; Case 2: $a = t_0 < b$; Case 3: $a < t_0 \leq b$.

Case 1: $a = b = t_0$. Clearly, θ exists and is unique. This case is proved.

Case 2: $a = t_0 < b$. “Uniqueness” Let $\theta_1 : [a, b] \rightarrow \mathcal{X}$ and $\theta_2 : [a, b] \rightarrow \mathcal{X}$ be continuous mappings that invert F along σ with $\theta_1(a) = \theta_2(a) = x_0$. Let $S := \{s \in [a, b] \subset \mathbb{R} \mid \theta_1(t) = \theta_2(t) \ \forall t \in [a, s] \subset \mathbb{R}\}$ and $\xi := \sup S$. Clearly, $a \in S$ and $a \leq \xi \leq b$. It is easy to show that $\theta_1(t) = \theta_2(t)$, $\forall t \in \mathbb{R}$ with $a \leq t < \xi$. There exists $(t_n)_{n=1}^\infty \subseteq S$ such that $\lim_{n \in \mathbb{N}} t_n = \xi$. By Proposition 3.66 and the continuity of θ_1 and θ_2 , we have $\theta_1(\xi) = \lim_{n \in \mathbb{N}} \theta_1(t_n) = \lim_{n \in \mathbb{N}} \theta_2(t_n) = \theta_2(t_n)$, where the limit operator makes sense since $\mathcal{X}$ is Hausdorff. Then, $\xi \in S$.

We will next show that $\xi = b$ by an argument of contradiction. Suppose that $\xi < b$. Let $x := \theta_1(\xi) = \theta_2(\xi)$ and $y := F(x)$. Then, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x \in U$ and $\exists V \in \mathcal{O}_{\mathcal{Y}}$ with $y \in V$ such that $F|_U : U \rightarrow V$ is a homeomorphism. By the continuity of θ_1 and θ_2 , $\exists \bar{\xi} \in (\xi, b]$ such that $\theta_1(t), \theta_2(t) \in U$, $\forall t \in [\xi, \bar{\xi}] \subset \mathbb{R}$. Then, $\sigma(t) = F(\theta_1(t)) = F(\theta_2(t)) \in V$, $\forall t \in [\xi, \bar{\xi}] \subset \mathbb{R}$. Since $F|_U : U \rightarrow V$ is a homeomorphism, then $\theta_1(t) = \theta_2(t)$, $\forall t \in [\xi, \bar{\xi}] \subset \mathbb{R}$. Then, $\bar{\xi} \in S$ and $\xi < \bar{\xi} \leq \sup S = \xi$. This is a contradiction. Therefore, we must have $\xi = b$.

Therefore, $\theta_1(t) = \theta_2(t)$, $\forall t \in [a, \xi] = [a, b] \subset \mathbb{R}$, since $\xi \in S$. This shows that $\theta_1 = \theta_2$. Hence, if θ exists then it must be unique.

“Existence” Let $S := \{s \in [a, b] \subset \mathbb{R} \mid \text{there exists a continuous } \theta : [a, s] \rightarrow \mathcal{X} \text{ that inverts } F \text{ along } \sigma|_{[a, s]} \text{ with } \theta(a) = x_0\} \subset \mathbb{R}$ and $\xi := \sup S$. Clearly, $a \in S$ and $a \leq \xi \leq b$.

We will show that $\xi \in S$ by an argument of contradiction. Suppose $\xi \notin S$. Then, $a < \xi \leq b$ and $\exists (t_n)_{n=1}^\infty \subseteq S$, which is nondecreasing, such that $\lim_{n \in \mathbb{N}} t_n = \xi$. $\forall n \in \mathbb{N}$, there exists a continuous $\theta_n : [a, t_n] \rightarrow \mathcal{X}$ that inverts F along $\sigma|_{[a, t_n]}$ with $\theta_n(a) = x_0$. By the uniqueness property that we have shown, we have $\theta_n = \theta_{n+1}|_{[a, t_n]}$. Hence, we may define $\theta : [a, \xi) \rightarrow \mathcal{X}$ such that $\theta(t) = \theta_n(t)$, $\forall t \in [a, t_n] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$. Then, θ is continuous and inverts F along $\sigma|_{[a, \xi)}$ with $\theta(a) = x_0$. Note that $\sigma(t_n) = F(\theta(t_n))$, $\forall n \in \mathbb{N}$. By continuity of σ and Proposition 3.66, we have $\lim_{n \in \mathbb{N}} \sigma(t_n) = \sigma(\xi) \in \mathcal{Y}$, where the limit operator makes sense since $\mathcal{Y}$ is Hausdorff. Then, $(\theta(t_n))_{n=1}^\infty \subseteq F_{\text{inv}}(M)$, where $M := \{\sigma(t_n) \in \mathcal{Y} \mid n \in \mathbb{N}\} \cup \{\sigma(\xi)\} \subseteq \mathcal{Y}$. Clearly, M is compact in $\mathcal{Y}$. Since F is countably proper, then $F_{\text{inv}}(M)$ is countably compact. By Proposition 5.26, $F_{\text{inv}}(M)$ have the Bolzano-Weierstrass property and $(\theta(t_n))_{n=1}^\infty$ admits a cluster point $x \in F_{\text{inv}}(M)$. By the continuity of F and Proposition 3.66, $(F(\theta(t_n)))_{n=1}^\infty = (\sigma(t_n))_{n=1}^\infty$ admits a cluster point $F(x)$. Since $\mathcal{Y}$ is Hausdorff, then $F(x) = \sigma(\xi) =: y$.

By the assumption of the lemma, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x \in U$ and $\exists V \in \mathcal{O}_{\mathcal{Y}}$ with $y \in V$ such that $F|_U : U \rightarrow V$ is a homeomorphism. Since σ is continuous, then $\sigma_{\text{inv}}(V)$ is open in $[a, b] \subset \mathbb{R}$. Since $\xi \in \sigma_{\text{inv}}(V)$, then $\exists \hat{\xi} \in [a, \xi) \subset \mathbb{R}$ such that $[\hat{\xi}, \xi] \subseteq \sigma_{\text{inv}}(V)$. By $\lim_{n \in \mathbb{N}} t_n = \xi$, $\exists N \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n \geq N$, $t_n \in [\hat{\xi}, \xi] \subset \mathbb{R}$. Since $(\theta(t_n))_{n=1}^\infty$ admits a cluster point $x \in U$, then $\exists n_0 \in \mathbb{N}$ with $n_0 \geq N$ such that $\theta(t_{n_0}) \in U$. Clearly, $\sigma([t_{n_0}, \xi]) \subseteq V$. Define $\theta_1 := (F|_U)_{\text{inv}} \circ \sigma|_{[t_{n_0}, \xi]} : [t_{n_0}, \xi] \rightarrow U$. Clearly, θ_1 is continuous and inverts F along $\sigma|_{[t_{n_0}, \xi]}$ with $\theta_1(t_{n_0}) = \theta(t_{n_0})$. Define

$\bar{\theta} : [a, \xi] \rightarrow \mathcal{X}$ by $\bar{\theta}(t) = \theta(t)$, $\forall t \in [a, t_{n_0}] \subset \mathbb{R}$, $\bar{\theta}(t) = \theta_1(t)$, $\forall t \in [t_{n_0}, \xi]$. By Theorem 3.11, $\bar{\theta}$ is continuous. Clearly, $\bar{\theta}$ inverts F along $\sigma|_{[a, \xi]}$ with $\bar{\theta}(a) = \theta(a) = x_0$. Hence, $\xi \in S$. This is a contradiction. Therefore, we must have $\xi \in S$.

Next, we will show that $\xi = b$ by an argument of contradiction. Suppose $\xi < b$. Since $\xi \in S$, then $\exists$ a continuous function $\theta : [a, \xi] \rightarrow \mathcal{X}$ that inverts F along $\sigma|_{[a, \xi]}$ with $\theta(a) = x_0$. Let $x := \theta(\xi) \in \mathcal{X}$ and $y := F(x) = \sigma(\xi) \in \mathcal{Y}$. Then, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x \in U$ and $\exists V \in \mathcal{O}_{\mathcal{Y}}$ with $y \in V$ such that $F|_U : U \rightarrow V$ is a homeomorphism. By the continuity of σ , $\exists \bar{\xi} \in (\xi, b] \subset \mathbb{R}$ such that $\sigma(t) \in V$, $\forall t \in [\xi, \bar{\xi}] \subset \mathbb{R}$. Define $\bar{\theta} : [a, \bar{\xi}] \rightarrow \mathcal{X}$ by $\bar{\theta}(t) = \theta(t)$, $\forall t \in [a, \xi]$, $\bar{\theta}(t) = (F|_U)^{-1}(\sigma(t))$, $\forall t \in [\xi, \bar{\xi}] \subset \mathbb{R}$. By Theorem 3.11, $\bar{\theta}$ is continuous. It is clear that $\bar{\theta}$ inverts F along $\sigma|_{[a, \bar{\xi}]}$ with $\bar{\theta}(a) = x_0$. Then, $\bar{\xi} \in S$. This implies that $\xi < \bar{\xi} \leq \sup S = \xi$, which is a contradiction. Hence, $\xi = b \in S$.

This case is proved.

Case 3: $a < t_0 \leq b$. By Case 1 and 2, $\exists! \theta_1 : [t_0, b] \rightarrow \mathcal{X}$ that is continuous and inverts F along $\sigma|_{[t_0, b]}$ with $\theta_1(t_0) = x_0$. Define $\bar{\sigma} : [t_0, 2t_0 - a] \rightarrow \mathcal{Y}$ by $\bar{\sigma}(t) = \sigma(2t_0 - t)$, $\forall t \in [t_0, 2t_0 - a] \subset \mathbb{R}$. By Proposition 3.12, $\bar{\sigma}$ is continuous and $\bar{\sigma}(t_0) = y_0$. By Case 2, $\exists! \theta_2 : [t_0, 2t_0 - a] \rightarrow \mathcal{X}$ that is continuous and inverts F along $\bar{\sigma}$ with $\theta_2(t_0) = x_0$. Define $\bar{\theta} : [a, b] \rightarrow \mathcal{X}$ by $\bar{\theta}(t) = \theta_1(t)$, $\forall t \in [t_0, b] \subset \mathbb{R}$, $\bar{\theta}(t) = \theta_2(2t_0 - t)$, $\forall t \in [a, t_0] \subset \mathbb{R}$. By Theorem 3.11, $\bar{\theta}$ is continuous and inverts F along σ with $\bar{\theta}(t_0) = x_0$. The uniqueness of $\bar{\theta}$ follows from the uniqueness of θ_1 and θ_2 . This case is proved.

This completes the proof of the lemma. $\square$

Lemma 9.63 *Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces, $F : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous and countably proper, $x_0 \in \mathcal{X}$, and $y_0 = F(x_0) \in \mathcal{Y}$. Assume that $\forall x \in \mathcal{X}$, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x \in U$ and $\exists V \in \mathcal{O}_{\mathcal{Y}}$ with $F(x) \in V$ such that $F|_U : U \rightarrow V$ is a homeomorphism. Then, given any continuous mapping $\sigma : [a, b] \times [c, d] \rightarrow \mathcal{Y}$ with $\sigma(a, c) = y_0$, where $a, b, c, d \in \mathbb{R}$, $a \leq b$, and $c \leq d$, there exists a unique continuous mapping $\theta : [a, b] \times [c, d] \rightarrow \mathcal{X}$ with $\theta(a, c) = x_0$ that inverts F along σ .*

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $a = b$; Case 2: $a < b$. Case 1: $a = b$. The result holds by Lemma 9.62.

Case 2: $a < b$. “Uniqueness” Let $\theta_1 : [a, b] \times [c, d] \rightarrow \mathcal{X}$ and $\theta_2 : [a, b] \times [c, d] \rightarrow \mathcal{X}$ be continuous mappings that invert F along σ with $\theta_1(a, c) = \theta_2(a, c) = x_0$. Fix any $(t, r) \in [a, b] \times [c, d] \subset \mathbb{R}^2$. Let $\bar{\sigma} : [0, 1] \rightarrow \mathcal{Y}$, $\bar{\theta}_1 : [0, 1] \rightarrow \mathcal{X}$, and $\bar{\theta}_2 : [0, 1] \rightarrow \mathcal{X}$ be defined by, $\forall \lambda \in [0, 1] \subset \mathbb{R}$,

$$\begin{aligned} \bar{\sigma}(\lambda) &= \sigma(\lambda t + (1 - \lambda)a, \lambda r + (1 - \lambda)c) \\ \bar{\theta}_1(\lambda) &= \theta_1(\lambda t + (1 - \lambda)a, \lambda r + (1 - \lambda)c) \\ \bar{\theta}_2(\lambda) &= \theta_2(\lambda t + (1 - \lambda)a, \lambda r + (1 - \lambda)c) \end{aligned}$$

Since $\sigma = F \circ \theta_1$ and $\sigma = F \circ \theta_2$, then, $\bar{\sigma} = F \circ \bar{\theta}_1$ and $\bar{\sigma} = F \circ \bar{\theta}_2$. By Proposition 3.12, $\bar{\sigma}$, $\bar{\theta}_1$, and $\bar{\theta}_2$ are continuous. This implies that $\bar{\theta}_1$ inverts F along $\bar{\sigma}$ with $\bar{\theta}_1(0) = \theta_1(a, c) = x_0$; and $\bar{\theta}_2$ inverts F along $\bar{\sigma}$ with $\bar{\theta}_2(0) = x_0$. By Lemma 9.62, we have $\bar{\theta}_1 = \bar{\theta}_2$. Then, $\theta_1(t, r) = \bar{\theta}_1(1) = \bar{\theta}_2(1) = \theta_2(t, r)$. Hence, $\theta_1 = \theta_2$. This shows that $\theta : [a, b] \times [c, d] \rightarrow \mathcal{X}$ is unique when it exists.

“Existence” Define $\sigma_{,c} : [a, b] \rightarrow \mathcal{Y}$ by $\sigma_{,c}(t) = \sigma(t, c)$, $\forall t \in [a, b] \subset \mathbb{R}$. By Proposition 3.12, $\sigma_{,c}$ is continuous with $\sigma_{,c}(a) = y_0$. By Lemma 9.62, there exists a unique continuous mapping $\theta_{,c} : [a, b] \rightarrow \mathcal{X}$ that inverts F along $\sigma_{,c}$ with $\theta_{,c}(a) = x_0$. Fix any $t \in [a, b] \subset \mathbb{R}$. Define $\sigma_t : [c, d] \rightarrow \mathcal{Y}$ by $\sigma_t(r) = \sigma(t, r)$, $\forall r \in [c, d] \subset \mathbb{R}$. By Proposition 3.12, σ_t is continuous with $\sigma_t(c) = \sigma(t, c) = \sigma_{,c}(t) = F(\theta_{,c}(t))$. By Lemma 9.62, there exists a unique continuous mapping $\theta_t : [c, d] \rightarrow \mathcal{X}$ that inverts F along σ_t with $\theta_t(c) = \theta_{,c}(t)$.

Define $\theta : [a, b] \times [c, d] \rightarrow \mathcal{X}$ by $\theta(t, r) = \theta_t(r)$, $\forall (t, r) \in [a, b] \times [c, d]$. Clearly, $\forall (t, r) \in [a, b] \times [c, d]$, $\sigma(t, r) = \sigma_t(r) = F(\theta_t(r)) = F(\theta(t, r))$. Hence, θ inverts F along σ . $\theta(a, c) = \theta_{,c}(a) = x_0$. All we need to show is that θ is continuous to complete the proof of the lemma. Define $S := \{r \in [c, d] \subset \mathbb{R} \mid \theta|_{[a, b] \times [c, r]} \text{ is continuous}\} \subset \mathbb{R}$ and $\xi = \sup S$. Clearly, $c \in S$ and $c \leq \xi \leq d$.

We will show that $\xi \in S$ by an argument of contradiction. Suppose $\xi \notin S$. Then, $c < \xi \leq d$ and $\exists (r_n)_{n=1}^\infty \subseteq S$, which is nondecreasing, such that $\lim_{n \in \mathbb{N}} r_n = \xi$. $\forall (t, r) \in [a, b] \times [c, \xi) \subset \mathbb{R}^2$, there exists $n_0 \in \mathbb{N}$ such that $\forall n \geq n_0$, we have $r_n > r$. Then, $\theta|_{[a, b] \times [c, r_n]}$ is continuous implies that θ is continuous at (t, r) . Hence, $\theta|_{[a, b] \times [c, \xi)}$ is continuous.

Fix any $t \in [a, b]$, let $x = \theta(t, \xi) \in \mathcal{X}$ and $y = F(x) = \sigma(t, \xi) \in \mathcal{Y}$. Then, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x \in U$ and $\exists V \in \mathcal{O}_{\mathcal{Y}}$ with $y \in V$ such that $F|_U : U \rightarrow V$ is a homeomorphism. Since σ is continuous, then $\exists a_t, b_t, c_t, d_t \in \mathbb{R}$ with $a_t < t < b_t$ and $c \leq c_t < \xi < d_t$ such that $\sigma(D_t) \subseteq V$, where $D_t := ((a_t, b_t) \times [c_t, d_t]) \cap ([a, b] \times [c, d]) \subset \mathbb{R}^2$. Define $\bar{\theta} : D_t \rightarrow U$ by $\bar{\theta}(\bar{t}, \bar{r}) = (F|_U)^{-1}(\sigma(\bar{t}, \bar{r}))$, $\forall (\bar{t}, \bar{r}) \in D_t$. By Proposition 3.12, $\bar{\theta}$ is continuous.

Claim 9.63.1 $\theta(\bar{t}, \bar{r}) = \bar{\theta}(\bar{t}, \bar{r})$, $\forall (\bar{t}, \bar{r}) \in D_t$.

Proof of claim: Fix any $(\bar{t}, \bar{r}) \in D_t$. Note that $\bar{\theta}$ inverts F along $\sigma|_{D_t}$. Define $\bar{\theta}_t : D_{t,2} \rightarrow \mathcal{X}$ by $\bar{\theta}_t(\hat{r}) = \bar{\theta}(\bar{t}, \hat{r})$, $\forall \hat{r} \in D_{t,2} := [c_t, d_t] \cap [c, d] \subset \mathbb{R}$. Then, $\bar{\theta}_t$ is continuous and inverts F along $\sigma_t|_{D_{t,2}}$ with $\bar{\theta}_t(\xi) = x$. Note that $\theta_t|_{D_{t,2}}$ is also continuous and inverts F along $\sigma_t|_{D_{t,2}}$ with $\theta_t(\xi) = x$. By Lemma 9.62, we have $\bar{\theta}_t = \theta_t|_{D_{t,2}}$ and, in particular, $\bar{\theta}(\bar{t}, c_t) = \bar{\theta}_t(c_t) = \theta_t(c_t) = \theta(\bar{t}, c_t)$. Define $\bar{\theta}_{,c_t} : D_{t,1} \rightarrow \mathcal{X}$ by $\bar{\theta}_{,c_t}(\hat{t}) = \bar{\theta}(\hat{t}, c_t)$, $\forall \hat{t} \in D_{t,1} := (a_t, b_t) \cap [a, b] \subset \mathbb{R}$. Define $\sigma_{,c_t} : [a, b] \rightarrow \mathcal{Y}$ by $\sigma_{,c_t}(\hat{t}) = \sigma(\hat{t}, c_t)$, $\forall \hat{t} \in [a, b]$. Define $\theta_{,c_t} : [a, b] \rightarrow \mathcal{X}$ by $\theta_{,c_t}(\hat{t}) = \theta(\hat{t}, c_t)$, $\forall \hat{t} \in [a, b]$. Then, $\bar{\theta}_{,c_t}$ is continuous and inverts F along $\sigma_{,c_t}|_{D_{t,1}}$ with $\bar{\theta}_{,c_t}(t) = \theta(\bar{t}, c_t)$. Since $c_t \in [c, \xi) \subset \mathbb{R}$, then $\theta_{,c_t}|_{D_{t,1}}$ is continuous and inverts F along $\sigma_{,c_t}|_{D_{t,1}}$.

with $\theta_{,c_t}(t) = \theta(t, c_t)$. By Lemma 9.62, we have $\theta_{,c_t}|_{D_{t,1}} = \bar{\theta}_{,c_t}$ and, in particular, $\theta(\bar{t}, c_t) = \theta_{,c_t}(\bar{t}) = \bar{\theta}_{,c_t}(\bar{t}) = \bar{\theta}(\bar{t}, c_t)$. Define $\bar{\theta}_{\bar{t}} : D_{t,2} \rightarrow \mathcal{X}$ by $\bar{\theta}_{\bar{t}}(\hat{r}) = \bar{\theta}(\bar{t}, \hat{r})$, $\forall \hat{r} \in D_{t,2}$. Then, $\bar{\theta}_{\bar{t}}$ is continuous and inverts F along $\sigma_{\bar{t}}|_{D_{t,2}}$ with $\bar{\theta}_{\bar{t}}(c_t) = \bar{\theta}(\bar{t}, c_t) = \theta(\bar{t}, c_t)$. Note that $\theta_{\bar{t}}|_{D_{t,2}}$ is also continuous and inverts F along $\sigma_{\bar{t}}|_{D_{t,2}}$ with $\theta_{\bar{t}}(c_t) = \theta(\bar{t}, c_t)$. By Lemma 9.62, we have $\bar{\theta}_{\bar{t}} = \theta_{\bar{t}}|_{D_{t,2}}$ and, in particular, $\bar{\theta}(\bar{t}, \bar{r}) = \bar{\theta}_{\bar{t}}(\bar{r}) = \theta_{\bar{t}}(\bar{r}) = \theta(\bar{t}, \bar{r})$. This completes the proof of the claim. $\square$

Then, $\theta|_{D_t}$ is continuous. Then, θ is continuous at (t, ξ) . Then, $\theta|_{[a,b] \times [c,\xi]}$ is continuous and $\xi \in S$. This contradicts with the hypothesis $\xi \notin S$. Therefore, $\xi \in S$.

Now, we show $\xi = d$ by an argument of contradiction. Suppose $\xi < d$. Fix any $t \in [a, b]$, let $x = \theta(t, \xi) \in \mathcal{X}$ and $y = F(x) = \sigma(t, \xi) \in \mathcal{Y}$. Then, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x \in U$ and $\exists V \in \mathcal{O}_{\mathcal{Y}}$ with $y \in V$ such that $F|_U : U \rightarrow V$ is a homeomorphism. Since σ is continuous, then $\exists a_t, b_t, c_t, d_t \in \mathbb{R}$ with $a_t < t < b_t$ and $c_t < \xi < d_t \leq d$ such that $\sigma(D_t) \subseteq V$, where $D_t := ((a_t, b_t) \times (c_t, d_t)) \cap ([a, b] \times [c, d]) \subset \mathbb{R}^2$. Define $\bar{\theta} : D_t \rightarrow U$ by $\bar{\theta}(\bar{t}, \bar{r}) = (F|_U)^{-1}(\sigma(\bar{t}, \bar{r}))$, $\forall (\bar{t}, \bar{r}) \in D_t$. By Proposition 3.12, $\bar{\theta}$ is continuous.

Claim 9.63.2 $\theta(\bar{t}, \bar{r}) = \bar{\theta}(\bar{t}, \bar{r})$, $\forall (\bar{t}, \bar{r}) \in D_t$.

Proof of claim: Fix any $(\bar{t}, \bar{r}) \in D_t$. Note that $\bar{\theta}$ inverts F along $\sigma|_{D_t}$. Define $\bar{\theta}_{,\xi} : D_{t,1} \rightarrow \mathcal{X}$ by $\bar{\theta}_{,\xi}(\hat{t}) = \bar{\theta}(\hat{t}, \xi)$, $\forall \hat{t} \in D_{t,1} := (a_t, b_t) \cap [a, b] \subset \mathbb{R}$. Define $\sigma_{,\xi} : [a, b] \rightarrow \mathcal{Y}$ by $\sigma_{,\xi}(\hat{t}) = \sigma(\hat{t}, \xi)$, $\forall \hat{t} \in [a, b]$. Define $\theta_{,\xi} : [a, b] \rightarrow \mathcal{X}$ by $\theta_{,\xi}(\hat{t}) = \theta(\hat{t}, \xi)$, $\forall \hat{t} \in [a, b]$. Then, $\bar{\theta}_{,\xi}$ is continuous and inverts F along $\sigma_{,\xi}|_{D_{t,1}}$ with $\bar{\theta}_{,\xi}(t) = x$. Since $\xi \in S$, then $\theta_{,\xi}|_{D_{t,1}}$ is continuous and inverts F along $\sigma_{,\xi}|_{D_{t,1}}$ with $\theta_{,\xi}(t) = \theta(t, \xi) = x$. By Lemma 9.62, we have $\theta_{,\xi}|_{D_{t,1}} = \bar{\theta}_{,\xi}$ and, in particular, $\theta(\bar{t}, \xi) = \theta_{,\xi}(\bar{t}) = \bar{\theta}_{,\xi}(\bar{t}) = \bar{\theta}(\bar{t}, \xi)$. Define $\bar{\theta}_{\bar{t}} : D_{t,2} \rightarrow \mathcal{X}$ by $\bar{\theta}_{\bar{t}}(\hat{r}) = \bar{\theta}(\bar{t}, \hat{r})$, $\forall \hat{r} \in D_{t,2} := (c_t, d_t) \cap [c, d] \subset \mathbb{R}$. Then, $\bar{\theta}_{\bar{t}}$ is continuous and inverts F along $\sigma_{\bar{t}}|_{D_{t,2}}$ with $\bar{\theta}_{\bar{t}}(\xi) = \bar{\theta}(\bar{t}, \xi) = \theta(\bar{t}, \xi)$. Note that $\theta_{\bar{t}}|_{D_{t,2}}$ is also continuous and inverts F along $\sigma_{\bar{t}}|_{D_{t,2}}$ with $\theta_{\bar{t}}(\xi) = \theta(\bar{t}, \xi)$. By Lemma 9.62, we have $\bar{\theta}_{\bar{t}} = \theta_{\bar{t}}|_{D_{t,2}}$ and, in particular, $\bar{\theta}(\bar{t}, \bar{r}) = \bar{\theta}_{\bar{t}}(\bar{r}) = \theta_{\bar{t}}(\bar{r}) = \theta(\bar{t}, \bar{r})$. This completes the proof of the claim. $\square$

Then, $\theta|_{D_t}$ is continuous. Note that $[a, b] \subseteq \bigcup_{t \in [a,b] \subset \mathbb{R}} (a_t, b_t)$. By the compactness of $[a, b] \subseteq \mathbb{R}$, there exists a finite set $T_N \subseteq [a, b] \subset \mathbb{R}$ such that $[a, b] \subseteq \bigcup_{t \in T_N} (a_t, b_t)$. Note that $\theta|_{D_t}$, $\forall t \in T_N$ and $\theta|_{[a,b] \times [c,\xi]}$ are continuous. Then, by Theorem 3.11, $\theta|_D$ is continuous, where $D := ([a, b] \times [c, \xi]) \cup (\bigcup_{t \in T_N} D_t) \subseteq [a, b] \times [c, d] \subseteq \mathbb{R}^2$ and all of the sets are relatively open. Set $\bar{d} = (\min_{t \in T_N} d_t + \xi)/2 \in (\xi, d]$. It is clear that $[a, b] \times [c, \bar{d}] \subseteq D \subseteq \mathbb{R}^2$. Then, $\theta|_{[a,b] \times [c,\bar{d}]}$ is continuous and $\bar{d} \in S$. This leads to the contradiction $\xi < \bar{d} \leq \sup S = \xi$. Therefore, $\xi = d$. Then, θ is continuous. This completes the proof of Case 2.

This completes the proof of the lemma. $\square$

Theorem 9.64 (Global Inverse Function Theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces, $\mathcal{X} \neq \emptyset$, $F : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous and countably proper. Assume that $\forall x \in \mathcal{X}$, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x \in U$ and $\exists V \in \mathcal{O}_{\mathcal{Y}}$ with $F(x) \in V$ such that $F|_U : U \rightarrow V$ is a homeomorphism, $\mathcal{X}$ is arcwise connected, and $\mathcal{Y}$ is simply connected. Then, $F : \mathcal{X} \rightarrow \mathcal{Y}$ is a homeomorphism.*

Proof Fix $x_0 \in \mathcal{X} \neq \emptyset$, let $y_0 = F(x_0) \in \mathcal{Y}$. $\forall y \in \mathcal{Y}$, since $\mathcal{Y}$ is simply connected, then $\mathcal{Y}$ is arcwise connected. Then, there exists a curve $\sigma : I \rightarrow \mathcal{Y}$, where $I := [0, 1] \subset \mathbb{R}$, such that $\sigma(0) = y_0$ and $\sigma(1) = y$. By Lemma 9.62, there exists a continuous mapping $\theta : I \rightarrow \mathcal{X}$ that inverts F along σ with $\theta(0) = x_0$. Then, $y = \sigma(1) = F(\theta(1))$. Hence, F is surjective.

Fix $x_1, x_2 \in \mathcal{X}$ such that $F(x_1) = F(x_2) = y$. Since $\mathcal{X}$ is arcwise connected, then there exists a curve $\delta : I \rightarrow \mathcal{X}$ such that $\delta(0) = x_1$ and $\delta(1) = x_2$. Consider the curve $\eta := F \circ \delta$, which is continuous by Proposition 3.12. η is a closed curve since $\eta(0) = F(x_1) = y = F(x_2) = \eta(1)$. Since $\mathcal{Y}$ is simply connected, then η is homotopic to a single point $\bar{y} \in \mathcal{Y}$. Then, there exists a continuous mapping $\gamma : I \times I \rightarrow \mathcal{Y}$ such that $\gamma(t, 0) = \eta(t)$, $\gamma(t, 1) = \bar{y}$, and $\gamma(0, t) = \gamma(1, t)$, $\forall t \in I$. By Lemma 9.63, there exists a continuous function $\zeta : I \times I \rightarrow \mathcal{X}$ that inverts F along γ with $\zeta(0, 0) = x_1$. Define $\gamma_0 : I \rightarrow \mathcal{Y}$ by $\gamma_0(t) = \gamma(0, t)$, $\gamma_1 : I \rightarrow \mathcal{Y}$ by $\gamma_1(t) = \gamma(1, t)$, $\gamma_{0,0} : I \rightarrow \mathcal{Y}$ by $\gamma_{0,0}(t) = \gamma(t, 0)$, and $\gamma_{1,1} : I \rightarrow \mathcal{Y}$ by $\gamma_{1,1}(t) = \gamma(t, 1)$, $\forall t \in I$. Then, $\gamma_0 = \gamma_1$. Define $\zeta_0 : I \rightarrow \mathcal{X}$ by $\zeta_0(t) = \zeta(0, t)$, $\zeta_1 : I \rightarrow \mathcal{X}$ by $\zeta_1(t) = \zeta(1, t)$, $\zeta_{0,0} : I \rightarrow \mathcal{X}$ by $\zeta_{0,0}(t) = \zeta(t, 0)$, and $\zeta_{1,1} : I \rightarrow \mathcal{X}$ by $\zeta_{1,1}(t) = \zeta(t, 1)$, $\forall t \in I$. Then, ζ_0 is continuous and inverts F along γ_0 . Set $\bar{x} = \zeta_0(1) = \zeta(0, 1) \in \mathcal{X}$, then, $\bar{y} = F(\bar{x})$. $\zeta_{1,1}$ is continuous and inverts F along $\gamma_{1,1}$ with $\zeta_{1,1}(0) = \zeta(0, 1) = \bar{x}$. Since $\gamma_{1,1}$ is a constant function with value $\bar{y}$. Then, the constant mapping $\lambda : I \rightarrow \mathcal{X}$ given by $\lambda(t) = \bar{x}$, $\forall t \in I$ is continuous and inverts F along $\gamma_{1,1}$ with $\lambda(0) = \bar{x}$. By Lemma 9.62, we have $\lambda = \zeta_{1,1}$ and, in particular, $\bar{x} = \lambda(1) = \zeta_{1,1}(1) = \zeta(1, 1)$. Note that ζ_1 is continuous and inverts F along γ_1 with $\zeta_1(1) = \zeta(1, 1) = \bar{x}$. Since $\gamma_1 = \gamma_0$. ζ_0 is continuous and inverts F along γ_1 with $\zeta_0(1) = \bar{x}$. By Lemma 9.62, we have $\zeta_0 = \zeta_1$ and in particular $\zeta(1, 0) = \zeta_1(0) = \zeta_0(0) = \zeta(0, 0) = x_1$. Note that $\zeta_{0,0}$ is continuous and inverts F along $\gamma_{0,0} = \eta$ with $\zeta_{0,0}(0) = \zeta(0, 0) = x_1$. By construction, δ is continuous and inverts F along η with $\delta(0) = x_1$. By Lemma 9.62, $\delta = \zeta_{0,0}$ and in particular, $x_2 = \delta(1) = \zeta_{0,0}(1) = \zeta(1, 0) = x_1$. Hence, $x_1 = x_2$. Therefore, F is injective.

Hence, F is bijective and admits inverse $F_i : \mathcal{Y} \rightarrow \mathcal{X}$. $\forall y \in \mathcal{Y}$, let $x = F_i(y) \in \mathcal{X}$. Then, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x \in U$ and $\exists V \in \mathcal{O}_{\mathcal{Y}}$ with $y \in V$ such that $F|_U : U \rightarrow V$ is a homeomorphism. Then, $F_i|_V$ is the inverse of $F|_U$ and is continuous. Then, F_i is continuous at y since V is open. By the arbitrariness of y , F_i is continuous. Hence, F is a homeomorphism.

This completes the proof of the theorem. $\square$

Theorem 9.65 (Global Inverse Function Theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces, $F : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous and countably*

proper. Let $H := \{x \in \mathcal{X} \mid \exists U \in \mathcal{O}_{\mathcal{X}} \text{ with } x \in U \ni \cdot \cdot F|_U : U \rightarrow F(U) \in \mathcal{O}_{\mathcal{Y}} \text{ is a homeomorphism}\} \subseteq \mathcal{X}$, $\Sigma := \mathcal{X} \setminus H$, $\Sigma_0 := F_{\text{inv}}(F(\Sigma))$, $X_0 := \mathcal{X} \setminus \Sigma_0$, and $Y_0 := \mathcal{Y} \setminus F(\Sigma)$. Assume that $X_0 \neq \emptyset$ is arcwise connected, and Y_0 is simply connected. Then, $G := F|_{X_0} : X_0 \rightarrow Y_0$ is a homeomorphism.

Proof Let $\mathcal{O}_{X_0}$ and $\mathcal{O}_{Y_0}$ be the subset topology on X_0 and Y_0 , respectively.

Claim 9.65.1 $\Sigma \subseteq \mathcal{X}$ is closed.

Proof of claim: $\forall x \in H$, $\exists U_x \in \mathcal{O}_{\mathcal{X}}$ with $x \in U_x$ such that $F|_{U_x} : U_x \rightarrow F(U_x) \in \mathcal{O}_{\mathcal{Y}}$ is a homeomorphism. $\forall \bar{x} \in U_x$, $\bar{x} \in H$. Then, $U_x \subseteq H$ and $H = \bigcup_{x \in H} U_x$. Hence, $H \in \mathcal{O}_{\mathcal{X}}$. Then, $\Sigma := \mathcal{X} \setminus H$ is closed. $\square$

Claim 9.65.2 $F_{\text{inv}}(Y_0) = X_0$. $G : X_0 \rightarrow Y_0$ is continuous and countably proper.

Proof of claim: By Proposition 2.5, $F_{\text{inv}}(Y_0) = F_{\text{inv}}(\mathcal{Y} \setminus F(\Sigma)) = F_{\text{inv}}(\mathcal{Y}) \setminus F_{\text{inv}}(F(\Sigma)) = \mathcal{X} \setminus \Sigma_0 = X_0$ and $F(X_0) = F(F_{\text{inv}}(Y_0)) \subseteq Y_0$. Hence, G is a function of X_0 to Y_0 .

Fix any $K \subseteq Y_0$ such that K is compact in $\mathcal{O}_{Y_0}$. Then K is a compact set in $\mathcal{O}_{\mathcal{Y}}$. Then, $G_{\text{inv}}(K) = F_{\text{inv}}(K) \subseteq X_0$. By the countable properness of F , we have $F_{\text{inv}}(K) \subseteq \mathcal{X}$ is countably compact in $\mathcal{O}_{\mathcal{X}}$. Then, it is easy to show that $G_{\text{inv}}(K)$ is countably compact in $\mathcal{O}_{X_0}$. Hence, G is countably proper.

$\forall V_0 \in \mathcal{O}_{Y_0}$, $V_0 = Y_0 \cap V$, where $V \in \mathcal{O}_{\mathcal{Y}}$. $G_{\text{inv}}(V_0) = F_{\text{inv}}(V_0) = F_{\text{inv}}(Y_0) \cap F_{\text{inv}}(V) = X_0 \cap F_{\text{inv}}(V)$. Since F is continuous, $F_{\text{inv}}(V) \in \mathcal{O}_{\mathcal{X}}$. Thus, $G_{\text{inv}}(V_0) \in \mathcal{O}_{X_0}$. Hence, G is continuous. $\square$

By Proposition 2.5, $\Sigma_0 = F_{\text{inv}}(F(\Sigma)) \supseteq \Sigma \cap \text{dom}(F) = \Sigma$. Then, $X_0 = \mathcal{X} \setminus \Sigma_0 \subseteq \mathcal{X} \setminus \Sigma = H$. $\forall x \in X_0 \subseteq H$, $\exists U \in \mathcal{O}_{\mathcal{X}}$ with $x \in U$ such that $F|_U : U \rightarrow F(U) \in \mathcal{O}_{\mathcal{Y}}$ is a homeomorphism. Let $U_0 := X_0 \cap U \in \mathcal{O}_{X_0}$, $V := F(U) \in \mathcal{O}_{\mathcal{Y}}$, and $V_0 := F(U) \cap Y_0 \in \mathcal{O}_{Y_0}$. Clearly, $x \in U_0$. By Proposition 2.5, $G(U_0) = F(U_0) \subseteq F(X_0) \cap F(U) \subseteq Y_0 \cap V = V_0$. Then, $G|_{U_0} : U_0 \rightarrow V_0$. Note that $G|_{U_0} = (F|_U)|_{U_0}$. Since $F|_U$ is injective, then $G|_{U_0}$ is injective. $\forall \bar{y} \in V_0$, $\bar{y} \in V$ then $\exists \bar{x} \in U$ such that $\bar{y} = F(\bar{x})$. Note that $\bar{y} \in Y_0$ and $F_{\text{inv}}(Y_0) = X_0$, then $\bar{x} \in X_0$. Hence, $\bar{x} \in U_0$ and $G(\bar{x}) = F(\bar{x}) = \bar{y}$. Then, $G|_{U_0} : U_0 \rightarrow V_0$ is surjective. Hence, $G|_{U_0}$ is bijective with inverse $G_i = F_i|_{V_0}$, where F_i is the inverse of $F|_U : U \rightarrow V$. Since $F|_U$ is homeomorphism, then $F|_U$ and F_i are continuous. Then, $G|_{U_0}$ and G_i are continuous. This shows that $G|_{U_0} : U_0 \rightarrow V_0$ is a homeomorphism.

By Global Inverse Function Theorem 9.64, $G : X_0 \rightarrow Y_0$ is a homeomorphism. This completes the proof of the theorem. $\square$

Proposition 9.66 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $A_{\alpha} \subseteq X$ be arcwise connected (in subset topology), $\forall \alpha \in \Lambda$, where Λ is an index set. Assume that $A_{\alpha_1} \cap A_{\alpha_2} \neq \emptyset$, $\forall \alpha_1, \alpha_2 \in \Lambda$. Then, $A := \bigcup_{\alpha \in \Lambda} A_{\alpha}$ is arcwise connected (in subset topology).

Proof $\forall x_1, x_2 \in A$, $\exists \alpha_1, \alpha_2 \in \Lambda$ such that $x_i \in A_{\alpha_i}$, $i = 1, 2$. By the assumption, let $x_0 \in A_{\alpha_1} \cap A_{\alpha_2} \neq \emptyset$. $\forall i \in \{1, 2\}$, since A_{α_i} is arcwise connected, $\exists$ curve $\gamma_i : [0, 1] \subset \mathbb{R} \rightarrow A_{\alpha_i}$ such that $\gamma_i(0) = x_0$ and $\gamma_i(1) = x_i$. Define $\gamma : [0, 1] \subset \mathbb{R} \rightarrow A$ by $\gamma(t) = \begin{cases} \gamma_1(1-2t) & 0 \leq t \leq 1/2 \\ \gamma_2(2t-1) & 1/2 < t \leq 1 \end{cases}$, $\forall t \in [0, 1] \subset \mathbb{R}$. Then $\gamma(0) = \gamma_1(1) = x_1$ and $\gamma(1) = \gamma_2(1) = x_2$. Note that $\gamma|_{[0, 1/2]}(t) = \gamma_1(1-2t)$, $\forall t \in [0, 1/2] \subset \mathbb{R}$, and $\gamma|_{[1/2, 1]}(t) = \gamma_2(2t-1)$, $\forall t \in [1/2, 1] \subset \mathbb{R}$, are continuous functions. By Theorem 3.11, γ is continuous. Therefore, γ is a curve connecting x_1 and x_2 . By the arbitrariness of x_1 and x_2 , we have A is arcwise connected. This completes the proof of the proposition. $\square$

Proposition 9.67 *Let $\mathcal{X}$ be a normed linear space and $O \subseteq \mathcal{X}$ be open and connected. Then, O is arcwise connected.*

Proof The result is trivial if $O = \emptyset$. Let $x_0 \in O \neq \emptyset$. Let $\mathcal{M} := \{A \subseteq O \mid x_0 \in A \text{ and } A \text{ is arcwise connected}\}$ and $A_0 := \bigcup_{A \in \mathcal{M}} A$. Then, $x_0 \in A_0 \subseteq O$ and, by Proposition 9.66, A_0 is arcwise connected. $\forall x \in A_0 \subseteq O$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{X}}(x, \delta) \subseteq O$. Clearly, $\mathcal{B}_{\mathcal{X}}(x, \delta)$ is arcwise connected. Then, by Proposition 9.66, $A_0 \cup \mathcal{B}_{\mathcal{X}}(x, \delta)$ is arcwise connected. This implies that $A_0 \cup \mathcal{B}_{\mathcal{X}}(x, \delta) \in \mathcal{M}$. Then, $\mathcal{B}_{\mathcal{X}}(x, \delta) \subseteq A_0$. By the arbitrariness of x , we have A_0 is open in $\mathcal{X}$. We will show that $A_0 = O$ by an argument of contradiction. Suppose $A_0 \subset O$. Let $E := O \setminus A_0 \neq \emptyset$. Let ∂E be the boundary of E in $\mathcal{X}$.

Claim 9.67.1 $\partial E \cap E \neq \emptyset$.

Proof of claim: Suppose $\partial E \cap E = \emptyset$. By Proposition 3.3, $E = \overline{E} \cap E = (\partial E \cup E^\circ) \cap E = E^\circ \cap E = E^\circ$. Hence, E is open in $\mathcal{X}$. Then, A_0 and E form a separation of O . This contradicts with the assumption that O is connected. Hence, the claim holds. $\square$

Let $x_1 \in \partial E \cap E \subseteq O$. Then, $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{X}}(x_1, \delta_1) \subseteq O$. $x_1 \in \partial E$ implies that $\mathcal{B}_{\mathcal{X}}(x_1, \delta_1) \cap \widetilde{E} \neq \emptyset$. This implies that $\exists x_2 \in \mathcal{B}_{\mathcal{X}}(x_1, \delta_1) \cap (O \setminus E) = \mathcal{B}_{\mathcal{X}}(x_1, \delta_1) \cap A_0$. By Proposition 9.66, $A_0 \cup \mathcal{B}_{\mathcal{X}}(x_1, \delta_1)$ is arcwise connected. Then, by the definition of A_0 , we have $\mathcal{B}_{\mathcal{X}}(x_1, \delta_1) \subseteq A_0$ and $x_1 \in A_0$. This contradicts with the fact that $x_1 \in E = O \setminus A_0$. Therefore, $A_0 = O$ and O is arcwise connected. This completes the proof of the proposition. $\square$

9.7 Interchange Differentiation and Limit

Proposition 9.68 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $x_0 \in D$, and $F_n : D \rightarrow \mathcal{Y}$, $\forall n \in \mathbb{N}$. Assume that*

- (i) $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $\bar{D} := D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) - x_0$ is a conic segment;

(ii) F_n is differentiable, $\forall n \in \mathbb{N}$;

(iii) $\left(F_n^{(1)}\right)_{n=1}^{\infty}$ converges uniformly to $G : D \rightarrow \mathcal{B}(\mathcal{X}, \mathcal{Y})$;

(iv) $\forall x \in D$, $\lim_{n \in \mathbb{N}} F_n(x) = F(x)$, where $F : D \rightarrow \mathcal{Y}$.

Then, F is differentiable at x_0 and $F^{(1)}(x_0) = G(x_0) = \lim_{n \in \mathbb{N}} F_n^{(1)}(x_0)$.

Proof By the differentiability of F_1 , we have $\overline{\text{span}(A_D(x))} = \mathcal{X}$, $\forall x \in D$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by (iii), $\exists n_0 \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n \geq n_0$, $\|F_n^{(1)}(x) - G(x)\| < \epsilon/4$, $\forall x \in D$. $\forall n, m \in \mathbb{N}$ with $n \geq n_0$ and $m \geq n_0$, by Proposition 9.15, $g : D \rightarrow \mathcal{Y}$, defined by $g(x) = F_n(x) - F_m(x)$, $\forall x \in D$, is differentiable and $g^{(1)}(x) = F_n^{(1)}(x) - F_m^{(1)}(x)$, $\forall x \in D$. By Mean Value Theorem 9.23, $\forall x \in \bar{D} + x_0$ and $\forall n, m \in \mathbb{N}$ with $n \geq n_0$ and $m \geq n_0$, $\|F_n(x) - F_m(x) - F_n(x_0) + F_m(x_0)\| = \|g(x) - g(x_0)\| \leq \|g^{(1)}(\bar{x})(x - x_0)\| \leq \|F_n^{(1)}(\bar{x}) - F_m^{(1)}(\bar{x})\| \|x - x_0\| \leq \epsilon \|x - x_0\| / 2$, where $\bar{x} = t_0 x + (1 - t_0)x_0 \in \bar{D} + x_0$ and $t_0 \in (0, 1) \subset \mathbb{R}$. By $F_{n_0}^{(1)}(x_0) \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, $\exists \delta \in (0, \delta_0] \subset \mathbb{R}$ such that, $\forall x \in \hat{D} := D \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) \subseteq \bar{D} + x_0$, we have $\|F_{n_0}(x) - F_{n_0}(x_0) - F_{n_0}^{(1)}(x_0)(x - x_0)\| \leq \epsilon \|x - x_0\| / 4$.

Fix any $x \in \hat{D}$. $\forall m \in \mathbb{N}$ with $m \geq n_0$, we have $\|F_m(x) - F_m(x_0) - G(x_0)(x - x_0)\| \leq \|F_m(x) - F_m(x_0) - F_{n_0}(x) + F_{n_0}(x_0)\| + \|F_{n_0}(x) - F_{n_0}(x_0) - F_{n_0}^{(1)}(x_0)(x - x_0)\| + \|F_{n_0}^{(1)}(x_0)(x - x_0) - G(x_0)(x - x_0)\| \leq \epsilon \|x - x_0\| / 2 + \epsilon \|x - x_0\| / 4 + \epsilon \|x - x_0\| / 4 = \epsilon \|x - x_0\|$. Take limit as $m \rightarrow \infty$, we have $\|F(x) - F(x_0) - G(x_0)(x - x_0)\| \leq \lim_{m \in \mathbb{N}} \|F_m(x) - F_m(x_0) - G(x_0)(x - x_0)\| \leq \epsilon \|x - x_0\|$. By the arbitrariness of x , F is differentiable at x_0 and $F^{(1)}(x_0) = G(x_0) = \lim_{n \in \mathbb{N}} DF_n(x_0)$. This completes the proof of the proposition. $\square$

Example 9.69 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $\Omega \subseteq \mathcal{X}$ be a compact set, which satisfies $\overline{\text{span}(A_{\Omega}(x))} = \mathcal{X}$, $\forall x \in \Omega$, $k \in \mathbb{N}$, and $\mathcal{W} := \mathcal{C}(\Omega, \mathcal{Y})$ be the normed linear space defined in Example 7.31. Define $Z := \{f \in \mathcal{C}(\Omega, \mathcal{Y}) \mid f \text{ is } \mathcal{C}_k\}$. By Proposition 9.40, Z is a subspace of $\mathcal{C}(\Omega, \mathcal{Y})$. Then, $\mathcal{Z} := (Z, \oplus_{\mathcal{W}}, \otimes_{\mathcal{W}}, \vartheta_{\mathcal{W}})$ is vector space over the field $\mathbb{K}$. $\forall f \in Z$, $\forall i \in \{0, \dots, k\}$, $f^{(i)} \in \mathcal{C}(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y}))$ with norm $\|f^{(i)}\|_{\mathcal{C}} \in \mathbb{R}$.

Now, define a norm on Z by $\|f\|_{\mathcal{C}_k} := \left(\sum_{i=0}^k \|f^{(i)}\|_{\mathcal{C}}^2\right)^{1/2} \in [0, \infty) \subset \mathbb{R}$, $\forall f \in Z$. If $f = \vartheta_{\mathcal{W}}$, by Proposition 9.33, f is $\mathcal{C}_{\infty}$ and $f^{(i)}(x) = \vartheta_{\mathcal{B}_i(\mathcal{X}, \mathcal{Y})}$, $\forall x \in \Omega$, $\forall i \in \mathbb{N}$, then, $\|f\|_{\mathcal{C}_k} = 0$. On the other hand, if $\|f\|_{\mathcal{C}_k} = 0$, then $f = \vartheta_{\mathcal{W}}$. $\forall f_1, f_2 \in Z$, by Proposition 9.40, $\|f_1 + f_2\|_{\mathcal{C}_k} = \left(\sum_{i=0}^k \|f_1^{(i)} + f_2^{(i)}\|_{\mathcal{C}}^2\right)^{1/2} \leq \left(\sum_{i=0}^k \|f_1^{(i)}\|_{\mathcal{C}}^2\right)^{1/2} + \left(\sum_{i=0}^k \|f_2^{(i)}\|_{\mathcal{C}}^2\right)^{1/2} = \|f_1\|_{\mathcal{C}_k} + \|f_2\|_{\mathcal{C}_k}$, where the inequality follows from Minkowski's Inequality, Theorem 7.9. $\forall \alpha \in \mathbb{K}$, $\forall f \in Z$, by Proposition 9.40, $\|\alpha f\|_{\mathcal{C}_k} =$

$\left(\sum_{i=0}^k \|\alpha f^{(i)}\|_{\mathcal{C}}^2\right)^{1/2} = \left(\sum_{i=0}^k |\alpha|^2 \|f^{(i)}\|_{\mathcal{C}}^2\right)^{1/2} = |\alpha| \|f\|_{\mathcal{C}_k}$. This shows that $(\mathcal{Z}, \mathbb{K}, \|\cdot\|_{\mathcal{C}_k})$ is a normed linear space, which will be denoted $\mathcal{C}_k(\Omega, \mathcal{Y})$. $\diamond$

Example 9.70 Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$ and $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $\Omega \subseteq \mathcal{X}$ be a compact set, and $k \in \mathbb{N}$. Assume that, $\forall x \in \Omega$, $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$ such that $\Omega \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x) - x$ is a conic segment and $\text{span}(\mathcal{A}_{\Omega}(x)) = \mathcal{X}$. Let $\mathcal{C}_k(\Omega, \mathcal{Y})$ be the normed linear space defined in Example 9.69. We will show that $\mathcal{C}_k(\Omega, \mathcal{Y})$ is also a Banach space over $\mathbb{K}$.

Fix any Cauchy sequence $(f_n)_{n=1}^{\infty} \subseteq \mathcal{C}_k(\Omega, \mathcal{Y})$. By the definition of the norm $\|\cdot\|_{\mathcal{C}_k}$, $(f_n^{(i)})_{n=1}^{\infty} \subseteq \mathcal{C}(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y})) =: \mathcal{W}_i$ is a Cauchy sequence, $\forall i \in \{0, \dots, k\}$. $\forall i \in \{0, \dots, k\}$, by Example 7.32 and Proposition 7.66, $\mathcal{C}(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y}))$ is a Banach space. Then, $\exists g_i \in \mathcal{C}(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y}))$ such that $\lim_{n \in \mathbb{N}} f_n^{(i)} = g_i$ in $\mathcal{C}(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y}))$. Then, $(f_n^{(i)})_{n=1}^{\infty}$ converges uniformly to g_i and g_i is continuous. By Proposition 9.68, we have $g_i^{(1)} = g_{i+1}$, $\forall i \in \{0, \dots, k-1\}$. Then, we have $g_0^{(i)} = g_i$, $\forall i \in \{0, \dots, k\}$, and $g_0 \in \mathcal{C}_k(\Omega, \mathcal{Y})$. Then, $\|f_n - g_0\|_{\mathcal{C}_k} = \left(\sum_{i=0}^k \|f_n^{(i)} - g_0^{(i)}\|_{\mathcal{C}}^2\right)^{1/2} = \left(\sum_{i=0}^k \|f_n^{(i)} - g_i\|_{\mathcal{C}}^2\right)^{1/2} \rightarrow 0$ as $n \rightarrow \infty$, where the first equality follows from Proposition 9.40. Then, $\lim_{n \in \mathbb{N}} f_n = g_0$ in $\mathcal{C}_k(\Omega, \mathcal{Y})$. Hence, $\mathcal{C}_k(\Omega, \mathcal{Y})$ is complete and therefore a Banach space. $\diamond$

Sometimes, we need to consider a normed linear space of continuous functions on a topological space which is not necessarily compact. This leads us to the following examples.

Example 9.71 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $\mathcal{Y}$ be a normed linear space over the field $\mathbb{K}$, and $\mathcal{C}_v(\mathcal{X}, \mathcal{Y})$ be the vector space of all continuous functions of $\mathcal{X}$ to $\mathcal{Y}$ as defined in Example 7.50 with null vector ϑ . Define a function $\|\cdot\| : \mathcal{C}_v(\mathcal{X}, \mathcal{Y}) \rightarrow \mathbb{R}_e$ by $\|f\| = \max\{\sup_{x \in \mathcal{X}} \|f(x)\|_{\mathcal{Y}}, 0\}$, $\forall f \in \mathcal{C}_v(\mathcal{X}, \mathcal{Y})$. Consider the set $\mathcal{M} := \{f \in \mathcal{C}_v(\mathcal{X}, \mathcal{Y}) \mid \|f\| < +\infty\}$. Clearly, $\vartheta \in \mathcal{M}$. $\forall f_1, f_2 \in \mathcal{M}$, $\forall \alpha, \beta \in \mathbb{K}$, $\|\alpha f_1 + \beta f_2\| = \max\{\sup_{x \in \mathcal{X}} \|\alpha f_1(x) + \beta f_2(x)\|_{\mathcal{Y}}, 0\} \leq \max\{\sup_{x \in \mathcal{X}} (|\alpha| \|f_1(x)\|_{\mathcal{Y}} + |\beta| \|f_2(x)\|_{\mathcal{Y}}), 0\} < +\infty$. Then, $\alpha f_1 + \beta f_2 \in \mathcal{M}$. Hence, $\mathcal{M}$ is a subspace of $\mathcal{C}_v(\mathcal{X}, \mathcal{Y})$.

Clearly, $\forall f \in \mathcal{M}$, $\|f\| \in [0, \infty) \subset \mathbb{R}$ and $\|f\| = 0 \Leftrightarrow f = \vartheta$. $\forall f_1, f_2 \in \mathcal{M}$, $\forall \alpha \in \mathbb{K}$, $\|f_1 + f_2\| = \max\{\sup_{x \in \mathcal{X}} \|f_1(x) + f_2(x)\|_{\mathcal{Y}}, 0\} \leq \max\{\sup_{x \in \mathcal{X}} \|f_1(x)\|_{\mathcal{Y}} + \sup_{x \in \mathcal{X}} \|f_2(x)\|_{\mathcal{Y}}, 0\} = \|f_1\| + \|f_2\|$, where the first inequality follows from Proposition 3.81. $\|\alpha f_1\| = \max\{\sup_{x \in \mathcal{X}} \|\alpha f_1(x)\|_{\mathcal{Y}}, 0\} = \max\{\sup_{x \in \mathcal{X}} |\alpha| \|f_1(x)\|_{\mathcal{Y}}, 0\} = \begin{cases} \max\{|\alpha| \sup_{x \in \mathcal{X}} \|f_1(x)\|_{\mathcal{Y}}, 0\} & \alpha \neq 0 \\ 0 & \alpha = 0 \end{cases} = |\alpha| \|f_1\|$, where the third equality follows from Proposition 3.81. Hence, $(\mathcal{M}, \mathbb{K}, \|\cdot\|)$ is a normed linear space, which will be denoted by $\mathcal{C}_b(\mathcal{X}, \mathcal{Y})$. $\diamond$

Example 9.72 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $\mathcal{Y}$ be a Banach space over the field $\mathbb{K}$ (with norm $\|\cdot\|_{\mathcal{Y}}$). Consider the normed linear space $\mathcal{C}_b(\mathcal{X}, \mathcal{Y})$ (with norm $\|\cdot\|$) defined in Example 9.71. We will show that this space is a Banach space. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X} \neq \emptyset$. Case 1: $\mathcal{X} = \emptyset$. Then, $\mathcal{C}_b(\mathcal{X}, \mathcal{Y})$ is a singleton set. Hence, any Cauchy sequence in $\mathcal{C}_b(\mathcal{X}, \mathcal{Y})$ must converge. Thus, $\mathcal{C}_b(\mathcal{X}, \mathcal{Y})$ is a Banach space. Case 2: $\mathcal{X} \neq \emptyset$. Take a Cauchy sequence $(f_n)_{n=1}^{\infty} \subseteq \mathcal{C}_b(\mathcal{X}, \mathcal{Y})$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $\forall n, m \geq N$, $0 \leq \|f_n(x) - f_m(x)\|_{\mathcal{Y}} \leq \|f_n - f_m\| < \epsilon$, $\forall x \in \mathcal{X}$. This shows that, $\forall x \in \mathcal{X}$, $(f_n(x))_{n=1}^{\infty} \subseteq \mathcal{Y}$ is a Cauchy sequence, which converges to $f(x) \in \mathcal{Y}$ since $\mathcal{Y}$ is complete. This defines a function $f : \mathcal{X} \rightarrow \mathcal{Y}$. It is easy to show that $(f_n)_{n=1}^{\infty}$, viewed as a sequence of functions of $\mathcal{X}$ to $\mathcal{Y}$, converges uniformly to f . By Proposition 4.26, f is continuous. $\forall x \in \mathcal{X}$, $\|f(x)\|_{\mathcal{Y}} \leq \|f_N(x) - f(x)\|_{\mathcal{Y}} + \|f_N(x)\|_{\mathcal{Y}} = \lim_{m \in \mathbb{N}} \|f_N(x) - f_m(x)\|_{\mathcal{Y}} + \|f_N(x)\|_{\mathcal{Y}} \leq \epsilon + \|f_N\|$. Hence, $\|f\| \leq \|f_N\| + \epsilon$. Then, $f \in \mathcal{C}_b(\mathcal{X}, \mathcal{Y})$. It is easy to show that $\lim_{n \in \mathbb{N}} \|f_n - f\| = 0$. Hence, $\lim_{n \in \mathbb{N}} f_n = f$ in $\mathcal{C}_b(\mathcal{X}, \mathcal{Y})$. Hence, $\mathcal{C}_b(\mathcal{X}, \mathcal{Y})$ is a Banach space. In both cases, we have shown that $\mathcal{C}_b(\mathcal{X}, \mathcal{Y})$ is a Banach space when $\mathcal{Y}$ is a Banach space. $\diamond$

Example 9.73 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $\Omega \subseteq \mathcal{X}$ be endowed with the subset topology, which satisfies $\overline{\text{span}(A_{\Omega}(x))} = \mathcal{X}$, $\forall x \in \Omega$, $k \in \mathbb{N}$, and $\mathcal{W} := \mathcal{C}_b(\Omega, \mathcal{Y})$ be the normed linear space defined in Example 9.71 with null vector $\vartheta_{\mathcal{W}}$. Define $Z := \{f \in \mathcal{C}_b(\Omega, \mathcal{Y}) \mid f \text{ is } \mathcal{C}_k \text{ and } f^{(i)} \in \mathcal{C}_b(\Omega, B_i(\mathcal{X}, \mathcal{Y})), i = 1, \dots, k\}$. By Proposition 9.40, Z is a subspace of $\mathcal{W}$. Then, $\mathcal{Z} := (Z, \oplus_{\mathcal{W}}, \otimes_{\mathcal{W}}, \vartheta_{\mathcal{W}})$ is vector space over the field $\mathbb{K}$. $\forall f \in Z$, $\forall i \in \{0, \dots, k\}$, $f^{(i)} \in \mathcal{C}_b(\Omega, B_i(\mathcal{X}, \mathcal{Y}))$ with norm $\|f^{(i)}\|_{\mathcal{C}_b} \in \mathbb{R}$. Now, define a norm on Z by $\|f\|_{\mathcal{C}_{b,k}} := \left(\sum_{i=0}^k \|f^{(i)}\|_{\mathcal{C}_b}^2\right)^{1/2} \in [0, \infty) \subset \mathbb{R}$, $\forall f \in Z$. If $f = \vartheta_{\mathcal{W}}$, by Proposition 9.33, f is $\mathcal{C}_{\infty}$ and $f^{(i)}(x) = \vartheta_{B_i(\mathcal{X}, \mathcal{Y})}$, $\forall x \in \Omega$, $\forall i \in \mathbb{N}$, then, $\|f\|_{\mathcal{C}_{b,k}} = 0$. On the other hand, if $\|f\|_{\mathcal{C}_{b,k}} = 0$, then $f = \vartheta_{\mathcal{W}}$. $\forall f_1, f_2 \in Z$, by Proposition 9.40, $\|f_1 + f_2\|_{\mathcal{C}_{b,k}} = \left(\sum_{i=0}^k \|f_1^{(i)} + f_2^{(i)}\|_{\mathcal{C}_b}^2\right)^{1/2} \leq \left(\sum_{i=0}^k \|f_1^{(i)}\|_{\mathcal{C}_b}^2\right)^{1/2} + \left(\sum_{i=0}^k \|f_2^{(i)}\|_{\mathcal{C}_b}^2\right)^{1/2} = \|f_1\|_{\mathcal{C}_{b,k}} + \|f_2\|_{\mathcal{C}_{b,k}}$, where the inequality follows from Minkowski's Inequality, Theorem 7.9. $\forall \alpha \in \mathbb{K}$, $\forall f \in Z$, by Proposition 9.40, $\|\alpha f\|_{\mathcal{C}_{b,k}} = \left(\sum_{i=0}^k \|\alpha f^{(i)}\|_{\mathcal{C}_b}^2\right)^{1/2} = \left(\sum_{i=0}^k |\alpha|^2 \|f^{(i)}\|_{\mathcal{C}_b}^2\right)^{1/2} = |\alpha| \|f\|_{\mathcal{C}_{b,k}}$. This shows that $(\mathcal{Z}, \mathbb{K}, \|\cdot\|_{\mathcal{C}_{b,k}})$ is a normed linear space, which will be denoted $\mathcal{C}_{b,k}(\Omega, \mathcal{Y})$. $\diamond$

Example 9.74 Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$ and $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $\Omega \subseteq \mathcal{X}$ be endowed with the subset topology, and $k \in \mathbb{N}$. Assume that, $\forall x \in \Omega$, $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$ such that $\Omega \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x) - x$ is a conic segment and $\overline{\text{span}(A_{\Omega}(x))} = \mathcal{X}$. Let $\mathcal{C}_{b,k}(\Omega, \mathcal{Y})$ be the normed

linear space defined in Example 9.73. We will show that $\mathcal{C}_{b,k}(\Omega, \mathcal{Y})$ is also a Banach space over $\mathbb{K}$.

Fix any Cauchy sequence $(f_n)_{n=1}^\infty \subseteq \mathcal{C}_{b,k}(\Omega, \mathcal{Y})$. By the definition of the norm $\|\cdot\|_{\mathcal{C}_{b,k}}$, $(f_n^{(i)})_{n=1}^\infty \subseteq \mathcal{C}_b(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y})) =: \mathcal{W}_i$ is a Cauchy sequence, $\forall i \in \{0, \dots, k\}$. $\forall i \in \{0, \dots, k\}$, by Example 9.72 and Proposition 7.66, $\mathcal{C}_b(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y}))$ is a Banach space. Then, $\exists g_i \in \mathcal{C}_b(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y}))$ such that $\lim_{n \in \mathbb{N}} f_n^{(i)} = g_i$ in $\mathcal{C}_b(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y}))$. Then, $(f_n^{(i)})_{n=1}^\infty$ converges uniformly to g_i and g_i is continuous. By Proposition 9.68, we have $g_i^{(1)} = g_{i+1}$, $\forall i \in \{0, \dots, k-1\}$. Then, we have $g_0^{(i)} = g_i$, $\forall i \in \{0, \dots, k\}$, and $g_0 \in \mathcal{C}_{b,k}(\Omega, \mathcal{Y})$. Furthermore, $\|f_n - g_0\|_{\mathcal{C}_{b,k}} = \left(\sum_{i=0}^k \|f_n^{(i)} - g_0^{(i)}\|_{\mathcal{C}_b}^2 \right)^{1/2} = \left(\sum_{i=0}^k \|f_n^{(i)} - g_i\|_{\mathcal{C}_b}^2 \right)^{1/2} \rightarrow 0$ as $n \rightarrow \infty$, where the first equality follows from Proposition 9.40. Then, $\lim_{n \in \mathbb{N}} f_n = g_0$ in $\mathcal{C}_{b,k}(\Omega, \mathcal{Y})$. Hence, $\mathcal{C}_{b,k}(\Omega, \mathcal{Y})$ is complete and therefore a Banach space. $\diamond$

9.8 Tensor Algebra

Definition 9.75 Let $m \in \mathbb{Z}_+$, $\mathcal{X}_i$ be a normed linear space over $\mathbb{K}$, $i = 1, \dots, m$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$. A bounded linear operator $A \in \mathcal{B}(\mathcal{X}_m, \mathcal{B}(\mathcal{X}_{m-1}, \dots, \mathcal{B}(\mathcal{X}_1, \mathcal{Z}) \cdots))$ is said to be an m th order $\mathcal{Z}$ -valued tensor. Let $B \in \mathcal{B}(\mathcal{Y}_n, \mathcal{B}(\mathcal{Y}_{n-1}, \dots, \mathcal{B}(\mathcal{Y}_1, \mathcal{X}_m) \cdots))$ be another n th order $\mathcal{X}_m$ -valued tensor. We define $AB := A \cdot B \in \mathcal{B}(\mathcal{Y}_n, \dots, \mathcal{B}(\mathcal{Y}_1, \mathcal{B}(\mathcal{X}_{m-1}, \dots, \mathcal{B}(\mathcal{X}_1, \mathcal{Z}) \cdots)) \cdots)$ to be an $(n+m-1)$ st order $\mathcal{Z}$ -valued tensor such that $(AB)(y_n) \cdots (y_1)(x_{m-1}) \cdots (x_1) = A(B(y_n) \cdots (y_1))(x_{m-1}) \cdots (x_1) \in \mathcal{Z}$, $\forall y_i \in \mathcal{Y}_i$, $i = 1, \dots, n$, $\forall x_j \in \mathcal{X}_j$, $j = 1, \dots, m-1$. Let $(n_1, \dots, n_m)$ be any permutation of $(1, \dots, m)$. Then, we may define the transpose of tensor A with permutation $(n_1, \dots, n_m)$ to be $A^{T_{n_1, \dots, n_m}} \in \mathcal{B}(\mathcal{X}_{n_m}, \dots, \mathcal{B}(\mathcal{X}_{n_1}, \mathcal{Z}) \cdots)$ such that

$$A^{T_{n_1, \dots, n_m}}(x_{n_m}) \cdots (x_{n_1}) = A(x_m) \cdots (x_1) \quad \forall x_i \in \mathcal{X}_i, \quad i = 1, \dots, m$$

Proposition 9.76 Let $m, n \in \mathbb{Z}_+$, $\mathcal{X}_i$, $i = 1, \dots, m$, $\mathcal{Y}_j$, $j = 1, \dots, n$, $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $A, A_k \in \mathcal{W}_1 := \mathcal{B}(\mathcal{X}_m, \dots, \mathcal{B}(\mathcal{X}_1, \mathcal{Z}) \cdots)$, $k = 1, 2$, be m th order $\mathcal{Z}$ -valued tensors, $B, B_l \in \mathcal{W}_2 := \mathcal{B}(\mathcal{Y}_n, \dots, \mathcal{B}(\mathcal{Y}_1, \mathcal{X}_m) \cdots)$ be n th order $\mathcal{X}_m$ -valued tensors, $l = 1, 2$, and $\mathcal{W}_3 := \mathcal{B}(\mathcal{Y}_n, \dots, \mathcal{B}(\mathcal{Y}_1, \mathcal{B}(\mathcal{X}_{m-1}, \dots, \mathcal{B}(\mathcal{X}_1, \mathcal{Z}) \cdots)) \cdots)$. Then, the following statements hold.

$$(i) \quad \|AB\| \leq \|A\| \|B\|.$$

$$(ii) \quad \forall \alpha_k, \beta_l \in \mathbb{K}, \quad k = 1, 2, \quad l = 1, 2, \quad \text{we have } (\alpha_1 A_1 + \alpha_2 A_2)(\beta_1 B_1 + \beta_2 B_2) = \alpha_1 \beta_1 A_1 B_1 + \alpha_1 \beta_2 A_1 B_2 + \alpha_2 \beta_1 A_2 B_1 + \alpha_2 \beta_2 A_2 B_2.$$

- (iii) Let $A \in B(\mathcal{X}_2, B(\mathcal{X}_1, \mathbb{K}))$ be a second order $\mathbb{K}$ -valued tensor. Then, $A^{T_{2,1}} = A' \phi_{\mathcal{X}_1}$, where $\phi_{\mathcal{X}_1} : \mathcal{X}_1 \rightarrow \mathcal{X}_1^{**}$ is the natural mapping as defined in Remark 7.88.
- (iv) Let $f : \mathcal{W}_1 \times \mathcal{W}_2 \rightarrow \mathcal{W}_3$ be defined by $f(A, B) = AB \in \mathcal{W}_3$, $\forall A \in \mathcal{W}_1$, $\forall B \in \mathcal{W}_2$. Then, f is $\mathcal{C}_\infty$, $f^{(1)}(A_0, B_0)(\Delta_{1,1}, \Delta_{2,1}) = A_0 \Delta_{2,1} + \Delta_{1,1} B_0$, $f^{(2)}(A_0, B_0)(\Delta_{1,1}, \Delta_{2,1})(\Delta_{1,2}, \Delta_{2,2}) = \Delta_{1,2} \Delta_{2,1} + \Delta_{1,1} \Delta_{2,2}$, and $f^{(i+2)}(A_0, B_0) = \vartheta_{B_{S_{i+2}}(\mathcal{W}_1 \times \mathcal{W}_2, \mathcal{W}_3)}$, $\forall (A_0, B_0) \in \mathcal{W}_1 \times \mathcal{W}_2$, $\forall (\Delta_{1,1}, \Delta_{2,1}) \in \mathcal{W}_1 \times \mathcal{W}_2$, $\forall (\Delta_{1,2}, \Delta_{2,2}) \in \mathcal{W}_1 \times \mathcal{W}_2$, $\forall i \in \mathbb{N}$.
- (v) Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $x_0 \in D \subseteq \mathcal{X}$, $A : D \rightarrow \mathcal{W}_1$ and $B : D \rightarrow \mathcal{W}_2$ be tensor-valued functions that are Fréchet differentiable at x_0 , and $C : D \rightarrow \mathcal{W}_3$ be defined by $C(x) = A(x)B(x)$, $\forall x \in D$. Then,

$$\begin{aligned} C^{(1)}(x_0) &= \left((A^{(1)}(x_0))^{T_{1,\dots,m-1,m+1,m}} B(x_0) \right)^{T_{1,\dots,m-1,m+1,\dots,n+m,m}} \\ &\quad + A(x_0) B^{(1)}(x_0) \end{aligned}$$

- (vi) Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $x_0 \in D \subseteq \mathcal{X}$, $A : D \rightarrow \mathcal{W}_1$ be a tensor-valued function that is Fréchet differentiable at x_0 , $(n_1, \dots, n_m)$ be a permutation of $(1, \dots, m)$, and $C : D \rightarrow \mathcal{W}_p := B(\mathcal{X}_{n_m}, \dots, B(\mathcal{X}_{n_1}, \mathcal{Z}) \cdots)$ be defined by $C(x) = (A(x))^{T_{n_1, \dots, n_m}}$, $\forall x \in D$. Then, we have

$$C^{(1)}(x_0) = (A^{(1)}(x_0))^{T_{n_1, \dots, n_m, m+1}}$$

Proof (i) and (ii) These are straightforward, and are therefore omitted.

(iii) Note that $A \in B(\mathcal{X}_2, \mathcal{X}_1^*)$. Then, $A' \in B(\mathcal{X}_1^{**}, \mathcal{X}_2^*)$ and $A^{T_{2,1}} \in B(\mathcal{X}_1, B(\mathcal{X}_2, \mathbb{K})) = B(\mathcal{X}_1, \mathcal{X}_2^*)$. Then, $A^{T_{2,1}}(x_1)(x_2) = A(x_2)(x_1) = \langle\langle A(x_2), x_1 \rangle\rangle = \langle\langle \phi_{\mathcal{X}_1}(x_1), A(x_2) \rangle\rangle = \langle\langle A'(\phi_{\mathcal{X}_1}(x_1)), x_2 \rangle\rangle = A'(\phi_{\mathcal{X}_1}(x_1))(x_2)$, $\forall x_i \in \mathcal{X}_i$, $i = 1, 2$. Then, we have $A^{T_{2,1}} = A' \phi_{\mathcal{X}_1}$.

(iv) is straightforward, and is therefore omitted.

(v) follows directly from (iv), the Chain Rule and Proposition 9.19.

(vi) is straightforward, and is therefore omitted. $\square$

Definition 9.77 Let $\mathcal{X}_i$, $i = 1, \dots, m$, $\mathcal{Z}$, $\mathcal{Y}_j$, $j = 1, \dots, n$, be normed linear spaces over $\mathbb{K}$, $A \in B(\mathcal{X}_m, \dots, B(\mathcal{X}_1, \mathcal{Z}) \cdots)$ be an m th order $\mathcal{Z}$ -valued tensor, and $B \in B(\mathcal{Y}_n, \dots, B(\mathcal{Y}_1, \mathbb{K}) \cdots)$ be an n th order $\mathbb{K}$ -valued tensor. Define the outer product of A and B to be an $(n+m)$ th order $\mathcal{Z}$ -valued tensor $C := A \otimes B \in B(\mathcal{Y}_n, \dots, B(\mathcal{Y}_1, B(\mathcal{X}_m, \dots, B(\mathcal{X}_1, \mathcal{Z}) \cdots)) \cdots)$ such that $C(y_n) \cdots (y_1)(x_m) \cdots (x_1) = B(y_n) \cdots (y_1)A(x_m) \cdots (x_1) \in \mathcal{Z}$, $\forall x_i \in \mathcal{X}_i$, $i = 1, \dots, m$, $\forall y_j \in \mathcal{Y}_j$, $j = 1, \dots, n$. Similarly, we may define the outer product of B and A to be an $(n+m)$ th order $\mathcal{Z}$ -valued tensor $\bar{C} := B \otimes A \in B(\mathcal{X}_m, \dots, B(\mathcal{X}_1, B(\mathcal{Y}_n, \dots, B(\mathcal{Y}_1, \mathcal{Z}) \cdots)) \cdots)$ such that

$$\bar{C}(x_m) \cdots (x_1)(y_n) \cdots (y_1) = B(y_n) \cdots (y_1)A(x_m) \cdots (x_1) \in \mathcal{Z}, \forall x_i \in \mathcal{X}_i, \\ i = 1, \dots, m, \forall y_j \in \mathcal{Y}_j, j = 1, \dots, n.$$

Proposition 9.78 *Let $m, n \in \mathbb{Z}_+$, $\mathcal{X}_i, i = 1, \dots, m, \mathcal{Y}_j, j = 1, \dots, n, \mathcal{Z}_\tau, \tau = 1, 2$, be normed linear spaces over $\mathbb{K}$ with $\mathcal{Z}_1 = \mathbb{K}$ or $\mathcal{Z}_2 = \mathbb{K}$, $\mathcal{Z} = \mathcal{Z}_1$ if $\mathcal{Z}_2 = \mathbb{K}$, $\mathcal{Z} = \mathcal{Z}_2$ if $\mathcal{Z}_1 = \mathbb{K}$, $A, A_k \in \mathcal{W}_1 := \mathcal{B}(\mathcal{X}_m, \dots, \mathcal{B}(\mathcal{X}_1, \mathcal{Z}_1) \cdots)$, $k = 1, 2$, be m th order $\mathcal{Z}_1$ -valued tensors, $B, B_l \in \mathcal{W}_2 := \mathcal{B}(\mathcal{Y}_n, \dots, \mathcal{B}(\mathcal{Y}_1, \mathcal{Z}_2) \cdots)$ be n th order $\mathcal{Z}_2$ -valued tensors, $l = 1, 2$, and $\mathcal{W}_3 := \mathcal{B}(\mathcal{Y}_n, \dots, \mathcal{B}(\mathcal{Y}_1, \mathcal{B}(\mathcal{X}_m, \dots, \mathcal{B}(\mathcal{X}_1, \mathcal{Z}) \cdots)) \cdots)$. Then, the following statements hold.*

$$(i) \|A \otimes B\| \leq \|A\| \|B\|.$$

$$(ii) \forall \alpha_k, \beta_l \in \mathbb{K}, k = 1, 2, l = 1, 2, \text{ we have } (\alpha_1 A_1 + \alpha_2 A_2) \otimes (\beta_1 B_1 + \beta_2 B_2) = \alpha_1 \beta_1 A_1 \otimes B_1 + \alpha_1 \beta_2 A_1 \otimes B_2 + \alpha_2 \beta_1 A_2 \otimes B_1 + \alpha_2 \beta_2 A_2 \otimes B_2.$$

$$(iii) \text{ Let } f : \mathcal{W}_1 \times \mathcal{W}_2 \rightarrow \mathcal{W}_3 \text{ be defined by } f(A, B) = A \otimes B \in \mathcal{W}_3, \forall A \in \mathcal{W}_1, \forall B \in \mathcal{W}_2. \text{ Then, } f \text{ is } \mathcal{C}_\infty, f^{(1)}(A_0, B_0)(\Delta_{1,1}, \Delta_{2,1}) = A_0 \otimes \Delta_{2,1} + \Delta_{1,1} \otimes B_0, f^{(2)}(A_0, B_0)(\Delta_{1,1}, \Delta_{2,1})(\Delta_{1,2}, \Delta_{2,2}) = \Delta_{1,2} \otimes \Delta_{2,1} + \Delta_{1,1} \otimes \Delta_{2,2}, \text{ and } f^{(i+2)}(A_0, B_0) = \vartheta_{\mathcal{B}_{S^{i+2}}(\mathcal{W}_1 \times \mathcal{W}_2, \mathcal{W}_3)}, \forall (A_0, B_0) \in \mathcal{W}_1 \times \mathcal{W}_2, \forall (\Delta_{1,1}, \Delta_{2,1}) \in \mathcal{W}_1 \times \mathcal{W}_2, \forall (\Delta_{1,2}, \Delta_{2,2}) \in \mathcal{W}_1 \times \mathcal{W}_2, \forall i \in \mathbb{N}.$$

$$(iv) \text{ Let } \mathcal{X} \text{ be a normed linear space over } \mathbb{K}, x_0 \in D \subseteq \mathcal{X}, A : D \rightarrow \mathcal{W}_1 \text{ and } B : D \rightarrow \mathcal{W}_2 \text{ be tensor-valued functions that are Fréchet differentiable at } x_0, \text{ and } C : D \rightarrow \mathcal{W}_3 \text{ be defined by } C(x) = A(x) \otimes B(x), \forall x \in D. \text{ Then,}$$

$$C^{(1)}(x_0) = (A^{(1)}(x_0) \otimes B(x_0))^{T_{1, \dots, m, m+2, \dots, n+m+1, m+1}} + A(x_0) \otimes B^{(1)}(x_0)$$

Proof These are straightforward, and are therefore omitted. $\square$

9.9 Analytic Functions

Definition 9.79 *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, and $x_0 \in D^\circ$. f is said to be analytic at x_0 if f is $\mathcal{C}_\infty$ at x_0 , and $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\exists c \in [0, \infty) \subset \mathbb{R}$, and $\exists M \in (0, \infty) \subset \mathbb{R}$, such that $\|f^{(n)}(x)\| \leq cn!M^n, \forall n \in \mathbb{Z}_+, \forall x \in \mathcal{B}_\mathcal{X}(x_0, \delta)$. In this case, $\forall \bar{\delta} \in (0, \delta] \subset \mathbb{R}$ with $M\bar{\delta} < 1$ is called an analytic radius of f at x_0 . If $D \supseteq D_1 \in \mathcal{O}_\mathcal{X}$ and f is analytic at $x, \forall x \in D_1$, then we say that f is analytic on D_1 . When f is analytic on D , then we say that f is analytic or an analytic function.*

Theorem 9.80 (Taylor Series) *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \rightarrow \mathcal{Y}$, and $x_0 \in D^\circ$. Assume that f is analytic*

at x_0 with an analytic radius $\delta \in (0, \infty) \subset \mathbb{R}$. Then, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta)$,

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0) \underbrace{(x - x_0) \cdots (x - x_0)}_{n\text{-times}}$$

Proof Since f is analytic at $x_0 \in D^\circ$ with an analytic radius $\delta \in (0, \infty) \subset \mathbb{R}$, then $\exists c \in [0, \infty) \subset \mathbb{R}$ and $\exists M \in (0, \infty) \subset \mathbb{R}$ such that $\|f^{(n)}(x)\| \leq cn!M^n$, $\forall n \in \mathbb{Z}_+$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta) \subseteq D$, and $M\delta < 1$. Then, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\forall n \in \mathbb{Z}_+$, let

$$R_n(x, x_0; f) := f(x) - \sum_{i=0}^n \frac{1}{i!} f^{(i)}(x_0) \underbrace{(x - x_0) \cdots (x - x_0)}_{i\text{-times}}$$

By Taylor Theorem 9.48, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ such that $\|R_n(x, x_0; f)\| \leq \frac{1}{(n+1)!} \|f^{(n+1)}(t_0x + (1 - t_0)x_0)\| \|x - x_0\|^{n+1} < c(M\delta)^{n+1}$. Clearly, $\lim_{n \in \mathbb{N}} \|R_n(x, x_0; f)\| = 0$. Hence, the desired equality holds. This completes the proof of the theorem. $\square$

Next, we establish a result that is needed in the further analysis of analytic functions.

Proposition 9.81 *Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$ be normed linear spaces over $\mathbb{K}$, $D_1 \subseteq \mathcal{X}$, $f : D_1 \rightarrow \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $g : D_1 \rightarrow \mathcal{B}(\mathcal{Z}, \mathcal{W})$, $n \in \mathbb{Z}_+$, and $x_0 \in D_1^\circ$. Assume that f and g are $\mathcal{C}_n$ at x_0 . Define $h : D_1 \rightarrow \mathcal{B}(\mathcal{Y}, \mathcal{W})$ by $h(x) = g(x)f(x)$, $\forall x \in D_1$. Then, h is $\mathcal{C}_n$ at x_0 , and $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that*

$$\begin{aligned} h^{(k)}(x)(\xi_{x_k}) \cdots (\xi_{x_1})(\xi_y) = \\ \sum_{i=0}^k \sum_{\substack{\Lambda_1 := \{j_1, \dots, j_i\} \text{ is any } i\text{-element subset of} \\ \Lambda := \{1, \dots, k\}; \Lambda_2 := \{j_{i+1}, \dots, j_k\} := \Lambda \setminus \Lambda_1}} g^{(i)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_i}}) \\ (f^{(k-i)}(x)(\xi_{x_{j_{i+1}}}) \cdots (\xi_{x_{j_k}})(\xi_y)) \end{aligned} \quad (9.1)$$

$\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\forall \xi_{x_l} \in \mathcal{X}$, $l = 1, \dots, k$, $\forall \xi_y \in \mathcal{Y}$, $\forall k \in \mathbb{Z}_+$ with $k \leq n$.

Proof By Propositions 9.45, 9.44, 3.12, and 9.42, h is $\mathcal{C}_n$ at x_0 . Since f and g are $\mathcal{C}_n$ at x_0 , then $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $f^{(n)}(x) \in \mathcal{B}_{\mathcal{S}_n}(\mathcal{X}, \mathcal{B}(\mathcal{Y}, \mathcal{Z}))$ and $g^{(n)}(x) \in \mathcal{B}_{\mathcal{S}_n}(\mathcal{X}, \mathcal{B}(\mathcal{Z}, \mathcal{W}))$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta)$. We will use mathematical induction on k to prove the result.

1° $k = 0$. We have $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\forall \xi_y \in \mathcal{Y}$,

$$\text{LHS} = h^{(0)}(x)(\xi_y) = h(x)(\xi_y) = g(x)f(x)\xi_y = g(x)(f(x)(\xi_y)) = \text{RHS}$$

This case is proved. If $n = 0$, the proposition holds. If $n \in \mathbb{N}$, continue to the following steps.

2° Assume that the result holds for $k = l \in \mathbb{Z}_+$ with $l < n$.

3° Consider the case when $k = l + 1$. By the inductive assumption, $\forall x \in \mathcal{B}_X(x_0, \delta)$, $\forall \xi_{x_s} \in \mathcal{X}$, $s = 1, \dots, l$, $\forall \xi_y \in \mathcal{Y}$,

$$\begin{aligned} h^{(l)}(x)(\xi_{x_l}) \cdots (\xi_{x_1})(\xi_y) = \\ \sum_{i=0}^l \sum_{\substack{\Lambda_1 := \{j_1, \dots, j_i\} \text{ is any } i\text{-element subset of} \\ \Lambda := \{1, \dots, l\}; \Lambda_2 := \{j_{i+1}, \dots, j_l\} := \Lambda \setminus \Lambda_1}} g^{(i)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_i}}) \\ (f^{(l-i)}(x)(\xi_{x_{j_{i+1}}}) \cdots (\xi_{x_{j_l}})(\xi_y)) \end{aligned}$$

Then, by Proposition 9.27, we have $\forall x \in \mathcal{B}_X(x_0, \delta)$, $\forall \xi_{x_s} \in \mathcal{X}$, $s = 1, \dots, l+1$, $\forall \xi_y \in \mathcal{Y}$,

$$h^{(l+1)}(x)(\xi_{x_{l+1}}) \cdots (\xi_{x_1}) = \bar{h}^{(1)}(x)(\xi_{x_{l+1}})$$

where $\bar{h} : \mathcal{B}_X(x_0, \delta) \rightarrow \mathcal{B}(\mathcal{Y}, \mathcal{W})$ is defined by $\bar{h}(x) = h^{(l)}(x)(\xi_{x_l}) \cdots (\xi_{x_1})$, $\forall x \in \mathcal{B}_X(x_0, \delta)$. This implies that, by the Chain Rule (Theorem 9.18) and Proposition 9.42,

$$\begin{aligned} & h^{(l+1)}(x)(\xi_{x_{l+1}}) \cdots (\xi_{x_1})(\xi_y) \\ &= \sum_{i=0}^l \sum_{\substack{\Lambda_1 := \{j_1, \dots, j_i\} \text{ is any } i\text{-element subset of} \\ \Lambda := \{1, \dots, l\}; \Lambda_2 := \{j_{i+1}, \dots, j_l\} := \Lambda \setminus \Lambda_1}} \left(g^{(i+1)}(x)(\xi_{x_{l+1}})(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_i}}) \right. \\ & \quad (f^{(l-i)}(x)(\xi_{x_{j_{i+1}}}) \cdots (\xi_{x_{j_l}})(\xi_y)) + g^{(i)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_i}}) \\ & \quad \left. (f^{(l-i+1)}(x)(\xi_{x_{l+1}})(\xi_{x_{j_{i+1}}}) \cdots (\xi_{x_{j_l}})(\xi_y)) \right) \\ &= \sum_{i=0}^{l+1} \left(\sum_{\substack{\Lambda_1 := \{j_1, \dots, j_i\} \text{ is any } i\text{-element subset of} \\ \Lambda := \{1, \dots, l\}; \Lambda_2 := \{j_{i+1}, \dots, j_l\} := \Lambda \setminus \Lambda_1}} g^{(i)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_i}}) \right. \\ & \quad (f^{(l+1-i)}(x)(\xi_{x_{l+1}})(\xi_{x_{j_{i+1}}}) \cdots (\xi_{x_{j_l}})(\xi_y)) \\ & \quad + \sum_{\substack{\Lambda_1 := \{j_1, \dots, j_{i-1}\} \text{ is any } (i-1)\text{-element subset of} \\ \Lambda := \{1, \dots, l\}; \Lambda_2 := \{j_i, \dots, j_l\} := \Lambda \setminus \Lambda_1}} g^{(i)}(x)(\xi_{x_{l+1}})(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_{i-1}}}) \\ & \quad \left. (f^{(l+1-i)}(x)(\xi_{x_{j_i}}) \cdots (\xi_{x_{j_l}})(\xi_y)) \right) \\ &= \sum_{i=0}^{l+1} \sum_{\substack{\Lambda_1 := \{j_1, \dots, j_i\} \text{ is any } i\text{-element subset of} \\ \Lambda := \{1, \dots, l+1\}; \Lambda_2 := \{j_{i+1}, \dots, j_{l+1}\} := \Lambda \setminus \Lambda_1}} g^{(i)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_i}}) \\ & \quad (f^{(l+1-i)}(x)(\xi_{x_{j_{i+1}}}) \cdots (\xi_{x_{j_{l+1}}})(\xi_y)) \end{aligned}$$

where the last equality follows since all i -element subsets of $\{1, \dots, l+1\}$ is divided into two disjoint classes: one with $l+1$ as an element; and the other without $l+1$ as an element. This completes the induction process.

This completes the proof of the proposition. $\square$

Now, we establish the important result that the composition of analytic functions is an analytic function.

Theorem 9.82 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D_1 \subseteq \mathcal{X}$, $D_2 \subseteq \mathcal{Y}$, $f : D_1 \rightarrow D_2$, $g : D_2 \rightarrow \mathcal{Z}$, $x_0 \in D_1^\circ$, and $y_0 := f(x_0) \in D_2^\circ$. Assume that f and g are analytic at x_0 and y_0 , respectively. Then, $h := g \circ f : D_1 \rightarrow \mathcal{Z}$ is analytic at x_0 .*

Proof Since g is analytic at y_0 , then $\exists \delta_g \in (0, \infty) \subset \mathbb{R}$, $\exists c_g \in [0, \infty) \subset \mathbb{R}$, and $\exists M_g \in (0, \infty) \subset \mathbb{R}$ such that $\|g^{(n)}(y)\| \leq c_g n! M_g^n$, $\forall n \in \mathbb{Z}_+$, $\forall y \in \mathcal{B}_\mathcal{Y}(y_0, \delta_g) \subseteq D_2$. Since f is analytic at x_0 , then $\exists \delta_f \in (0, \infty) \subset \mathbb{R}$, $\exists c_f \in [0, \infty) \subset \mathbb{R}$, and $\exists M_f \in (0, \infty) \subset \mathbb{R}$ such that $\|f^{(n)}(x)\| \leq c_f n! M_f^n$ and $f(x) \in \mathcal{B}_\mathcal{Y}(y_0, \delta_g)$, $\forall n \in \mathbb{Z}_+$, $\forall x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)$. By Proposition 9.45, h is $\mathcal{C}_\infty$ at $\forall x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)$.

Define nonnegative constants $K_{n,i} := \sup_{x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)} \|D^n(g^{(i)} \circ f)(x)\|$, $\forall i, n \in \mathbb{Z}_+$. Clearly, $\forall i \in \mathbb{Z}_+$,

$$K_{0,i} = \sup_{x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)} \|g^{(i)} \circ f(x)\| \leq \sup_{y \in \mathcal{B}_\mathcal{Y}(y_0, \delta_g)} \|g^{(i)}(y)\| \leq c_g i! M_g^i$$

Note that, $\forall i, n \in \mathbb{Z}_+$, by Chain Rule (Theorem 9.18),

$$\begin{aligned} K_{n+1,i} &= \sup_{x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)} \|D^{n+1}(g^{(i)} \circ f)(x)\| \\ &= \sup_{x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)} \|D^n((g^{(i+1)} \circ f)(f^{(1)}))(x)\| \end{aligned}$$

Define $l_{i+1} : \mathcal{B}_\mathcal{X}(x_0, \delta_f) \rightarrow \mathcal{B}_{S_{i+1}}(\mathcal{Y}, \mathcal{Z})$ by $l_{i+1}(x) := g^{(i+1)}(f(x))$, $\forall x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)$, $\forall i \in \mathbb{Z}_+$. Then, $\forall i, n \in \mathbb{Z}_+$,

$$\begin{aligned} K_{n+1,i} &= \sup_{x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)} \|D^n((g^{(i+1)} \circ f)(f^{(1)}))(x)\| \\ &= \sup_{x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)} \|D^n(l_{i+1} f^{(1)})(x)\| \\ &= \sup_{x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)} \sup_{\substack{\xi_{x_s} \in \mathcal{X}, s=1, \dots, n, \xi_x \in \mathcal{X}, \\ \|\xi_{x_s}\| \leq 1, \|\xi_x\| \leq 1}} \|D^n(l_{i+1} f^{(1)})(x)(\xi_{x_n}) \cdots (\xi_{x_1})(\xi_x)\| \\ &= \sup_{x \in \mathcal{B}_\mathcal{X}(x_0, \delta_f)} \sup_{\substack{\xi_{x_s} \in \mathcal{X}, s=1, \dots, n, \xi_x \in \mathcal{X}, \\ \|\xi_{x_s}\| \leq 1, \|\xi_x\| \leq 1}} \left\| \sum_{k=0}^n l_{i+1}^{(k)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_k}}) \right. \\ &\quad \left. \sum_{\substack{\Lambda_1 := \{j_1, \dots, j_k\} \text{ is any } k\text{-element subset of} \\ \Lambda := \{1, \dots, n\}; \Lambda_2 := \{j_{k+1}, \dots, j_n\} := \Lambda \setminus \Lambda_1}} (f^{(n-k+1)}(x)(\xi_{x_{j_{k+1}}}) \cdots (\xi_{x_{j_n}})(\xi_x)) \right\| \end{aligned}$$

$$\begin{aligned}
&\leq \sup_{x \in \mathcal{B}_X(x_0, \delta_f)} \sum_{k=0}^n \sum_{\substack{\Lambda_1 := \{j_1, \dots, j_k\} \text{ is any } k\text{-element subset of} \\ \Lambda := \{1, \dots, n\}; \Lambda_2 := \{j_{k+1}, \dots, j_n\} := \Lambda \setminus \Lambda_1}} \left\| l_{i+1}^{(k)}(x) \right\| \cdot \\
&\quad \left\| f^{(n-k+1)}(x) \right\| \\
&= \sup_{x \in \mathcal{B}_X(x_0, \delta_f)} \sum_{k=0}^n \binom{n}{k} \left\| D^k(g^{(i+1)} \circ f)(x) \right\| \left\| f^{(n-k+1)}(x) \right\| \\
&\leq \sum_{k=0}^n \binom{n}{k} K_{k,i+1} c_f (n+1-k)! M_f^{n+1-k}
\end{aligned}$$

where the fourth equality follows from Proposition 9.81. Hence, $0 \leq K_{n+1,i} \leq \sum_{k=0}^n \binom{n}{k} K_{k,i+1} c_f (n-k+1)! M_f^{n-k+1}$, $\forall i, n \in \mathbb{Z}_+$.

We will use mathematical induction on $n \in \mathbb{N}$ to show that $\forall n \in \mathbb{N}$, $\forall i \in \mathbb{Z}_+$,

$$0 \leq K_{n,i} \leq (n+i)! c_g M_g^{i+1} M_f^n \sum_{j=0}^{n-1} \binom{n}{j+1} \binom{n-1}{j} \binom{n+i}{n-j-1}^{-1} c_f^{j+1} M_g^j$$

1° $n = 1$. By the recursive formula and the inequality for $K_{0,i}$, $\forall i \in \mathbb{Z}_+$,

$$K_{1,i} \leq K_{0,i+1} c_f M_f \leq (i+1)! c_f M_f c_g M_g^{i+1} = \text{RHS}$$

This case is proved.

2° Assume that the result holds for $n = 1, \dots, l \in \mathbb{N}$.

3° Consider the case $n = l+1$. $\forall i \in \mathbb{Z}_+$, by the recursive formula,

$$\begin{aligned}
K_{l+1,i} &\leq \sum_{k=0}^l \binom{l}{k} K_{k,i+1} c_f (l-k+1)! M_f^{l-k+1} \\
&\leq c_f (l+1)! M_f^{l+1} c_g (i+1)! M_g^{i+1} + \sum_{k=1}^l \binom{l}{k} c_f (l-k+1)! M_f^{l-k+1} \\
&\quad \cdot (k+i+1)! c_g M_g^{i+2} M_f^k \sum_{j=0}^{k-1} \binom{k}{j+1} \binom{k-1}{j} \binom{k+i+1}{k-j-1}^{-1} c_f^{j+1} M_g^j \\
&= c_f c_g M_g^{i+1} M_f^{l+1} (l+i+1)! (l+1) \binom{l+i+1}{l}^{-1} \\
&\quad + \sum_{k=1}^l \sum_{j=0}^{k-1} \frac{l!(l-k+1)! k!(k-1)!(i+j+2)!}{k!(l-k)! j!(j+1)!(k-j-1)!} c_f^{j+2} c_g M_f^{l+1} M_g^{j+i+2} \\
&= c_f c_g M_g^{i+1} M_f^{l+1} (l+i+1)! (l+1) \binom{l+i+1}{l}^{-1} \\
&\quad + \sum_{j=0}^{l-1} \sum_{k=j+1}^l \frac{l!(l-k+1)! (k-1)!(i+j+2)!}{(l-k)! j!(j+1)!(k-j-1)!} c_f^{j+2} c_g M_f^{l+1} M_g^{j+i+2}
\end{aligned}$$

$$\begin{aligned}
&= (l+i+1)!c_g M_g^{i+1} M_f^{l+1} \left(c_f(l+1) \binom{l+i+1}{l} \right)^{-1} \\
&\quad + \sum_{j=0}^{l-1} \sum_{k=j+1}^l \frac{(l-k+1)l!(k-1)!(i+j+2)!}{j!(j+1)!(k-j-1)!(l+i+1)!} c_f^{j+2} M_g^{j+1} \Big) \\
&= (l+i+1)!c_g M_g^{i+1} M_f^{l+1} \left(c_f(l+1) \binom{l+i+1}{l} \right)^{-1} \\
&\quad + \sum_{j=0}^{l-1} \sum_{k=j+1}^l (l-k+1) \binom{k-1}{j} \binom{l}{j+1} \binom{l+i+1}{l-j-1}^{-1} c_f^{j+2} M_g^{j+1} \Big) \\
&= (l+i+1)!c_g M_g^{i+1} M_f^{l+1} \left(c_f(l+1) \binom{l+i+1}{l} \right)^{-1} \\
&\quad + \sum_{j=1}^l \binom{l}{j} \binom{l+i+1}{l-j}^{-1} c_f^{j+1} M_g^j \sum_{k=j}^l (l-k+1) \binom{k-1}{j-1} \Big)
\end{aligned}$$

where the second inequality follows from the inductive assumption and the bound for $K_{0,i}$. Note that $\sum_{k=j}^l (l-k+1) \binom{k-1}{j-1} = \sum_{k=j}^l (l-k+1) \binom{k-1}{j-1} = \sum_{k=j}^l \left(\binom{k}{j} - \binom{k-1}{j} \right) = \sum_{k=j}^l (l-k+1) \binom{k}{j} - \sum_{k=j-1}^{l-1} (l-k) \binom{k}{j} = \sum_{k=j}^l \binom{k}{j} = \sum_{k=j}^l \left(\binom{k+1}{j+1} - \binom{k}{j+1} \right) = \binom{l+1}{j+1}$, $\forall 1 \leq j \leq l$. This leads to

$$\begin{aligned}
K_{l+1,i} &\leq (l+i+1)!c_g M_g^{i+1} M_f^{l+1} \left(c_f(l+1) \binom{l+i+1}{l} \right)^{-1} \\
&\quad + \sum_{j=1}^l \binom{l}{j} \binom{l+1}{j+1} \binom{l+i+1}{l-j}^{-1} c_f^{j+1} M_g^j \Big) \\
&= (l+i+1)!c_g M_g^{i+1} M_f^{l+1} \sum_{j=0}^l \binom{l}{j} \binom{l+1}{j+1} \binom{l+i+1}{l-j}^{-1} c_f^{j+1} M_g^j \\
&= \text{RHS}
\end{aligned}$$

This completes the induction process.

Hence, $\forall n \in \mathbb{N}$,

$$\begin{aligned}
\sup_{x \in \mathcal{B}_X(x_0, \delta_f)} \|h^{(n)}(x)\| &= \sup_{x \in \mathcal{B}_X(x_0, \delta_f)} \|D^n(g \circ f)(x)\| = K_{n,0} \\
&\leq n!c_g M_g M_f^n \sum_{j=0}^{n-1} \binom{n}{j+1} \binom{n-1}{j} \binom{n}{n-j-1}^{-1} c_f^{j+1} M_g^j \\
&= n!c_g M_g M_f^n \sum_{j=0}^{n-1} \binom{n-1}{j} c_f^{j+1} M_g^j = n!c_f c_g M_g M_f^n (1 + c_f M_g)^{n-1}
\end{aligned}$$

$$= n! \frac{c_f c_g M_g}{1 + c_f M_g} (M_f (1 + c_f M_g))^n$$

Hence, h is analytic at x_0 . This completes the proof of the theorem. $\square$

Proposition 9.83 *The constant mapping as defined in Proposition 9.33, the identity mapping as delineated in Proposition 9.36, the projection mapping as delineated in Proposition 9.37, the vector addition, the scalar multiplication, the operation of a linear operator on a vector as delineated in Proposition 9.41, the operation of composition of two linear operators as delineated in Proposition 9.42, and the building up of linear operator on the product space using linear operators on individual spaces as delineated in Proposition 9.43 are analytic on the interior of the domain of the mapping.*

Proof This is immediate from Definition 9.79. $\square$

Proposition 9.84 *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D := \{L \in B(\mathcal{X}, \mathcal{Y}) \mid L \text{ is bijective.}\}$, and $f : D \rightarrow B(\mathcal{Y}, \mathcal{X})$ be defined by $f(A) = A^{-1}$, $\forall A \in D \subseteq B(\mathcal{X}, \mathcal{Y})$. Then, f is analytic.*

Proof By Proposition 9.56, D is open in $B(\mathcal{X}, \mathcal{Y})$, f is $\mathcal{C}_\infty$, and $f^{(1)}(A)(\Delta) = -f(A)\Delta f(A)$, $\forall A \in D$, $\forall \Delta \in B(\mathcal{X}, \mathcal{Y})$. Fix any $A \in D$. Since D is open, then $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $U := \mathcal{B}_{B(\mathcal{X}, \mathcal{Y})}(A, \delta) \subseteq D$. Without loss of generality, assume that $\delta \|f(A)\| < 1$. By Proposition 9.55, $\forall T \in U$, we have $\|f(T) - f(A)\| \leq \frac{\|f(A)\|^2 \|T - A\|}{1 - \|f(A)\| \|T - A\|} \leq \|f(A)\|^2 \delta (1 - \|f(A)\| \delta)^{-1}$. This implies that $\|f(T)\| \leq \|f(A)\| + \|f(A)\|^2 \delta (1 - \|f(A)\| \delta)^{-1} = \|f(A)\| (1 - \delta \|f(A)\|)^{-1} =: \alpha \in [0, \infty) \subset \mathbb{R}$. This leads to

$$\|f^{(1)}(T)\| = \sup_{\Delta \in B(\mathcal{X}, \mathcal{Y}), \|\Delta\| \leq 1} \|f^{(1)}(T)(\Delta)\| \leq \|f(T)\|^2 \leq \alpha^2$$

We will use mathematical induction on $n \in \mathbb{Z}_+$ to show that $\|f^{(n)}(T)\| \leq n! \alpha^{n+1}$, $\forall n \in \mathbb{Z}_+$, $\forall T \in U$.

1° Clearly, $n = 0$ and $n = 1$ cases are proved.

2° Assume that the result holds for $n = 0, \dots, k \in \mathbb{N}$.

3° Consider the case when $n = k + 1$. $\forall T \in U$, $\forall \Delta_1 \in B(\mathcal{X}, \mathcal{Y})$, $f^{(1)}(T)(\Delta_1) = -f(T)\Delta_1 f(T) =: g_{\Delta_1}(T)$, where $g_{\Delta_1} : U \rightarrow B(\mathcal{Y}, \mathcal{X})$. By Proposition 9.27, we have, $\forall \Delta_i \in B(\mathcal{X}, \mathcal{Y})$, $i = 2, \dots, k + 1$,

$$g_{\Delta_1}^{(k)}(T)(\Delta_{k+1}) \cdots (\Delta_2) = f^{(k+1)}(T)(\Delta_{k+1}) \cdots (\Delta_2)(\Delta_1)$$

This implies that

$$\|f^{(k+1)}(T)\| = \sup_{\Delta_i \in B(\mathcal{X}, \mathcal{Y}), \|\Delta_i\| \leq 1, i=1, \dots, k+1} \|f^{(k+1)}(T)(\Delta_{k+1}) \cdots (\Delta_1)\|$$

$$= \sup_{\Delta_i \in B(\mathcal{X}, \mathcal{Y}), \|\Delta_i\| \leq 1, i=1, \dots, k+1} \left\| g_{\Delta_1}^{(k)}(T)(\Delta_{k+1}) \cdots (\Delta_2) \right\|$$

By Proposition 9.81, we have the following equality,

$$\begin{aligned} g_{\Delta_1}^{(k)}(T)(\Delta_{k+1}) \cdots (\Delta_2) &= \\ &- \sum_{i=0}^k \sum_{\substack{\Lambda_1 := \{j_1, \dots, j_i\} \text{ is any } i\text{-element subset of} \\ \Lambda := \{2, \dots, k+1\}; \Lambda_2 := \{j_{i+1}, \dots, j_k\} := \Lambda \setminus \Lambda_1}} f^{(i)}(T)(\Delta_{j_1}) \cdots (\Delta_{j_i}) \\ &(\Delta_1 f^{(k-i)}(T)(\Delta_{j_{i+1}}) \cdots (\Delta_{j_k})) \end{aligned}$$

Then,

$$\begin{aligned} \|f^{(k+1)}(T)\| &\leq \sum_{i=0}^k \sum_{\substack{\Lambda_1 := \{j_1, \dots, j_i\} \text{ is any } i\text{-element subset of} \\ \Lambda := \{2, \dots, k+1\}; \Lambda_2 := \{j_{i+1}, \dots, j_k\} := \Lambda \setminus \Lambda_1}} \|f^{(i)}(T)\| \|f^{(k-i)}(T)\| \\ &= \sum_{i=0}^k \binom{k}{i} \|f^{(i)}(T)\| \|f^{(k-i)}(T)\| \\ &\leq \sum_{i=0}^k \binom{k}{i} i! \alpha^{i+1} (k-i)! \alpha^{k+1-i} = \sum_{i=0}^k k! \alpha^{k+2} = (k+1)! \alpha^{k+2} \end{aligned}$$

where the last inequality follows from the inductive assumption. This completes the induction process.

Hence, f is analytic at T . By the arbitrariness of T , f is an analytic function. This completes the proof of the proposition. $\square$

Finally, we establish the analytic versions of the Inverse Function Theorem and the Implicit Function Theorem.

Theorem 9.85 (Inverse Function Theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $F : D \rightarrow \mathcal{Y}$ be analytic at $x_0 \in D^\circ$, and $y_0 = F(x_0) \in \mathcal{Y}$. Assume that $F^{(1)}(x_0) \in B(\mathcal{X}, \mathcal{Y})$ is bijective. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\exists U \subseteq D$ with $x_0 \in U \in \mathcal{O}_{\mathcal{X}}$ such that*

(i) $F|_U : U \rightarrow \mathcal{B}_{\mathcal{Y}}(y_0, \delta) =: V \subseteq \mathcal{Y}$ is bijective;

(ii) the inverse mapping $F_i : V \rightarrow U$ is $\mathcal{C}_\infty$ and $F_i^{(1)} : V \rightarrow B(\mathcal{Y}, \mathcal{X})$ is given by $F_i^{(1)}(y) = (F^{(1)}(F_i(y)))^{-1}$, $\forall y \in V$;

(iii) the inverse mapping F_i is analytic at $y_0 \in V$.

Proof The result (i) and (ii) are implied by Inverse Function Theorem 9.57. All we need to show is (iii). Since F is analytic at x_0 , then $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$, $\exists c \in [0, \infty) \subset \mathbb{R}$, and $\exists M \in (0, \infty) \subset \mathbb{R}$ such that $\bar{U} := \mathcal{B}_{\mathcal{X}}(x_0, \delta_1) \subseteq U$, $\|F^{(n)}(x)\| \leq cn!M^n$, $\forall n \in \mathbb{Z}_+$, $\forall x \in \bar{U}$. Then, $\exists \delta_2 \in (0, \delta] \subset \mathbb{R}$ such that $F_i(\mathcal{B}_{\mathcal{Y}}(y_0, \delta_2)) \subseteq \bar{U}$.

Define $h : \bar{U} \rightarrow B(\mathcal{Y}, \mathcal{X})$ by $h(x) = (F^{(1)}(x))^{-1}$, $\forall x \in \bar{U}$. Since $F|_{\bar{U}}$ is an analytic function, then $F^{(1)}|_{\bar{U}}$ is an analytic function. By Proposition 9.84 and Theorem 9.82, h is an analytic function. Then, $\exists \bar{\delta}_1 \in (0, \delta_1] \subset \mathbb{R}$, $\exists c_h \in [0, \infty) \subset \mathbb{R}$, and $\exists M_h \in (0, \infty) \subset \mathbb{R}$ such that $\|h^{(n)}(x)\| \leq c_h n! M_h^n$, $\forall n \in \mathbb{Z}_+$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \bar{\delta}_1)$. By choosing δ_1 sufficiently small, without loss of generality, we may take $\bar{\delta}_1 = \delta_1$.

Note that $F_i^{(1)} \circ F(x) = h(x)$, $\forall x \in \bar{U}$. Then, $h^{(1)}(x) = D(F_i^{(1)} \circ F)(x) = (F_i^{(2)} \circ F(x))F^{(1)}(x)$ and $F_i^{(2)} \circ F(x) = (h^{(1)}(x))h(x)$, $\forall x \in \bar{U}$. Define $h_k : \bar{U} \rightarrow B_{\mathcal{S}k}(\mathcal{Y}, \mathcal{X})$ by $h_k(x) = F_i^{(k)} \circ F(x)$, $\forall x \in \bar{U}$, $\forall 2 \leq k \in \mathbb{N}$. This leads to the recursive relationship

$$(Dh_k(x))h(x) = F_i^{(k+1)} \circ F(x) = h_{k+1}(x), \quad \forall x \in \bar{U}, \quad \forall 2 \leq k \in \mathbb{N}$$

with the initial condition $h_2(x) = h^{(1)}(x)h(x)$, $\forall x \in \bar{U}$. Define the nonnegative constants $K_{k,j} := \sup_{x \in \bar{U}} \|D^j h_k(x)\|$, $\forall j \in \mathbb{Z}_+$, $\forall 2 \leq k \in \mathbb{N}$. We will use mathematical induction on $2 \leq k \in \mathbb{N}$ to show that

$$0 \leq K_{k,j} \leq \frac{(j+2k-2)!}{(k-1)!2^{k-1}} c_h^k M_h^{j+k-1}, \quad \forall j \in \mathbb{Z}_+, \quad \forall 2 \leq k \in \mathbb{N}$$

1° $k = 2$. $\forall j \in \mathbb{Z}_+$, by the initial condition of the recursion,

$$\begin{aligned} K_{2,j} &= \sup_{x \in \bar{U}} \|D^j h_2(x)\| = \sup_{x \in \bar{U}} \|D^j (h^{(1)}h)(x)\| \\ &= \sup_{x \in \bar{U}} \sup_{\substack{\xi_{x_l} \in \mathcal{X}, \|\xi_{x_l}\| \leq 1, l=1, \dots, j \\ \xi_y \in \mathcal{Y}, \|\xi_y\| \leq 1}} \|D^j (h^{(1)}h)(x)(\xi_{x_j}) \cdots (\xi_{x_1})(\xi_y)\| \\ &= \sup_{x \in \bar{U}} \sup_{\substack{\xi_{x_l} \in \mathcal{X}, \|\xi_{x_l}\| \leq 1, l=1, \dots, j \\ \xi_y \in \mathcal{Y}, \|\xi_y\| \leq 1}} \left\| \sum_{l=0}^j \right. \\ &\quad \sum_{\substack{\Lambda_1 := \{s_1, \dots, s_l\} \text{ is any } l\text{-element subset of} \\ \Lambda := \{1, \dots, j\}; \Lambda_2 := \{s_{l+1}, \dots, s_j\} := \Lambda \setminus \Lambda_1}} h^{(l+1)}(x)(\xi_{x_{s_1}}) \cdots (\xi_{x_{s_l}}) \\ &\quad \left. (h^{(j-l)}(x)(\xi_{x_{s_{l+1}}}) \cdots (\xi_{x_{s_j}})(\xi_y)) \right\| \\ &\leq \sup_{x \in \bar{U}} \sum_{l=0}^j \sum_{\substack{\Lambda_1 := \{s_1, \dots, s_l\} \text{ is any } l\text{-element subset of} \\ \Lambda := \{1, \dots, j\}; \Lambda_2 := \{s_{l+1}, \dots, s_j\} := \Lambda \setminus \Lambda_1}} \|h^{(l+1)}(x)\| \|h^{(j-l)}(x)\| \\ &= \sup_{x \in \bar{U}} \sum_{l=0}^j \binom{j}{l} \|h^{(l+1)}(x)\| \|h^{(j-l)}(x)\| \\ &\leq \sum_{l=0}^j \frac{j!}{l!(j-l)!} c_h^{l+1} M_h^{l+1} c_h^{j-l} M_h^{j-l} \end{aligned}$$

$$= j!c_h^2 M_h^{j+1} \sum_{l=0}^j (l+1) = \frac{(j+2)!}{2} c_h^2 M_h^{j+1} = \text{RHS}$$

where the fourth equality follows from Proposition 9.81. This case is proved.

2° Assume that the result holds for $k = n \geq 2$ with $n \in \mathbb{N}$.

3° Consider the case when $k = n + 1$. $\forall j \in \mathbb{Z}_+$, by the recursive relationship,

$$\begin{aligned} K_{n+1,j} &= \sup_{x \in \bar{U}} \|D^j h_{n+1}(x)\| = \sup_{x \in \bar{U}} \|D^j(h_n^{(1)}h)(x)\| \\ &= \sup_{x \in \bar{U}} \sup_{\substack{\xi_{x_l} \in \mathcal{X}, \|\xi_{x_l}\| \leq 1, l=1, \dots, j \\ \xi_y \in \mathcal{Y}, \|\xi_y\| \leq 1}} \|D^j(h_n^{(1)}h)(x)(\xi_{x_j}) \cdots (\xi_{x_1})(\xi_y)\| \\ &= \sup_{x \in \bar{U}} \sup_{\substack{\xi_{x_l} \in \mathcal{X}, \|\xi_{x_l}\| \leq 1, l=1, \dots, j \\ \xi_y \in \mathcal{Y}, \|\xi_y\| \leq 1}} \left\| \sum_{l=0}^j \right. \\ &\quad \sum_{\substack{\Lambda_1 := \{s_1, \dots, s_l\} \text{ is any } l\text{-element subset of} \\ \Lambda := \{1, \dots, j\}; \Lambda_2 := \{s_{l+1}, \dots, s_j\} := \Lambda \setminus \Lambda_1}} h_n^{(l+1)}(x)(\xi_{x_{s_1}}) \cdots (\xi_{x_{s_l}}) \\ &\quad \left. (h^{(j-l)}(x)(\xi_{x_{s_{l+1}}}) \cdots (\xi_{x_{s_j}})(\xi_y)) \right\| \\ &\leq \sup_{x \in \bar{U}} \sum_{l=0}^j \sum_{\substack{\Lambda_1 := \{s_1, \dots, s_l\} \text{ is any } l\text{-element subset of} \\ \Lambda := \{1, \dots, j\}; \Lambda_2 := \{s_{l+1}, \dots, s_j\} := \Lambda \setminus \Lambda_1}} \|h_n^{(l+1)}(x)\| \|h^{(j-l)}(x)\| \\ &= \sup_{x \in \bar{U}} \sum_{l=0}^j \binom{j}{l} \|h_n^{(l+1)}(x)\| \|h^{(j-l)}(x)\| \\ &\leq \sup_{x \in \bar{U}} \sum_{l=0}^j \binom{j}{l} K_{n,l+1} \|h^{(j-l)}(x)\| \\ &\leq \sum_{l=0}^j \frac{j!}{l!(j-l)!} \frac{(l+2n-1)!}{(n-1)!2^{n-1}} c_h^n M_h^{l+n} c_h(j-l)! M_h^{j-l} \\ &= c_h^{n+1} M_h^{j+n} \frac{j!}{(n-1)!2^{n-1}} \sum_{l=0}^j \prod_{s=1}^{2n-1} (l+s) \\ &= c_h^{n+1} M_h^{j+n} \frac{j!}{(n-1)!2^{n-1}} \frac{1}{2n} \frac{(j+2n)!}{j!} = \frac{(j+2n)!}{n!2^n} c_h^{n+1} M_h^{j+n} = \text{RHS} \end{aligned}$$

where the fourth equality follows from Proposition 9.81; the second inequality follows the inductive assumption. This completes the induction process.

Based on this inequality, we have the following bounds for the high order derivatives of F_i around y_0 . $\forall y \in \mathcal{B}_Y(y_0, \delta_2)$, $\forall 2 \leq n \in \mathbb{N}$,

$$\begin{aligned} \|F_i^{(n)}(y)\| &\leq \sup_{x \in \bar{U}} \|F_i^{(n)} \circ F(x)\| = \sup_{x \in \bar{U}} \|h_n(x)\| \leq K_{n,0} \\ &\leq \frac{(2n-2)!}{(n-1)!2^{n-1}} c_h^n M_h^{n-1} \leq (n-1)!2^{n-1} c_h^n M_h^{n-1} = \frac{1}{2M_h n} n! (2c_h M_h)^n \end{aligned}$$

Hence, F_i is analytic at y_0 . This completes the proof of the theorem. $\square$

Theorem 9.86 (Implicit Function Theorem) *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be Banach spaces over $\mathbb{K}$, $D \subseteq \mathcal{X} \times \mathcal{Y}$, $F : D \rightarrow \mathcal{Z}$, and $(x_0, y_0) \in D^\circ$. Assume that F is analytic at (x_0, y_0) , $F(x_0, y_0) = \vartheta_Z$, and $\frac{\partial F}{\partial y}(x_0, y_0) \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ is bijective. Then, the following statements hold.*

(i) *There exists $r_0, r_1 \in (0, \infty) \subset \mathbb{R}$ such that $U \times V := \mathcal{B}_X(x_0, r_0) \times \mathcal{B}_Y(y_0, r_1) \subseteq D$ and $\forall x \in U$, $\exists! y \in V$ satisfying $F(x, y) = \vartheta_Z$. This defines a function $\phi : U \rightarrow V$ by $\phi(x) = y$, $\forall x \in U$. Then, ϕ is $\mathcal{C}_\infty$.*

(ii) $\phi^{(1)}(x) = -(\frac{\partial F}{\partial y}(x, \phi(x)))^{-1} \frac{\partial F}{\partial x}(x, \phi(x))$, $\forall x \in U$.

(iii) ϕ is analytic at x_0 .

Proof The results (i) and (ii) follows immediately from Implicit Function Theorem 9.59. All we need to show is (iii). Since F is analytic at (x_0, y_0) , we may choose r_0 and r_1 to be sufficiently small such that $\exists c \in [0, \infty) \subset \mathbb{R}$ and $\exists M \in (0, \infty) \subset \mathbb{R}$, $\|F^{(n)}(x, y)\| \leq cn!M^n$, $\forall n \in \mathbb{Z}_+$, $\forall (x, y) \in U \times V$.

Define mapping $\bar{F} : D \rightarrow \mathcal{X} \times \mathcal{Z}$ by $\bar{F}(x, y) = (x, F(x, y))$, $\forall (x, y) \in U \times V$. By Proposition 9.83 and Theorem 9.82, $\bar{F}$ is analytic at $(x_0, y_0) \in D^\circ$. By Propositions 9.44, 9.13, and 9.24,

$$\bar{F}^{(1)}(x_0, y_0) = \begin{bmatrix} \text{id}_X & \vartheta_{\mathcal{B}(\mathcal{Y}, \mathcal{X})} \\ \frac{\partial F}{\partial x}(x_0, y_0) & \frac{\partial F}{\partial y}(x_0, y_0) \end{bmatrix}$$

Clearly, $\bar{F}^{(1)}(x_0, y_0)$ is bijective. By Proposition 4.31, $\mathcal{X} \times \mathcal{Y}$ and $\mathcal{X} \times \mathcal{Z}$ are Banach spaces over $\mathbb{K}$. By the analytic version of the Inverse Function Theorem 9.85, $\exists \bar{U} \subseteq D$ with $(x_0, y_0) \in \bar{U} \in \mathcal{O}_{\mathcal{X} \times \mathcal{Y}}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\bar{F}|_{\bar{U}} : \bar{U} \rightarrow \mathcal{B}_{\mathcal{X} \times \mathcal{Z}}((x_0, \vartheta_Z), \delta)$ is bijective, whose inverse function $\bar{F}_i : \mathcal{B}_{\mathcal{X} \times \mathcal{Z}}((x_0, \vartheta_Z), \delta) \rightarrow \bar{U}$ is analytic at (x_0, ϑ_Z) . By taking r_0 and r_1 sufficiently small, we may assume that $U \times V \subseteq \bar{U}$. Then, $\exists \delta_1 \in (0, \min\{\delta, r_0\}] \subset \mathbb{R}$, $\exists W \subseteq U \times V$ such that $\bar{F}|_W \rightarrow \mathcal{B}_{\mathcal{X} \times \mathcal{Z}}((x_0, \vartheta_Z), \delta_1)$ is bijective. By the uniqueness of $y \in V$ that solves $F(x, y) = \vartheta_Z$, we have $\phi(x) = \pi_Y \circ \bar{F}_i(x, \vartheta_Z)$, $\forall x \in \mathcal{B}_X(x_0, \delta_1)$. By Theorem 9.82 and Proposition 9.83, ϕ is analytic at x_0 . This completes the proof of the theorem. $\square$

Chapter 10

Local Theory of Optimization

In this chapter, we will develop a number of tools for optimization of sufficiently many times differentiable functions. As in Chapter 8, we will be mainly concerned with real spaces, rather than complex ones.

10.1 Basic Notion

Definition 10.1 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $f : \mathcal{X} \rightarrow \mathbb{R}$, and $x_0 \in \mathcal{X}$. x_0 is said to be a point of minimum for f if $f(x_0) \leq f(x)$, $\forall x \in \mathcal{X}$. It is said to be the point of strict minimum for f if $f(x_0) < f(x)$, $\forall x \in \mathcal{X} \setminus \{x_0\}$. It is said to be a point of relative minimum for f if $\exists O \in \mathcal{O}$ with $x_0 \in O$ such that $f(x_0) \leq f(x)$, $\forall x \in O$. It is said to be a point of relative strict minimum for f if $\exists O \in \mathcal{O}$ with $x_0 \in O$ such that $f(x_0) < f(x)$, $\forall x \in O \setminus \{x_0\}$. Similar definitions for points of maxima. Moreover, x_0 is said to be a point of relative extremum if it is a point of relative minimum or relative maximum. It is said to be a point of relative strict extremum if it is a point of relative strict minimum or relative strict maximum.

Proposition 10.2 Let $\mathcal{X}$ be a real normed linear space, $D \subseteq \mathcal{X}$, $x_0 \in D$, $f : D \rightarrow \mathbb{R}$, $u \in A_D(x_0)$. Assume that the directional derivative of f at x_0 along u exists and x_0 is a point of relative minimum for f . Then, $Df(x_0; u) \geq 0$.

Proof This is immediate from Definition 9.4. □

Definition 10.3 Let $\mathcal{X}$ be a real normed linear space and $A \in B_{S_2}(\mathcal{X}, \mathbb{R})$. A is said to be positive definite if $\exists m \in (0, \infty) \subset \mathbb{R}$ such that $A(x)(x) = \langle Ax, x \rangle \geq m \|x\|^2$, $\forall x \in \mathcal{X}$. Let the set of all such positive definite

operators be denoted by $\mathcal{S}_+ \mathcal{X}$. A is said to be positive semi-definite if $A(x)(x) \geq 0, \forall x \in \mathcal{X}$. Let the set of all such positive semi-definite operators be denoted by $\mathcal{S}_{\text{psd}} \mathcal{X}$. A is said to be negative definite if $\exists m \in (0, \infty) \subset \mathbb{R}$ such that $A(x)(x) \leq -m \|x\|^2, \forall x \in \mathcal{X}$. Let the set of all such negative definite operators be denoted by $\mathcal{S}_- \mathcal{X}$. A is said to be negative semi-definite if $A(x)(x) \leq 0, \forall x \in \mathcal{X}$. Let the set of all such negative semi-definite operators be denoted by $\mathcal{S}_{\text{nsd}} \mathcal{X}$. We will denote $\text{B}_{\mathbb{S}_2}(\mathcal{X}, \mathbb{R})$ by $\mathcal{S}_{\mathcal{X}}$.

Proposition 10.4 *Let $\mathcal{X}$ be a real normed linear space. Then,*

- (i) $\mathcal{S}_- \mathcal{X} = -\mathcal{S}_+ \mathcal{X}$ and $\mathcal{S}_{\text{nsd}} \mathcal{X} = -\mathcal{S}_{\text{psd}} \mathcal{X}$;
- (ii) $\mathcal{S}_+ \mathcal{X}$ and $\mathcal{S}_- \mathcal{X}$ are open sets in $\text{B}_{\mathbb{S}_2}(\mathcal{X}, \mathbb{R}) = \mathcal{S}_{\mathcal{X}}$;
- (iii) $\mathcal{S}_{\text{psd}} \mathcal{X}$ and $\mathcal{S}_{\text{nsd}} \mathcal{X}$ are closed convex cones in $\mathcal{S}_{\mathcal{X}}$;
- (iv) $\mathcal{S}_+ \mathcal{X} \subseteq \mathcal{S}_{\text{psd}}^\circ \mathcal{X}$ and $\mathcal{S}_- \mathcal{X} \subseteq \mathcal{S}_{\text{nsd}}^\circ \mathcal{X}$.

Proof (i) This is clear. (ii) Fix $A \in \mathcal{S}_+ \mathcal{X}$. Then, $\exists m \in (0, \infty) \subset \mathbb{R}$ such that $A(x)(x) \geq m \|x\|^2, \forall x \in \mathcal{X}$. $\forall B \in \text{B}_{\mathbb{S}_2}(\mathcal{X}, \mathbb{R})$ with $\|B - A\| < m/2, \forall x \in \mathcal{X}$, we have $B(x)(x) = A(x)(x) + (B - A)(x)(x) \geq m \|x\|^2 - \|B - A\| \|x\|^2 \geq m \|x\|^2 / 2$, where the first inequality follows from Proposition 7.64. This implies that $B \in \mathcal{S}_+ \mathcal{X}$. Then, $A \in \mathcal{S}_+^\circ \mathcal{X}$. Hence, $\mathcal{S}_+ \mathcal{X}$ is open in $\mathcal{S}_{\mathcal{X}}$. Therefore, $\mathcal{S}_- \mathcal{X} = -\mathcal{S}_+ \mathcal{X}$ is open in $\mathcal{S}_{\mathcal{X}}$.

(iii) Clearly, $\vartheta_{\text{B}_{\mathbb{S}_2}(\mathcal{X}, \mathbb{R})} \in \mathcal{S}_{\text{psd}} \mathcal{X}$. $\forall A \in \mathcal{S}_{\text{psd}} \mathcal{X}, \forall \alpha \in [0, \infty) \subset \mathbb{R}, \alpha A \in \mathcal{S}_{\text{psd}} \mathcal{X}$. Hence, $\mathcal{S}_{\text{psd}} \mathcal{X}$ is a cone. $\forall A, B \in \mathcal{S}_{\text{psd}} \mathcal{X}, \forall x \in \mathcal{X}, (A + B)(x)(x) = A(x)(x) + B(x)(x) \geq 0$. Hence, $A + B \in \mathcal{S}_{\text{psd}} \mathcal{X}$. Then, $\mathcal{S}_{\text{psd}} \mathcal{X}$ is a convex cone.

Let $M := \mathcal{S}_{\mathcal{X}} \setminus \mathcal{S}_{\text{psd}} \mathcal{X}$. $\forall A \in M, \exists x_0 \in \mathcal{X}$ such that $A(x_0)(x_0) < 0$. Then, there exists $m \in (0, \infty) \subset \mathbb{R}$ such that $A(x_0)(x_0) < -m \|x_0\|^2 < 0$. $\forall B \in \text{B}_{\mathbb{S}_2}(\mathcal{X}, \mathbb{R})$ with $\|B - A\| < m/2$, we have $B(x_0)(x_0) = A(x_0)(x_0) + (B - A)(x_0)(x_0) < -m \|x_0\|^2 + \|B - A\| \|x_0\|^2 \leq -m \|x_0\|^2 / 2 < 0$, where the first inequality follows from Proposition 7.64. This implies that $B \in M$ and $A \in M^\circ$. Then, $M = M^\circ$ is open. Therefore, $\mathcal{S}_{\text{psd}} \mathcal{X}$ is closed.

$\mathcal{S}_{\text{nsd}} \mathcal{X} = -\mathcal{S}_{\text{psd}} \mathcal{X}$ is clearly also a closed convex cone.

(iv) Clearly, $\mathcal{S}_+ \mathcal{X} \subseteq \mathcal{S}_{\text{psd}} \mathcal{X}$. Then, $\mathcal{S}_+ \mathcal{X} \subseteq \mathcal{S}_{\text{psd}}^\circ \mathcal{X}$. This further implies that $\mathcal{S}_- \mathcal{X} \subseteq \mathcal{S}_{\text{nsd}}^\circ \mathcal{X}$. This completes the proof of the proposition. $\square$

Definition 10.5 *Let $(\mathcal{X}, \mathbb{K})$ be a vector space and $K \subseteq \mathcal{X}$ be convex. $M \subseteq K$ is said to be an extreme subset of K if $M \neq \emptyset$ and $\forall x_1, x_2 \in K, \forall \alpha \in (0, 1) \subset \mathbb{R}, \alpha x_1 + (1 - \alpha)x_2 \in M$ implies that $x_1, x_2 \in M$. If a singleton set $\{x_0\} \subseteq K$ is an extreme subset, then x_0 is called an extreme point of K .*

Proposition 10.6 *Let $\mathcal{X}$ be a real normed linear space, $K \subseteq \mathcal{X}$ be a nonempty convex set, and $H := \{x \in \mathcal{X} \mid \langle x_{*0}, x \rangle = c\}$ be a supporting hyperplane of K , where $x_{*0} \in \mathcal{X}^*$ with $x_{*0} \neq \vartheta_{\mathcal{X}^*}$ and $c \in \mathbb{R}$. Then, any extreme subset of $K_1 := K \cap H$ is also an extreme subset of K .*

Proof Without loss of generality, assume that $\inf_{k \in K} \langle \langle x_{*0}, x \rangle \rangle = c$. Let $M \subseteq K_1$ be an extreme subset of K_1 . Then, $M \neq \emptyset$. $\forall x_1, x_2 \in \mathbb{K}$, $\forall \alpha \in (0, 1) \subset \mathbb{R}$, let $\bar{x} := \alpha x_1 + (1 - \alpha)x_2 \in M \subseteq H$. Then, $\langle \langle x_{*0}, x_1 \rangle \rangle \geq c$ and $\langle \langle x_{*0}, x_2 \rangle \rangle \geq c$ and $\langle \langle x_{*0}, \bar{x} \rangle \rangle = c$. This implies that $\alpha(\langle \langle x_{*0}, x_1 \rangle \rangle - c) + (1 - \alpha)(\langle \langle x_{*0}, x_2 \rangle \rangle - c) = 0$ and $\langle \langle x_{*0}, x_1 \rangle \rangle = c = \langle \langle x_{*0}, x_2 \rangle \rangle$. Hence, $x_1, x_2 \in K_1$. Since, M is an extreme subset of K_1 , then $x_1, x_2 \in M$. Therefore, M is an extreme subset of K . This completes the proof of the proposition. $\square$

Proposition 10.7 (Krein-Milman) *Let $\mathcal{X}$ be a real reflexive Banach space, $K \subseteq \mathcal{X}$ be a nonempty bounded closed convex set, and $M \subseteq K$ be a weakly compact extreme subset of K . Then, M contains at least one extreme point of K .*

Proof By Proposition 8.11, K is compact in $\mathcal{X}_{\text{weak}}$. Let $\mathcal{M} = \{E \subseteq M \mid E \text{ is a weakly compact extreme subset of } K\}$. Clearly, $M \in \mathcal{M} \neq \emptyset$. Clearly, $\supseteq$ defines an antisymmetric partial ordering on $\mathcal{M}$, where smaller sets are further down the stream. Next, we will use Zorn's Lemma to show that $\mathcal{M}$ admits a maximal element.

Let $\mathcal{E} \subseteq \mathcal{M}$ be a nonempty totally ordered (by $\supseteq$) subcollection. Let $E_0 := \bigcap_{E \in \mathcal{E}} E$. $\forall E \in \mathcal{E}$, $E \subseteq M$ is weakly compact extreme subset of K . Then, by Propositions 7.116, 5.5, and 3.61, E is weakly closed. By Proposition 5.5, E_0 is weakly compact. $\forall x_1, x_2 \in K$, $\forall \alpha \in (0, 1) \subset \mathbb{R}$, let $\alpha x_1 + (1 - \alpha)x_2 \in E_0$. $\forall E \in \mathcal{E}$, $\alpha x_1 + (1 - \alpha)x_2 \in E$. Since E is an extreme subset of K , then $x_1, x_2 \in E$. By the arbitrariness of E , we have $x_1, x_2 \in E_0$. Hence, E_0 is an extreme subset of K if $E_0 \neq \emptyset$. $\forall E \in \mathcal{E}$, E is nonempty. Since $\mathcal{E}$ is totally ordered by $\supseteq$, then the intersection of finite number of sets in $\mathcal{E}$ is again in $\mathcal{E}$, and hence nonempty. By Proposition 5.12, $E_0 \neq \emptyset$. Then, E_0 is a weakly compact extreme subset of K and $E_0 \in \mathcal{M}$. Clearly, E_0 is an upper bound of $\mathcal{E}$ (in terms of $\supseteq$). Then, by Zorn's Lemma, $\mathcal{M}$ admits a maximal element E_M . Then, $E_M \subseteq M$ is a weakly compact extreme subset of K and $E_M \neq \emptyset$.

We will show that E_M is a singleton set, which then proves that M contains an extreme point of K . Suppose that $\exists x_1, x_2 \in E_M$ with $x_1 \neq x_2$. Let $N := \text{span}(\{x_2 - x_1\})$ and define a functional $f : N \rightarrow \mathbb{R}$ by $f(\alpha(x_2 - x_1)) = \alpha$, $\forall \alpha \in \mathbb{R}$. Clearly, f is a linear functional on N , and $\|f\|_N = 1/\|x_2 - x_1\| < \infty$. By Hahn-Banach Theorem 7.83, there exists $x_{*0} \in \mathcal{X}^*$ with $\|x_{*0}\| = 1/\|x_2 - x_1\|$ such that $\langle \langle x_{*0}, \alpha(x_2 - x_1) \rangle \rangle = \alpha$, $\forall \alpha \in \mathbb{R}$. Clearly, $\langle \langle x_{*0}, x_1 \rangle \rangle \neq \langle \langle x_{*0}, x_2 \rangle \rangle$. Note that E_M is nonempty and compact in $\mathcal{X}_{\text{weak}}$ and x_{*0} is weakly continuous. By Proposition 5.29, $c := \langle \langle x_{*0}, x_0 \rangle \rangle = \inf_{x \in E_M} \langle \langle x_{*0}, x \rangle \rangle \in \mathbb{R}$ for some $x_0 \in E_M$. Define $H := \{x \in \mathcal{X} \mid \langle \langle x_{*0}, x \rangle \rangle = c\}$. Let $E_m := E_M \cap H$. Then, at least one of x_1 and x_2 is not in E_m . Hence, $E_M \supset E_m$. Clearly, $E_m \ni x_0$ is nonempty. Note that H is weakly closed. Then, by Proposition 5.5, $E_m \subseteq M$ is weakly compact. $\forall \bar{x}_1, \bar{x}_2 \in K$, $\forall \bar{\alpha} \in (0, 1) \subset \mathbb{R}$, let $\bar{\alpha}\bar{x}_1 + (1 - \bar{\alpha})\bar{x}_2 \in E_m \subseteq E_M$. Since E_M is an extreme subset of K , then we have $\bar{x}_1, \bar{x}_2 \in E_M$.

This further implies that $\langle\langle x_{*0}, \bar{x}_1 \rangle\rangle \geq c$ and $\langle\langle x_{*0}, \bar{x}_2 \rangle\rangle \geq c$. Note that $\langle\langle x_{*0}, \bar{\alpha}\bar{x}_1 + (1 - \bar{\alpha})\bar{x}_2 \rangle\rangle = c$. Then, we must have $\langle\langle x_{*0}, \bar{x}_1 \rangle\rangle = c = \langle\langle x_{*0}, \bar{x}_2 \rangle\rangle$ and $\bar{x}_1, \bar{x}_2 \in E_m$. This shows that E_m is an extreme subset of K . Then $E_m \in \mathcal{M}$. This contradicts with the fact that E_M is maximal with respect to $\supseteq$. Therefore, E_M is a singleton set.

This completes the proof of the proposition. $\square$

Proposition 10.8 *Let $\mathcal{X}$ be a real reflexive Banach space, $K \subseteq \mathcal{X}$ be a nonempty bounded closed convex set, and E be the set of extreme points of K . Then, $K = \overline{\text{co}(E)}$.*

Proof Let $C := \overline{\text{co}(E)}$. By Proposition 7.15, C is closed and convex and $C \subseteq K$. We will prove the result by an argument of contradiction. Suppose $K \supset C$. Then, $\exists x_0 \in K \setminus C$. By Proposition 8.10, there exists $x_{*0} \in \mathcal{X}^*$ such that $\langle\langle x_{*0}, x_0 \rangle\rangle < \inf_{x \in C} \langle\langle x_{*0}, x \rangle\rangle$. By Proposition 8.11, K is compact in the weak topology and x_{*0} is continuous in the weak topology. Then, by Proposition 5.29, $c_0 := \langle\langle x_{*0}, x_1 \rangle\rangle = \inf_{x \in K} \langle\langle x_{*0}, x \rangle\rangle \in \mathbb{R}$ for some $x_1 \in K$. Clearly $c_0 \leq \langle\langle x_{*0}, x_0 \rangle\rangle < \inf_{x \in C} \langle\langle x_{*0}, x \rangle\rangle$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x_{*0} = \vartheta_{\mathcal{X}^*}$; Case 2: $x_{*0} \neq \vartheta_{\mathcal{X}^*}$. Case 1: $x_{*0} = \vartheta_{\mathcal{X}^*}$. Then, $c_0 = 0$ and $C = \emptyset$. By Proposition 10.7, $E \neq \emptyset$. Then, $C \neq \emptyset$. This is a contradiction. Case 2: $x_{*0} \neq \vartheta_{\mathcal{X}^*}$. Let $H := \{x \in \mathcal{X} \mid \langle\langle x_{*0}, x \rangle\rangle = c_0\}$. H is a supporting hyperplane of K and $H \cap C = \emptyset$. Let $C_m := K \cap H \neq \emptyset$. Clearly, C_m is bounded closed and convex. By Proposition 10.7, there is an extreme point $x_m \in C_m$ of C_m . Then, $\{x_m\} \subseteq C_m$ is an extreme subset of C_m . By Proposition 10.6, $\{x_m\}$ is an extreme subset of K . Then, x_m is an extreme point of K and $x_m \in E$. This leads to the contradiction $x_m \in C_m \subseteq H$, $x_m \in E \subseteq C$ and $C \cap H = \emptyset$. Thus, in both cases, we have arrived at a contradiction. Then, the hypothesis must be false. Hence, $K = \overline{\text{co}(E)}$. This completes the proof of the proposition. $\square$

Proposition 10.9 *Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be a convex set, and $f_1 : \Omega \rightarrow \mathbb{R}$ and $f_2 : \Omega \rightarrow \mathbb{R}$ be convex functionals. Then, the following statements hold.*

- (i) $\forall n \in \mathbb{N}, \forall x_1, \dots, x_n \in \Omega, \forall \alpha_1, \dots, \alpha_n \in [0, 1] \subset \mathbb{R}$ with $\sum_{i=1}^n \alpha_i = 1$, we have $f_1(\sum_{i=1}^n \alpha_i x_i) \leq \sum_{i=1}^n \alpha_i f_1(x_i)$. If, in addition, f_1 is strictly convex, $x_1, \dots, x_n$ are distinct, and $\alpha_1, \dots, \alpha_n \in (0, 1) \subset \mathbb{R}$, then $f_1(\sum_{i=1}^n \alpha_i x_i) < \sum_{i=1}^n \alpha_i f_1(x_i)$.
- (ii) $\forall \alpha_1, \alpha_2 \in [0, \infty) \subset \mathbb{R}$, $\alpha_1 f_1 + \alpha_2 f_2$ is convex. If, in addition, f_1 is strictly convex and $\alpha_1 \in (0, \infty) \subset \mathbb{R}$, then $\alpha_1 f_1 + \alpha_2 f_2$ is strictly convex.
- (iii) $\forall c \in \mathbb{R}, \{x \in \Omega \mid f_1(x) \leq c\}$ is convex.
- (iv) Let $\mathcal{Y}$ be a real vector space, $A : \mathcal{Y} \rightarrow \mathcal{X}$ be an affine operator, $D \subseteq \mathcal{Y}$ be convex, and $A(D) \subseteq \Omega$. Then, $f_1 \circ A|_D : D \rightarrow \mathbb{R}$ is convex.

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 10.10 Let $f : [a, b] \rightarrow \mathbb{R}$, where $a, b \in \mathbb{R}$ with $a < b$. Then, the following statements hold.

- (i) f is convex if, and only if, $\forall x_1, x_2, x_3 \in [a, b] \subset \mathbb{R}$ with $x_1 < x_2 < x_3$, we have

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}$$

- if, and only if, $\forall x_1, x_2, x_3 \in [a, b] \subset \mathbb{R}$ with $x_1 < x_2 < x_3$, we have

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1}$$

- (ii) f is strictly convex if, and only if, $\forall x_1, x_2, x_3 \in [a, b] \subset \mathbb{R}$ with $x_1 < x_2 < x_3$, we have

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} < \frac{f(x_3) - f(x_2)}{x_3 - x_2}$$

- if, and only if, $\forall x_1, x_2, x_3 \in [a, b] \subset \mathbb{R}$ with $x_1 < x_2 < x_3$, we have

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} < \frac{f(x_3) - f(x_1)}{x_3 - x_1}$$

- (iii) If f is convex and $c \in (a, b) \subset \mathbb{R}$, then, the one-sided derivatives

$$\lim_{x \rightarrow c^+} \frac{f(x) - f(c)}{x - c}; \quad \lim_{x \rightarrow c^-} \frac{f(x) - f(c)}{x - c}$$

exist.

Proof This is straightforward, and is therefore omitted. $\square$

Proposition 10.11 Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$ be convex, and $f : \Omega \rightarrow \mathbb{R}$ be differentiable. Then, the following statements hold.

- (i) f is convex if, and only if, $\forall x, y \in \Omega$, we have $f(y) \geq f(x) + f^{(1)}(x)(y - x)$.

- (ii) f is strictly convex if, and only if, $\forall x, y \in \Omega$ with $x \neq y$, we have $f(y) > f(x) + f^{(1)}(x)(y - x)$.

Proof (i) “Necessity” Let f be convex. $\forall x, y \in \Omega$, by the convexity of Ω , we have $y - x \in A_\Omega(x)$. $\forall \alpha \in (0, 1] \subset \mathbb{R}$, we have $f(\alpha y + (1 - \alpha)x) \leq \alpha f(y) + (1 - \alpha)f(x)$, which further implies that $f(y) - f(x) \geq (f(x + \alpha(y - x)) - f(x))/\alpha$. Then, we have

$$\begin{aligned} f(y) - f(x) &\geq \lim_{\alpha \rightarrow 0^+} (f(x + \alpha(y - x)) - f(x))/\alpha = Df(x; y - x) \\ &= f^{(1)}(x)(y - x) \end{aligned}$$

where the first equality follows from Definition 9.4 and the last equality follows from Proposition 9.5.

“Sufficiency” Let $f(y) \geq f(x) + f^{(1)}(x)(y - x)$, $\forall x, y \in \Omega$. $\forall x_1, x_2 \in \Omega$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, let $x := \alpha x_1 + (1 - \alpha)x_2 \in \Omega$. Then, $f(x_2) = f(x + \alpha(x_2 - x_1)) \geq f(x) + \alpha f^{(1)}(x)(x_2 - x_1)$. Note also, $f(x_1) = f(x + (1 - \alpha)(x_1 - x_2)) \geq f(x) + (1 - \alpha)f^{(1)}(x)(x_1 - x_2)$. Then, $\alpha f(x_1) + (1 - \alpha)f(x_2) \geq f(x)$. Hence, f is convex.

(ii) “Necessity” Let f be strictly convex. $\forall x, y \in \Omega$ with $x \neq y$, by the convexity of Ω , we have $y - x \in A_\Omega(x)$. $\forall \alpha \in (0, 1) \subset \mathbb{R}$, we have $f(\alpha y + (1 - \alpha)x) < \alpha f(y) + (1 - \alpha)f(x)$, which further implies that $f(y) - f(x) > (f(x + \alpha(y - x)) - f(x))/\alpha$. Define $A : \mathbb{R} \rightarrow \mathcal{X}$ by $A(\beta) = x + \beta(y - x)$, $\forall \beta \in \mathbb{R}$. Clearly, A is an affine operator and $A(I) \subseteq \Omega$, where $I := [0, 1] \subset \mathbb{R}$. By Proposition 10.9, $g := f \circ A|_I$ is convex. Then, we have

$$\begin{aligned} f(y) - f(x) &> 2(f(x + 0.5(y - x)) - f(x)) = \frac{g(0.5) - g(0)}{0.5} \\ &\geq \lim_{\alpha \rightarrow 0^+} \frac{g(\alpha) - g(0)}{\alpha} = \lim_{\alpha \rightarrow 0^+} \frac{f(x + \alpha(y - x)) - f(x)}{\alpha} \\ &= Df(x; y - x) = f^{(1)}(x)(y - x) \end{aligned}$$

where the second inequality follows from Proposition 10.10 and the third equality follows from Definition 9.4 and the last equality follows from Proposition 9.5.

“Sufficiency” Let $f(y) > f(x) + f^{(1)}(x)(y - x)$, $\forall x, y \in \Omega$ with $x \neq y$. $\forall x_1, x_2 \in \Omega$ with $x_1 \neq x_2$, $\forall \alpha \in (0, 1) \subset \mathbb{R}$, let $x := \alpha x_1 + (1 - \alpha)x_2 \in \Omega$. Then, $f(x_2) = f(x + \alpha(x_2 - x_1)) > f(x) + \alpha f^{(1)}(x)(x_2 - x_1)$. Note also, $f(x_1) = f(x + (1 - \alpha)(x_1 - x_2)) > f(x) + (1 - \alpha)f^{(1)}(x)(x_1 - x_2)$. Then, $\alpha f(x_1) + (1 - \alpha)f(x_2) > f(x)$. Hence, f is strictly convex.

This completes the proof of the proposition. $\square$

Proposition 10.12 *Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$ be convex, and $f : \Omega \rightarrow \mathbb{R}$ be twice differentiable. Then, the following statements hold.*

- (i) *If $f^{(2)}(x)$ is positive semi-definite, $\forall x \in \Omega$, then f is convex.*
- (ii) *If f is convex and $\mathcal{C}_2$, then $f^{(2)}(x)$ is positive semi-definite, $\forall x \in \bar{\Omega}^\circ \cap \Omega$.*
- (iii) *If f is convex and $f^{(2)}(x)$ is positive definite for all $x \in \Omega \setminus E$, where $E \subseteq \Omega$ does not contain any line segment, then f is strictly convex.*

Proof (i) $\forall x, y \in \Omega$, by Taylor’s Theorem, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ we have

$$f(y) = f(x) + f^{(1)}(x)(y - x) + \frac{1}{2}f^{(2)}(t_0 y + (1 - t_0)x)(y - x)(y - x)$$

By the assumption, $f^{(2)}(t_0 y + (1 - t_0)x)(y - x)(y - x) \geq 0$. Then, $f(y) \geq f(x) + f^{(1)}(x)(y - x)$. By the arbitrariness of x, y and Proposition 10.11, f is convex.

(ii) We will prove this statement by an argument of contradiction. Suppose $\exists x_0 \in \overline{\Omega^\circ} \cap \Omega$ such that $f^{(2)}(x_0)$ is not positive semi-definite. Then, $\exists h \in \mathcal{X}$ such that $f^{(2)}(x_0)(h)(h) < 0$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x_0 \in \Omega^\circ$; Case 2: $x_0 \notin \Omega^\circ$. Case 1: $x_0 \in \Omega^\circ$. By the continuity of $f^{(2)}$ and $x_0 \in \Omega^\circ$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $x_1 := x_0 + \delta h \in \Omega$ and $f^{(2)}(x_0 + \alpha \delta h)(h)(h) < 0$, $\forall \alpha \in (0, 1) \subset \mathbb{R}$. By Taylor's Theorem, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ we have

$$f(x_1) = f(x_0) + f^{(1)}(x_0)(x_1 - x_0) + \frac{\delta^2}{2} f^{(2)}(x_0 + t_0 \delta h)(h)(h)$$

Then, $f(x_1) < f(x_0) + f^{(1)}(x_0)(x_1 - x_0)$. This contradicts with the fact that f is convex and Proposition 10.11.

Case 2: $x_0 \notin \Omega^\circ$. Then, $x_0 \in (\overline{\Omega^\circ} \cap \Omega) \setminus \Omega^\circ$. Then, by Proposition 4.13, there exists $(x_n)_{n=1}^\infty \subseteq \Omega^\circ$ such that $\lim_{n \in \mathbb{N}} x_n = x_0$. By the continuity of $f^{(2)}$, we have $\exists n_0 \in \mathbb{N}$ such that $f^{(2)}(x_{n_0})(h)(h) < 0$. Hence, $f^{(2)}(x_{n_0})$ is not positive semi-definite and $x_{n_0} \in \Omega^\circ$. By Case 1, there is a contradiction. Hence, the statement must be true.

(iii) We will prove this statement by an argument of contradiction. Suppose f is not strictly convex. By Proposition 10.11, $\exists x, y \in \Omega$ with $x \neq y$ such that $f(y) = f(x) + f^{(1)}(x)(y - x)$. By the convexity of f and Proposition 10.11, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we have $f(x + \alpha(y - x)) \geq f(x) + \alpha f^{(1)}(x)(y - x) = f(x) + \alpha(f(y) - f(x))$. By the convexity of f , we have $f(x + \alpha(y - x)) = f(x) + \alpha(f(y) - f(x))$. Define $g : [0, 1] \rightarrow \mathbb{R}$ by $g(\alpha) = f(\alpha y + (1 - \alpha)x) - \alpha f(y) - (1 - \alpha)f(x)$, $\forall \alpha \in I := [0, 1] \subset \mathbb{R}$. Then, $g(\alpha) = 0$, $\forall \alpha \in I$. This implies that $g^{(2)}(\alpha) = 0$, $\forall \alpha \in I$, and $f^{(2)}(\alpha y + (1 - \alpha)x)(y - x)(y - x) = 0$, $\forall \alpha \in I$. Hence, $f^{(2)}(\alpha y + (1 - \alpha)x)$ is not positive definite, $\forall \alpha \in I$. Hence, E contains the line segment connecting x and y . This contradicts with the assumption. Therefore, the statement must be true.

This completes the proof of the proposition. $\square$

10.2 Unconstrained Optimization

The basic problem to be considered in this section is

$$\mu_0 := \inf_{x \in \Omega} f(x) \tag{10.1}$$

where $\mathcal{X}$ is a real normed linear space, $\Omega \subseteq \mathcal{X}$ is a set, and $f : \Omega \rightarrow \mathbb{R}$ is a functional.

Proposition 10.13 *Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$ be convex, $f : \Omega \rightarrow \mathbb{R}$ be a convex functional, and $\mu_0 := \inf_{x \in \Omega} f(x) \in \mathbb{R}$. Then, the following statements hold.*

- (i) *The set of all points of minimum for f , which is given by $\{x \in \Omega \mid f(x) = \mu_0\}$, is convex.*

(ii) Any point of relative minimum for f is a point of minimum for f .

(iii) Any point of relative strict minimum for f is the point of strict minimum for f .

Proof (i) Note that $\{x \in \Omega \mid f(x) = \mu_0\} = \{x \in \Omega \mid f(x) \leq \mu_0\}$, which is convex by Proposition 10.9.

(ii) Fix any $x_0 \in \Omega$ that is a point of relative minimum for f . Then, $\exists \epsilon \in (0, \infty) \subset \mathbb{R}$ such that $f(x) \geq f(x_0)$, $\forall x \in \Omega \cap \mathcal{B}_{\mathcal{X}}(x_0, \epsilon)$. $\forall y \in \Omega$, $\exists \alpha \in (0, 1) \subset \mathbb{R}$ such that $x_0 + \alpha(y - x_0) \in \mathcal{B}_{\mathcal{X}}(x_0, \epsilon)$. Then, $f(x_0) \leq f(x_0 + \alpha(y - x_0)) = f(\alpha y + (1 - \alpha)x_0) \leq \alpha f(y) + (1 - \alpha)f(x_0)$. This implies that $f(x_0) \leq f(y)$. Hence, x_0 is a point of minimum for f .

(iii) Fix any $x_0 \in \Omega$ that is a point of relative strict minimum for f . Then, $\exists \epsilon \in (0, \infty) \subset \mathbb{R}$ such that $f(x) > f(x_0)$, $\forall x \in (\Omega \cap \mathcal{B}_{\mathcal{X}}(x_0, \epsilon)) \setminus \{x_0\}$. $\forall y \in \Omega$ with $y \neq x_0$, $\exists \alpha \in (0, 1) \subset \mathbb{R}$ such that $x_0 + \alpha(y - x_0) \in \mathcal{B}_{\mathcal{X}}(x_0, \epsilon)$. Then, $f(x_0) < f(x_0 + \alpha(y - x_0)) = f(\alpha y + (1 - \alpha)x_0) \leq \alpha f(y) + (1 - \alpha)f(x_0)$. This implies that $f(x_0) < f(y)$. Hence, x_0 is a point of strict minimum for f . This completes the proof of the proposition. $\square$

Proposition 10.14 Let $\mathcal{X}$ be a real reflexive Banach space, $\Omega \subseteq \mathcal{X}$ be a nonempty bounded closed convex set, and $f : \Omega \rightarrow \mathbb{R}$ be a weakly upper semicontinuous convex functional. Then, there exist an extreme point $x_0 \in \Omega$ of Ω such that x_0 is a point of maximum for f .

Proof By Proposition 8.11, Ω is weakly compact. Let $\mathcal{O}_{\text{weak}}(\mathcal{X})$ be the weak topology on $\mathcal{X}$ and $\mathcal{O}_{\text{weak}, \Omega}$ be the subset topology on Ω with respect to $\mathcal{O}_{\text{weak}}(\mathcal{X})$. By Proposition 5.30, there exists $x_1 \in \Omega$ that is a point of maximum for f . Let $M := \{x \in \Omega \mid f(x) \geq f(x_1)\}$. Then $x_1 \in M \neq \emptyset$. Note that, by Proposition 2.5, $M = f_{\text{inv}}(\mathbb{R} \setminus I) = \Omega \setminus f_{\text{inv}}(I)$, where $I = (-\infty, f(x_1)) \subset \mathbb{R}$. Since I is open in $\mathbb{R}$ and f is weakly upper semicontinuous, then, $f_{\text{inv}}(I) \in \mathcal{O}_{\text{weak}, \Omega}$. Then, M is closed in $\mathcal{O}_{\text{weak}, \Omega}$. By Proposition 5.5, M is weakly compact. Since $f(x_1) = \max_{x \in \Omega} f(x)$, then M is an extreme subset of Ω . By Proposition 10.7, there exists $x_0 \in M$ that is an extreme point of Ω . This completes the proof of the proposition. $\square$

Proposition 10.15 Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^\circ$, and $f : \Omega \rightarrow \mathbb{R}$ be differentiable at x_0 . Assume that x_0 is a point of relative extremum of f , then $f^{(1)}(x_0) = \vartheta_{\mathcal{X}^*}$.

Proof Without loss of generality, assume that x_0 is a point of relative minimum for f . The case when x_0 is a point of relative maximum for f can be proved similarly. Since $x_0 \in \Omega^\circ$, then $A_\Omega(x_0) = \mathcal{X}$. $\forall u \in \mathcal{X}$, by Propositions 10.2 and 9.5, $Df(x_0; u) = f^{(1)}(x_0)u \geq 0$ and $Df(x_0; -u) = -f^{(1)}(x_0)u \geq 0$. Then, $f^{(1)}(x_0)u = 0$. By the arbitrariness of u , we have $f^{(1)}(x_0) = \vartheta_{\mathcal{X}^*}$. This completes the proof of the proposition. $\square$

Proposition 10.16 *Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^\circ$, and $f : \Omega \rightarrow \mathbb{R}$ be $\mathcal{C}_2$ at x_0 . Assume that x_0 is a point of relative minimum for f , then $f^{(1)}(x_0) = \vartheta_{\mathcal{X}^*}$ and $f^{(2)}(x_0)$ is positive semi-definite.*

Proof By Proposition 10.15, $f^{(1)}(x_0) = \vartheta_{\mathcal{X}^*}$. We will show that $f^{(2)}(x_0)$ is positive semi-definite by an argument of contradiction. Suppose $f^{(2)}(x_0)$ is not positive semi-definite. Then, $\exists h \in \mathcal{X}$ such that $f^{(2)}(x_0)(h)(h) < 0$. By the continuity of $f^{(2)}$ at x_0 , there exists $\delta \in (0, \infty) \subset \mathbb{R}$ with $D := \mathcal{B}_{\mathcal{X}}(x_0, \delta) \subseteq \Omega$ such that $f^{(2)}(x)(h)(h) < 0, \forall x \in D$. Let $x_1 := x_0 + \alpha h \in D$, where $\alpha \in (0, \infty) \subset \mathbb{R}$ is an arbitrary constant. By Taylor's Theorem 9.48, there exists $t_0 \in (0, 1) \subset \mathbb{R}$ such that

$$\begin{aligned} f(x_1) &= f(x_0) + \frac{1}{2}f^{(2)}(x_0 + t_0\alpha h)(x_1 - x_0)(x_1 - x_0) \\ &= f(x_0) + \frac{\alpha^2}{2}f^{(2)}(x_0 + t_0\alpha h)(h)(h) < f(x_0) \end{aligned}$$

This contradicts with the fact that x_0 is a point of relative minimum for f . Hence, the result must be true. This completes the proof of the proposition. $\square$

Proposition 10.17 *Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^\circ$, and $f : \Omega \rightarrow \mathbb{R}$ be twice differentiable at x , $\forall x \in D := \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) \subseteq \Omega$, where $\delta_0 \in (0, \infty) \subset \mathbb{R}$ is some constant. Assume that $f^{(1)}(x_0) = \vartheta_{\mathcal{X}^*}$ and $f^{(2)}(x)$ is positive semi-definite, $\forall x \in D$. Then, x_0 is a point of relative minimum for f .*

Proof $\forall x \in D$, by Taylor's Theorem 9.48, there exists $t_0 \in (0, 1) \subset \mathbb{R}$ such that

$$f(x) = f(x_0) + \frac{1}{2}f^{(2)}(x_0 + t_0(x - x_0))(x - x_0)(x - x_0) \geq f(x_0)$$

Hence, x_0 is a point of relative minimum for f . This completes the proof of the proposition. $\square$

Proposition 10.18 *Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^\circ$, and $f : \Omega \rightarrow \mathbb{R}$ be $\mathcal{C}_2$ at x_0 . Assume that $f^{(1)}(x_0) = \vartheta_{\mathcal{X}^*}$ and $f^{(2)}(x_0)$ is positive definite. Then, x_0 is a point of relative strict minimum for f .*

Proof By Proposition 10.4 and the continuity of $f^{(2)}$ at x_0 , $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $f^{(2)}(x)$ is positive definite, $\forall x \in D := \mathcal{B}_{\mathcal{X}}(x_0, \delta) \subseteq \Omega$. $\forall x \in D \setminus \{x_0\}$, by Taylor's Theorem, there exists $t_0 \in (0, 1) \subset \mathbb{R}$ such that

$$f(x) = f(x_0) + \frac{1}{2}f^{(2)}(x_0 + t_0(x - x_0))(x - x_0)(x - x_0) > f(x_0)$$

Hence, x_0 is a point of relative strict minimum for f . This completes the proof of the proposition. $\square$

10.3 Optimization with Equality Constraints

The basic problem to be considered in this section is

$$\begin{aligned} \mu_0 := \quad & \inf && f(x) \\ \text{subject to} & && x \in \Omega \\ & && H(x) = \vartheta_{\mathcal{Y}} \end{aligned} \quad (10.2)$$

where $\mathcal{X}$ and $\mathcal{Y}$ are real normed linear spaces, $\Omega \subseteq \mathcal{X}$ is a set, $f : \Omega \rightarrow \mathbb{R}$ is a functional, and $H : \Omega \rightarrow \mathcal{Y}$ is a function.

Definition 10.19 Let $\mathcal{X}$ and $\mathcal{Y}$ be real normed linear spaces, $\Omega \subseteq \mathcal{X}$, $H : \Omega \rightarrow \mathcal{Y}$ be $\mathcal{C}_1$ at $x_0 \in \Omega^\circ$. x_0 is said to be a regular point of H if $H^{(1)}(x_0) \in B(\mathcal{X}, \mathcal{Y})$ is surjective.

Lemma 10.20 Let $\mathcal{X}$ and $\mathcal{Y}$ be real Banach spaces, $\Omega \subseteq \mathcal{X}$, $f : \Omega \rightarrow \mathbb{R}$ be $\mathcal{C}_1$ at $x_0 \in \Omega^\circ$, and $H : \Omega \rightarrow \mathcal{Y}$ be $\mathcal{C}_1$ at x_0 . Consider the optimization problem (10.2). Assume that x_0 is a point of relative minimum for f on the set $\Omega_c := \{x \in \Omega \mid H(x) = \vartheta_{\mathcal{Y}}\}$; and x_0 is a regular point of H . Then, $\forall u \in \mathcal{X}$ with $H^{(1)}(x_0)u = \vartheta_{\mathcal{Y}}$, we have $f^{(1)}(x_0)u = 0$.

Proof Define $T : \Omega \rightarrow \mathbb{R} \times \mathcal{Y}$ by $T(x) = (f(x), H(x))$, $\forall x \in \Omega$. By Proposition 9.44, T is $\mathcal{C}_1$ at x_0 and $T^{(1)}(x_0) = \begin{bmatrix} f^{(1)}(x_0) \\ H^{(1)}(x_0) \end{bmatrix}$.

We will prove the result using an argument of contradiction. Suppose the result is not true. Then, $\exists u_0 \in \mathcal{X}$ with $H^{(1)}(x_0)u_0 = \vartheta_{\mathcal{Y}}$, we have $f^{(1)}(x_0)u_0 \neq 0$. We will show that $T^{(1)}(x_0)$ is surjective. $\forall (r, y) \in \mathbb{R} \times \mathcal{Y}$, by x_0 being a regular point of H , $\exists u_1 \in \mathcal{X}$ such that $H^{(1)}(x_0)u_1 = y$. Let $u = \frac{r - f^{(1)}(x_0)u_1}{f^{(1)}(x_0)u_0}u_0 + u_1$. Then, $T^{(1)}(x_0)u = (r, y)$. Hence, $T^{(1)}(x_0)$ is surjective.

Note that $T(x_0) = (f(x_0), \vartheta_{\mathcal{Y}})$ and $\mathbb{R} \times \mathcal{Y}$ is a real Banach space, by Propositions 7.22 and 4.31. By Surjective Mapping Theorem 9.53, $\exists \delta_r \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, and $c_1 \in [0, \infty) \subset \mathbb{R}$ with $c_1\delta \leq \delta_r$ such that $\forall (r, y) \in \mathcal{B}_{\mathbb{R} \times \mathcal{Y}}(T(x_0), \delta/2)$, $\exists x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta_r) \subseteq \Omega$ with $\|x - x_0\| \leq c_1\|(r, y) - T(x_0)\|$, we have $T(x) = (r, y)$. Then, $\forall \delta_r \in (0, \delta/2) \subset \mathbb{R}$, let $r_1 = f(x_0) - \delta_r \in \mathbb{R}$. $(r_1, \vartheta_{\mathcal{Y}}) \in \mathcal{B}_{\mathbb{R} \times \mathcal{Y}}(T(x_0), \delta/2)$. Then, $\exists x_1 \in \Omega$ with $\|x_1 - x_0\| \leq c_1\|(r_1, \vartheta_{\mathcal{Y}}) - T(x_0)\| = c_1|r_1 - f(x_0)| = c_1\delta_r < (1 + c_1)\delta_r$ such that $T(x_1) = (r_1, \vartheta_{\mathcal{Y}})$. Then, $H(x_1) = \vartheta_{\mathcal{Y}}$ and $x_1 \in \Omega_c \cap \mathcal{B}_{\mathcal{X}}(x_0, (1 + c_1)\delta_r)$. Furthermore, $f(x_1) = r_1 < f(x_0)$. This contradicts with the assumption that x_0 is a point of relative minimum for f on Ω_c . Therefore, the result must be true.

This completes the proof of the lemma. $\square$

Proposition 10.21 (Lagrange Multiplier) Let $\mathcal{X}$ and $\mathcal{Y}$ be real Banach spaces, $\Omega \subseteq \mathcal{X}$, $f : \Omega \rightarrow \mathbb{R}$ be $\mathcal{C}_1$ at $x_0 \in \Omega^\circ$, and $H : \Omega \rightarrow \mathcal{Y}$ be $\mathcal{C}_1$ at x_0 . Consider the optimization problem (10.2). Assume that x_0 is a

point of relative minimum for f on the set $\Omega_c := \{x \in \Omega \mid H(x) = \vartheta_{\mathcal{Y}}\}$; and x_0 is a regular point of H . Then, there exists a Lagrange multiplier $y_{*0} \in \mathcal{Y}^*$ such that the Lagrangian $L : \Omega \times \mathcal{Y}^* \rightarrow \mathbb{R}$ defined by $L(x, y_*) = f(x) + \langle y_*, H(x) \rangle$, $\forall (x, y_*) \in \Omega \times \mathcal{Y}^*$, is stationary at (x_0, y_{*0}) , that is $L^{(1)}(x_0, y_{*0}) = \vartheta_{\mathcal{B}(\mathcal{X} \times \mathcal{Y}^*, \mathbb{R})}$.

Proof By Lemma 10.20, $\forall u \in \mathcal{X}$ with $H^{(1)}(x_0)u = \vartheta_{\mathcal{Y}}$, we have $f^{(1)}(x_0)u = 0$. Then, $f^{(1)}(x_0) \in (\mathcal{N}(H^{(1)}(x_0)))^\perp$. Since x_0 is a regular point of H , then $\mathcal{R}(H^{(1)}(x_0)) = \mathcal{Y}$ is closed. By Proposition 7.114, we have $(\mathcal{N}(H^{(1)}(x_0)))^\perp = \mathcal{R}((H^{(1)}(x_0))')$. Then, there exists $y_{*0} \in \mathcal{Y}^*$ such that $f^{(1)}(x_0) = -(H^{(1)}(x_0))' y_{*0}$. By Propositions 9.34, 9.41, 9.38, 9.37, 9.44, and 9.45, the Lagrangian L is $\mathcal{C}_1$ at (x_0, y_{*0}) and, $\forall (u, v_*) \in \mathcal{X} \times \mathcal{Y}^*$,

$$\begin{aligned} L^{(1)}(x_0, y_{*0})(u, v_*) &= f^{(1)}(x_0)u + \langle v_*, H(x_0) \rangle + \langle y_{*0}, H^{(1)}(x_0)u \rangle \\ &= f^{(1)}(x_0)u + \left\langle (H^{(1)}(x_0))' y_{*0}, u \right\rangle = 0 \end{aligned}$$

where the second equality follows from that fact that $x_0 \in \Omega_c$. Hence, $L^{(1)}(x_0, y_{*0}) = \vartheta_{\mathcal{B}(\mathcal{X} \times \mathcal{Y}^*, \mathbb{R})}$. This completes the proof of the proposition. $\square$

Proposition 10.22 (Generalized Lagrange Multiplier) Let $\mathcal{X}$ and $\mathcal{Y}$ be real Banach spaces, $\Omega \subseteq \mathcal{X}$, $f : \Omega \rightarrow \mathbb{R}$ be $\mathcal{C}_1$ at $x_0 \in \Omega^\circ$, and $H : \Omega \rightarrow \mathcal{Y}$ be $\mathcal{C}_1$ at x_0 . Consider the optimization problem (10.2). Assume that x_0 is a point of relative minimum for f on the set $\Omega_c := \{x \in \Omega \mid H(x) = \vartheta_{\mathcal{Y}}\}$; and $\mathcal{R}(H^{(1)}(x_0)) \subseteq \mathcal{Y}$ is closed. Then, there exists a Lagrange multiplier $(r_0, y_{*0}) \in \mathbb{R} \times \mathcal{Y}^*$ with $(r_0, y_{*0}) \neq (0, \vartheta_{\mathcal{Y}^*})$ such that the Lagrangian $L : \Omega \times \mathcal{Y}^* \rightarrow \mathbb{R}$ defined by $L(x, y_*) = r_0 f(x) + \langle y_*, H(x) \rangle$, $\forall (x, y_*) \in \Omega \times \mathcal{Y}^*$, is stationary at (x_0, y_{*0}) , that is $L^{(1)}(x_0, y_{*0}) = \vartheta_{\mathcal{B}(\mathcal{X} \times \mathcal{Y}^*, \mathbb{R})}$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{R}(H^{(1)}(x_0)) = \mathcal{Y}$; Case 2: $\mathcal{R}(H^{(1)}(x_0)) \subset \mathcal{Y}$.

Case 1: $\mathcal{R}(H^{(1)}(x_0)) = \mathcal{Y}$. Take $r_0 = 1$. Clearly, x_0 is a regular point of H . The result follows from Proposition 10.21.

Case 2: $\mathcal{R}(H^{(1)}(x_0)) \subset \mathcal{Y}$. Take $r_0 = 0$. Clearly, x_0 is not a regular point of H . Then, $\exists y_0 \in \mathcal{Y} \setminus \mathcal{R}(H^{(1)}(x_0))$. Let $M = \mathcal{R}(H^{(1)}(x_0))$, which is a closed subspace of $\mathcal{Y}$ by the assumption. Then, by Proposition 4.10, $\delta := \inf_{m \in M} \|y_0 - m\| > 0$. By Proposition 7.97, we have $\delta = \max_{y_* \in M^\perp, \|y_*\| \leq 1} \langle y_*, y_0 \rangle$, where the maximum is achieved at $y_{*0} \in M^\perp$. Then, $y_{*0} \neq \vartheta_{\mathcal{Y}^*}$ and, by Proposition 7.112, $y_{*0} \in \mathcal{N}((H^{(1)}(x_0))')$. The Lagrangian L is given by $L(x, y_*) = \langle y_*, H(x) \rangle$, $\forall (x, y_*) \in \Omega \times \mathcal{Y}^*$. By Propositions 9.34, 9.41, 9.37, 9.44, and 9.45, L is $\mathcal{C}_1$ at (x_0, y_{*0}) and, $\forall (u, v_*) \in \mathcal{X} \times \mathcal{Y}^*$,

$$\begin{aligned} L^{(1)}(x_0, y_{*0})(u, v_*) &= \langle v_*, H(x_0) \rangle + \langle y_{*0}, H^{(1)}(x_0)u \rangle \\ &= \left\langle (H^{(1)}(x_0))' y_{*0}, u \right\rangle = \langle \vartheta_{\mathcal{X}^*}, u \rangle = 0 \end{aligned}$$

where the second equality follows from that fact that $x_0 \in \Omega_c$. Hence, $L^{(1)}(x_0, y_{*0}) = \vartheta_{B(X \times Y^*, \mathbb{R})}$. This case is proved.

This completes the proof of the proposition. $\square$

Proposition 10.23 *Let X and Y be real Banach spaces, $\Omega \subseteq X$, $f : \Omega \rightarrow \mathbb{R}$ be $\mathcal{C}_2$ at $x_0 \in \Omega^\circ$, $H : \Omega \rightarrow Y$ be $\mathcal{C}_2$ at x_0 . Consider the optimization problem (10.2). Assume that*

- (i) $H(x_0) = \vartheta_Y$ and $\mathcal{R}(H^{(1)}(x_0)) \subseteq Y$ is closed;
- (ii) the Lagrangian $L : \Omega \times Y^* \rightarrow \mathbb{R}$ defined by $L(x, y_*) = f(x) + \langle y_*, H(x) \rangle$, $\forall (x, y_*) \in \Omega \times Y^*$, is stationary at (x_0, y_{*0}) , where $y_{*0} \in Y^*$ is a Lagrange multiplier;
- (iii) $\frac{\partial^2 L}{\partial x^2}(x_0, y_{*0})$ is positive definite on the subspace $M := \mathcal{N}(H^{(1)}(x_0))$, that is, $\exists m \in (0, \infty) \subset \mathbb{R}$ such that $\frac{\partial^2 L}{\partial x^2}(x_0, y_{*0})(h)(h) \geq m \|h\|^2$, $\forall h \in \mathcal{N}(H^{(1)}(x_0))$.

Then, x_0 is a point of relative strict minimum for f on the set $\Omega_c := \{x \in \Omega \mid H(x) = \vartheta_Y\}$.

Proof By Propositions 9.34, 9.37, 9.38, 9.41, 9.44, and 9.45, L is $\mathcal{C}_2$ at (x_0, y_{*0}) . By Proposition 9.9 and (ii),

$$L^{(1)}(x_0, y_{*0}) = \begin{bmatrix} \frac{\partial L}{\partial x}(x_0, y_{*0}) & \frac{\partial L}{\partial y_*}(x_0, y_{*0}) \end{bmatrix} = \vartheta_{B(X \times Y^*, \mathbb{R})}$$

Then, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $f^{(2)}(x)$, $H^{(2)}(x)$, and $\frac{\partial^2 L}{\partial x^2}(x, y_*)$ exists, $\forall x \in \mathcal{B}_X(x_0, \delta_0) \subseteq \Omega$ and $\forall y_* \in \mathcal{B}_{Y^*}(y_{*0}, \delta_0)$, and, by Proposition 9.46, $\frac{\partial^2 L}{\partial x^2}$ is continuous at (x_0, y_{*0}) . Then, $\exists \delta_1 \in (0, \delta_0] \subset \mathbb{R}$ and $\exists c_1 \in [0, \infty) \subset \mathbb{R}$ such that $\left\| \frac{\partial^2 L}{\partial x^2}(x, y_{*0}) - \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right\| < m/5$ and $\|H^{(2)}(x)\| \leq c_1$, $\forall x \in D_1 := \mathcal{B}_X(x_0, \delta_1)$. By Propositions 7.114 and 7.98 and (i), $M^\perp = \mathcal{R}\left((H^{(1)}(x_0))'\right)$ is closed. Then, by Proposition 7.113, $\exists c_2 \in [0, \infty) \subset \mathbb{R}$ such that, $\forall x_* \in \mathcal{R}\left((H^{(1)}(x_0))'\right)$, there exists $y_* \in Y^*$ such that $x_* = (H^{(1)}(x_0))' y_*$ and $\|y_*\| \leq c_2 \|x_*\|$.

Claim 10.23.1 $\forall x \in \Omega_c \cap D_1$, $\exists h_0 \in M$ such that $\|x - x_0 - h_0\| \leq c_1 c_2 \|x - x_0\|^2$, $\|x - x_0 - c_1 c_2 \|x - x_0\|^2\| \leq \|h_0\| \leq \|x - x_0\| + c_1 c_2 \|x - x_0\|^2$.

Proof of claim: Fix any $x \in \Omega_c \cap D_1$. Then, $H(x) = H(x_0) = \vartheta_Y$. By Taylor's Theorem 9.48, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ such that $\|H^{(1)}(x_0)(x - x_0)\| = \|H(x) - H(x_0) - H^{(1)}(x_0)(x - x_0)\| \leq \frac{1}{2} \|H^{(2)}(t_0 x + (1 - t_0)x_0)\| \|x - x_0\|^2 \leq c_1 \|x - x_0\|^2 / 2$. By Proposition 7.97, we have $\inf_{h \in M} \|x - x_0 - h\| = \max_{x_* \in M^\perp, \|x_*\| \leq 1} \langle x_*, x - x_0 \rangle$, where the maximum is achieved at $x_{*1} \in M^\perp = \mathcal{R}\left((H^{(1)}(x_0))'\right)$ with $\|x_{*1}\| \leq 1$. Then, $\exists y_{*1} \in Y^*$ such that

$x_{*1} = (H^{(1)}(x_0))' y_{*1}$ and $\|y_{*1}\| \leq c_2$. By Proposition 7.68, M is closed. Then, $\exists h_0 \in M$ such that $\|x - x_0 - h_0\| \leq 2 \inf_{h \in M} \|x - x_0 - h\| = 2 \langle \langle x_{*1}, x - x_0 \rangle \rangle = 2 \langle \langle y_{*1}, H^{(1)}(x_0)(x - x_0) \rangle \rangle \leq 2 \|y_{*1}\| \|H^{(1)}(x_0)(x - x_0)\| \leq c_1 c_2 \|x - x_0\|^2$.

Then, $\|h_0\| \geq \|x - x_0\| - \|x - x_0 - h_0\| \geq \|x - x_0\| - c_1 c_2 \|x - x_0\|^2$; and $\|h_0\| \leq \|x - x_0\| + \|x - x_0 - h_0\| \leq \|x - x_0\| + c_1 c_2 \|x - x_0\|^2$. This completes the proof of the claim. $\square$

Let $c_3 = \left\| \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right\|$. Let $\delta \in (0, \delta_1] \subset \mathbb{R}$ such that $c_1 c_2 \delta \leq 1/4$, $c_1^2 c_2^2 (c_3 + m/5) \delta^2 \leq m/5$, and $5c_1 c_2 (c_3 + m/5) \delta/2 \leq m/5$. $\forall x \in \Omega_c \cap \mathcal{B}_X(x_0, \delta)$, by Claim 10.23.1, $\exists h_0 \in M$ such that $\|x - x_0 - h_0\| \leq c_1 c_2 \|x - x_0\|^2 \leq \|x - x_0\|/4$, $3\|x - x_0\|/4 \leq \|h_0\| \leq 5\|x - x_0\|/4$. By Taylor's Theorem 9.48, $\exists t_1 \in (0, 1) \subset \mathbb{R}$ such that

$$\begin{aligned}
f(x) - f(x_0) &= L(x, y_{*0}) - L(x_0, y_{*0}) - \frac{\partial L}{\partial x}(x_0, y_{*0})(x - x_0) \\
&= \frac{1}{2} \frac{\partial^2 L}{\partial x^2}(\underbrace{t_1 x + (1 - t_1)x_0}_{\bar{x}}, y_{*0})(x - x_0)(x - x_0) \\
&= \frac{1}{2} \left(\frac{\partial^2 L}{\partial x^2}(\bar{x}, y_{*0})(h_0)(h_0) + \frac{\partial^2 L}{\partial x^2}(\bar{x}, y_{*0})(x - x_0 - h_0)(h_0) \right. \\
&\quad \left. + \frac{\partial^2 L}{\partial x^2}(\bar{x}, y_{*0})(h_0)(x - x_0 - h_0) \right. \\
&\quad \left. + \frac{\partial^2 L}{\partial x^2}(\bar{x}, y_{*0})(x - x_0 - h_0)(x - x_0 - h_0) \right) \\
&\geq \frac{1}{2} \left(\frac{\partial^2 L}{\partial x^2}(x_0, y_{*0})(h_0)(h_0) + \left(\frac{\partial^2 L}{\partial x^2}(\bar{x}, y_{*0}) - \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right)(h_0)(h_0) \right. \\
&\quad \left. - \left\| \frac{\partial^2 L}{\partial x^2}(\bar{x}, y_{*0}) \right\| (2\|x - x_0 - h_0\| \|h_0\| + \|x - x_0 - h_0\|^2) \right) \\
&\geq \frac{1}{2} \left(m\|h_0\|^2 - m\|h_0\|^2/5 - \left(\left\| \frac{\partial^2 L}{\partial x^2}(\bar{x}, y_{*0}) - \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right\| \right. \right. \\
&\quad \left. \left. + \left\| \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right\| \right) (2\|x - x_0 - h_0\| \|h_0\| + \|x - x_0 - h_0\|^2) \right) \\
&\geq \frac{1}{2} \left(\frac{4m}{5} \|h_0\|^2 - (c_3 + m/5) \left(\frac{5}{2} c_1 c_2 \|x - x_0\|^3 + c_1^2 c_2^2 \|x - x_0\|^4 \right) \right) \\
&\geq \frac{1}{2} \left(\frac{4m}{5} \frac{9}{16} - \frac{m}{5} - \frac{m}{5} \right) \|x - x_0\|^2 = \frac{m}{40} \|x - x_0\|^2
\end{aligned}$$

Hence, x_0 is a point of relative strict minimum for f on the set Ω_c . This completes the proof of the proposition. $\square$

Proposition 10.24 *Let $\mathcal{X}$ and $\mathcal{Y}$ be real Banach spaces, $\Omega \subseteq \mathcal{X}$, $f : \Omega \rightarrow \mathbb{R}$ be $\mathcal{C}_2$ at $x_0 \in \Omega^\circ$, $H : \Omega \rightarrow \mathcal{Y}$ be $\mathcal{C}_2$ at x_0 . Consider the optimization problem (10.2). Assume that x_0 is a regular point of H and is a point of relative minimum for f on the set $\Omega_c := \{x \in \Omega \mid H(x) = \vartheta_Y\}$. Then, there exists a Lagrange multiplier $y_{*0} \in \mathcal{Y}^*$ such that*

- (i) the Lagrangian $L : \Omega \times \mathcal{Y}^* \rightarrow \mathbb{R}$ defined by $L(x, y_*) = f(x) + \langle y_*, H(x) \rangle$, $\forall (x, y_*) \in \Omega \times \mathcal{Y}^*$, is stationary at (x_0, y_{*0}) ;
- (ii) $\frac{\partial^2 L}{\partial x^2}(x_0, y_{*0})$ is positive semi-definite on the subspace $\mathcal{N}(H^{(1)}(x_0)) =: M$, that is, $\frac{\partial^2 L}{\partial x^2}(x_0, y_{*0})(h)(h) \geq 0$, $\forall h \in \mathcal{N}(H^{(1)}(x_0))$.

Proof Under the assumption of the proposition, by Proposition 10.21, there exists a Lagrange multiplier $y_{*0} \in \mathcal{Y}^*$ such that (i) holds. We will show that (ii) also holds by an argument of contradiction. Suppose (ii) does not hold. Then, $\exists h_0 \in M$ with $\|h_0\| = 1$ such that $\frac{\partial^2 L}{\partial x^2}(x_0, y_{*0})(h_0)(h_0) < -m < 0$ for some $m \in (0, \infty) \subset \mathbb{R}$. By Surjective Mapping Theorem 9.53, $\exists r_1 \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$, and $\exists c_1 \in [0, \infty) \subset \mathbb{R}$ with $c_1 \delta_1 \leq r_1$ such that $\forall \bar{y} \in \mathcal{B}_Y(\vartheta_Y, \delta_1/2)$, $\forall \bar{x} \in \mathcal{B}_X(x_0, r_1/2)$ with $\bar{y} = H(\bar{x})$, $\forall y \in \mathcal{B}_Y(\bar{y}, \delta_1/2)$, $\exists x \in \mathcal{B}_X(x_0, r_1) \subseteq \Omega$ with $\|x - \bar{x}\| \leq c_1 \|y - \bar{y}\|$, we have $y = H(x)$. By Propositions 9.34, 9.37, 9.38, 9.41, 9.44, and 9.45, L is $\mathcal{C}_2$ at (x_0, y_{*0}) . By Proposition 9.9 and (i),

$$L^{(1)}(x_0, y_{*0}) = \begin{bmatrix} \frac{\partial L}{\partial x}(x_0, y_{*0}) & \frac{\partial L}{\partial y_*}(x_0, y_{*0}) \end{bmatrix} = \vartheta_{B(X \times Y^*, \mathbb{R})}$$

By Proposition 9.46, $\exists r_2 \in (0, r_1] \subset \mathbb{R}$ such that $\frac{\partial^2 L}{\partial x^2}(x, y_{*0})$ exists and $\left\| \frac{\partial^2 L}{\partial x^2}(x, y_{*0}) - \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right\| < m/5$, $\forall x \in \mathcal{B}_X(x_0, r_2)$. Since H is $\mathcal{C}_2$ at x_0 , then $\exists c_2 \in [0, \infty) \subset \mathbb{R}$ and $\exists r_3 \in (0, r_2] \subset \mathbb{R}$ such that $\|H^{(2)}(x)\| \leq c_2$, $\forall x \in \mathcal{B}_X(x_0, r_3)$. Let $c_3 := \left\| \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right\|$.

$\forall \delta \in (0, r_3/2) \subset \mathbb{R}$ such that $c_2 \delta^2 < \delta_1$, $c_1 c_2 \delta/2 \leq 1/4$, $c_1^2 c_2^2 (c_3 + m/5) \delta^2/4 \leq m/5$, and $c_1 c_2 (c_3 + m/5) \delta \leq m/5$. By Taylor's Theorem 9.48, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ such that

$$\begin{aligned} \|H(x_0 + \delta h_0)\| &= \|H(x_0 + \delta h_0) - H(x_0) - \delta H^{(1)}(x_0)h_0\| \\ &\leq \frac{1}{2} \|H^{(2)}(x_0 + t_0 \delta h_0)\| \delta^2 \leq c_2 \delta^2/2 < \delta_1/2 \end{aligned}$$

Let $\bar{y}_\delta := H(x_0 + \delta h_0) \in \mathcal{B}_Y(\vartheta_Y, \delta_1/2)$ and $\bar{x}_\delta := x_0 + \delta h_0 \in \mathcal{B}_X(x_0, r_1/2)$. Note that $\bar{y}_\delta = H(\bar{x}_\delta)$ and $\vartheta_Y \in \mathcal{B}_Y(\bar{y}_\delta, \delta_1/2)$. Then, $\exists x_\delta \in \mathcal{B}_X(x_0, r_1)$ with $\|x_\delta - \bar{x}_\delta\| \leq c_1 \|\bar{y}_\delta\| \leq c_1 c_2 \delta^2/2$ such that $H(x_\delta) = \vartheta_Y$. Then, we have $\|x_\delta - x_0\| \geq \|\delta h_0\| - \|x_\delta - \bar{x}_\delta\| \geq \delta - c_1 c_2 \delta^2/2 \geq 3\delta/4$; and $\|x_\delta - x_0\| \leq \|\delta h_0\| + \|x_\delta - \bar{x}_\delta\| \leq \delta + c_1 c_2 \delta^2/2 \leq 5\delta/4 < r_3$. Hence, $x_\delta \in \Omega_c \cap \mathcal{B}_X(x_0, r_3)$. By Taylor's Theorem 9.48, $\exists t_1 \in (0, 1) \subset \mathbb{R}$ such that

$$\begin{aligned} f(x_\delta) - f(x_0) &= L(x_\delta, y_{*0}) - L(x_0, y_{*0}) - \frac{\partial L}{\partial x}(x_0, y_{*0})(x_\delta - x_0) \\ &= \frac{1}{2} \frac{\partial^2 L}{\partial x^2}(\underbrace{t_1 x_\delta + (1 - t_1) x_0}_{\bar{x}}, y_{*0})(x_\delta - x_0)(x_\delta - x_0) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \left(\frac{\partial^2 L}{\partial x^2}(\hat{x}, y_{*0})(\bar{x}_\delta - x_0)(\bar{x}_\delta - x_0) \right. \\
&\quad + \frac{\partial^2 L}{\partial x^2}(\hat{x}, y_{*0})(x_\delta - \bar{x}_\delta)(\bar{x}_\delta - x_0) + \frac{\partial^2 L}{\partial x^2}(\hat{x}, y_{*0})(\bar{x}_\delta - x_0)(x_\delta - \bar{x}_\delta) \\
&\quad \left. + \frac{\partial^2 L}{\partial x^2}(\hat{x}, y_{*0})(x_\delta - \bar{x}_\delta)(x_\delta - \bar{x}_\delta) \right) \\
&\leq \frac{1}{2} \left(\delta^2 \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0})(h_0)(h_0) + \delta^2 \left(\frac{\partial^2 L}{\partial x^2}(\hat{x}, y_{*0}) \right. \right. \\
&\quad \left. \left. - \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right)(h_0)(h_0) + \left\| \frac{\partial^2 L}{\partial x^2}(\hat{x}, y_{*0}) \right\| (2 \|x_\delta - \bar{x}_\delta\| \delta \|h_0\| \right. \\
&\quad \left. + \|x_\delta - \bar{x}_\delta\|^2) \right) \\
&< \frac{1}{2} \left(-m\delta^2 + m\delta^2/5 + \left(\left\| \frac{\partial^2 L}{\partial x^2}(\hat{x}, y_{*0}) - \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right\| \right. \right. \\
&\quad \left. \left. + \left\| \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right\| \right) (c_1 c_2 \delta^3 + c_1^2 c_2^2 \delta^4/4) \right) \\
&\leq \frac{1}{2} \left(-\frac{4m}{5} \delta^2 + (c_3 + m/5) (c_1 c_2 \delta^3 + c_1^2 c_2^2 \delta^4/4) \right) \\
&\leq \frac{1}{2} \left(-\frac{4m}{5} \delta^2 + \frac{m}{5} \delta^2 + \frac{m}{5} \delta^2 \right) = -\frac{m}{5} \delta^2 < 0
\end{aligned}$$

Thus, $\lim_{\delta \rightarrow 0^+} x_\delta = x_0$, $x_\delta \in \Omega_c$, and $f(x_\delta) < f(x_0)$, for sufficiently small δ . Then, x_0 is not a point of relative minimum for f on the set Ω_c . This contradicts with the assumption. Hence, (ii) must hold. This completes the proof of the proposition. $\square$

10.4 Inequality Constraints

The basic problem to be considered in this section is

$$\begin{aligned}
\mu_0 := & \inf_{\substack{x \in \Omega \\ G(x) \leq \vartheta_{\mathcal{Z}}}} f(x) \tag{10.3}
\end{aligned}$$

where $\mathcal{X}$ is a real normed linear spaces, $\Omega \subseteq \mathcal{X}$ is a set, $f : \Omega \rightarrow \mathbb{R}$ is a functional, $\mathcal{Z}$ is a real normed linear space with a positive cone $P \subseteq \mathcal{Z}$, and $G : \Omega \rightarrow \mathcal{Z}$ is a function.

Theorem 10.25 (Generalized Kuhn-Tucker Theorem) *Let $\mathcal{X}$ be a real Banach space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^\circ$, $\mathcal{Z}$ be a real Banach space with a positive cone $P \subseteq \mathcal{Z}$, and $f : \Omega \rightarrow \mathbb{R}$ and $G : \Omega \rightarrow \mathcal{Z}$ be $\mathcal{C}_1$ at x_0 . Consider the optimization problem (10.3). Assume that x_0 is a regular point of G and is a point of relative minimum for f on the set $\Omega_c := \left\{ x \in \Omega \mid G(x) \leq \vartheta_{\mathcal{Z}} \right\}$. Then, there exists a Lagrange multiplier $z_{*0} \in \mathcal{Z}^*$ with $z_{*0} \geq \vartheta_{\mathcal{Z}^*}$ such that*

the Lagrangian $L : \Omega \rightarrow \mathbb{R}$ defined by $L(x) = f(x) + \langle z_{*0}, G(x) \rangle$, $\forall x \in \Omega$, is stationary at x_0 and $\langle z_{*0}, G(x_0) \rangle = 0$.

Proof Consider the small variation equivalent problem of (10.3) at x_0 :

$$\begin{aligned} \bar{\mu}_0 := \inf_{\text{subject to}} \quad & f^{(1)}(x_0)u \\ & u \in \mathcal{X} \\ & G(x_0) + G^{(1)}(x_0)u \leq \vartheta_Z \end{aligned} \quad (10.4)$$

This problem is convex, and belongs to the class of problems studied in Section 8.8. The primal function for (10.4) is $\omega : \bar{\Gamma} \rightarrow \mathbb{R}_e$, where

$$\bar{\Gamma} := \left\{ \bar{z} \in \mathcal{Z} \mid \exists u \in \mathcal{X} \ni G(x_0) + G^{(1)}(x_0)u \leq \bar{z} \right\} \quad (10.5a)$$

$$\omega(\bar{z}) = \inf_{u \in \mathcal{X}, G(x_0) + G^{(1)}(x_0)u \leq \bar{z}} f^{(1)}(x_0)u; \quad \forall \bar{z} \in \bar{\Gamma} \quad (10.5b)$$

Since $G^{(1)}(x_0) \in B(\mathcal{X}, \mathcal{Z})$ is surjective, then $\bar{\Gamma} = \mathcal{Z}$. Note that $\bar{\mu}_0 = \omega(\vartheta_Z)$. Since $x_0 \in \Omega_c$, then $z_0 := G(x_0) \leq \vartheta_Z$. Hence, $\bar{\mu}_0 \leq 0$.

Claim 10.25.1 $\bar{\mu}_0 = 0$.

Proof of claim: We will prove this using an argument of contradiction. Suppose $\bar{\mu}_0 < 0$. Then, $\exists u_0 \in \mathcal{X}$ with $\|u_0\| > 0$ such that $f^{(1)}(x_0)u_0 = -m\|u_0\| < 0$ and $G(x_0) + G^{(1)}(x_0)u_0 \leq \vartheta_Z$. By Surjective Mapping Theorem 9.53, $\exists r_1 \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$, and $\exists c_1 \in [0, \infty) \subset \mathbb{R}$ with $c_1\delta_1 \leq r_1$ such that $\forall \bar{z} \in \mathcal{B}_Z(z_0, \delta_1/2)$, $\forall \bar{x} \in \mathcal{B}_X(x_0, r_1/2)$ with $\bar{z} = G(\bar{x})$, $\forall z \in \mathcal{B}_Z(\bar{z}, \delta_1/2)$, $\exists x \in \mathcal{B}_X(x_0, r_1) \subseteq \Omega$ with $\|x - \bar{x}\| \leq c_1\|z - \bar{z}\|$, we have $z = G(x)$. Let $c_2 := \|G^{(1)}(x_0)\|$ and $c_3 := \|f^{(1)}(x_0)\|$. By f and G being $\mathcal{C}_1$ at x_0 , $\exists r_2 \in (0, r_1] \subset \mathbb{R}$ such that $\forall x \in \mathcal{B}_X(x_0, r_2)$, $\|f^{(1)}(x) - f^{(1)}(x_0)\| < m/3$ and $\|G^{(1)}(x) - G^{(1)}(x_0)\| < \frac{m}{3(1+c_1+c_1(c_3+m/3))}$.

$\forall \delta \in (0, 1) \subset \mathbb{R}$ such that $\delta\|u_0\| < r_2/2$, $(c_2 + m/3)\delta\|u_0\| < \delta_1/2$, and $\delta m\|u_0\|/3 < r_2/2$. By Mean Value Theorem 9.23, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ such that $\|G(x_0 + \delta u_0) - G(x_0) - \delta G^{(1)}(x_0)u_0\| \leq \|(G^{(1)}(x_0 + t_0\delta u_0) - G^{(1)}(x_0))(\delta u_0)\| < \frac{m}{3(1+c_1+c_1(c_3+m/3))}\delta\|u_0\| < \delta_1/2$. Let $\bar{x}_\delta := x_0 + \delta u_0 \in \mathcal{B}_X(x_0, r_2/2)$, $\bar{z}_\delta := G(\bar{x}_\delta)$, and $z_\delta := G(x_0) + \delta G^{(1)}(x_0)u_0 \in \mathcal{B}_Z(\bar{z}_\delta, \delta_1/2)$. Note that $\|\bar{z}_\delta - z_0\| \leq \|G(x_0 + \delta u_0) - G(x_0) - \delta G^{(1)}(x_0)u_0\| + \|\delta G^{(1)}(x_0)u_0\| < c_2\delta\|u_0\| + \frac{m}{3(1+c_1+c_1(c_3+m/3))}\delta\|u_0\| < \delta_1/2$ and $\bar{z}_\delta \in \mathcal{B}_Z(z_0, \delta_1/2)$. Then, $\exists x_\delta \in \mathcal{B}_X(x_0, r_1)$ with $\|x_\delta - \bar{x}_\delta\| \leq c_1\|z_\delta - \bar{z}_\delta\| \leq \frac{c_1 m}{3(1+c_1+c_1(c_3+m/3))}\delta\|u_0\| < r_2/2$ such that $z_\delta = G(x_\delta)$. Note that $G(x_\delta) = z_\delta = (1 - \delta)G(x_0) + \delta(G(x_0) + G^{(1)}(x_0)u_0) \leq \vartheta_Z$. Then, $x_\delta \in \Omega_c$. Note also that $x_\delta \in \mathcal{B}_X(x_0, r_2)$. By Mean Value Theorem 9.20, $\exists t_1, t_2 \in (0, 1) \subset \mathbb{R}$ such that

$$f(x_\delta) - f(x_0) = f(x_\delta) - f(\bar{x}_\delta) + f(\bar{x}_\delta) - f(x_0)$$

$$\begin{aligned}
&= f^{(1)}(t_1 x_\delta + (1 - t_1) \bar{x}_\delta) (x_\delta - \bar{x}_\delta) + \delta f^{(1)}(x_0 + t_2 \delta u_0) u_0 \\
&\leq (\|f^{(1)}(x_0)\| + \|f^{(1)}(t_1 x_\delta + (1 - t_1) \bar{x}_\delta) - f^{(1)}(x_0)\|) \|x_\delta - \bar{x}_\delta\| \\
&\quad + \delta f^{(1)}(x_0) u_0 + \delta \|f^{(1)}(x_0 + t_2 \delta u_0) - f^{(1)}(x_0)\| \|u_0\| \\
&\leq -m\delta \|u_0\| + (c_3 + m/3) \frac{c_1 m}{3(1 + c_1 + c_1(c_3 + m/3))} \delta \|u_0\| \\
&\quad + m\delta \|u_0\| / 3 < -m\delta \|u_0\| / 3 < 0
\end{aligned}$$

Note that $\|x_\delta - x_0\| \leq \|x_\delta - \bar{x}_\delta\| + \delta \|u_0\| < (1 + m/3)\delta \|u_0\|$. Then, we have shown that $f(x_\delta) < f(x_0)$, $x_\delta \in \Omega_c$, and $\lim_{\delta \rightarrow 0^+} x_\delta = x_0$. This contradicts with the assumption that x_0 is a point of relative minimum for f on Ω_c . Hence, $\bar{\mu}_0 = 0$.

This completes the proof of the claim. $\square$

Then, $\omega(\vartheta_Z) = 0$. By Fact 8.52, $\omega : Z \rightarrow \mathbb{R}$ is real-valued and convex.

Next, we will show that ω is continuous at ϑ_Z by Proposition 8.22. By Proposition 7.113, $\exists c_4 \in [0, \infty) \subset \mathbb{R}$, $\forall \bar{z} \in Z$, $\exists \bar{u} \in \mathcal{X}$ such that $\bar{z} = G^{(1)}(x_0) \bar{u}$ and $\|\bar{u}\| \leq c_4 \|\bar{z}\|$. Then, $G(x_0) + G^{(1)}(x_0) \bar{u} = G(x_0) + \bar{z} \stackrel{\leq}{=} \bar{z}$ and $f^{(1)}(x_0) \bar{u} \leq \|f^{(1)}(x_0)\| \|\bar{u}\| \leq c_4 \|f^{(1)}(x_0)\| \|\bar{z}\|$. This implies that $\forall \bar{z} \in \mathcal{B}_Z(\vartheta_Z, 1/2)$, $\omega(\bar{z}) \leq c_4 \|f^{(1)}(x_0)\| / 2 =: r_0 - 1/2$. Then, $\mathcal{B}_{\mathbb{R} \times Z}((r_0, \vartheta_Z), 1/2) \subseteq [\omega, \bar{\Gamma}]$. Hence, $(r_0, \vartheta_Z) \in [\omega, \bar{\Gamma}]^\circ$. By Proposition 8.22, ω is continuous at ϑ_Z . By Proposition 8.23, ω is continuous.

By Proposition 8.34, we have

$$0 = \bar{\mu}_0 = \omega(\vartheta_Z) = \max_{z_* \in \bar{\Gamma}_{\text{conj}}} (-\omega_{\text{conj}}(z_*))$$

where $\bar{\Gamma}_{\text{conj}}$ and $\omega_{\text{conj}} : \bar{\Gamma}_{\text{conj}} \rightarrow \mathbb{R}$ are the conjugate set and conjugate functional of ω . By Fact 8.53, $\bar{\Gamma}_{\text{conj}} \subseteq P^\ominus$. Let the maximum be achieved at $-z_{*0} \in \bar{\Gamma}_{\text{conj}}$, where $z_{*0} \stackrel{\geq}{=} \vartheta_{Z^*}$. Define $\bar{\omega} : P^\ominus \rightarrow \mathbb{R}_e$ by $\bar{\omega}(z_*) = \sup_{z \in \bar{\Gamma}} (\langle z_*, z \rangle - \omega(z))$, $\forall z_* \in P^\ominus$. By Fact 8.54, we have $\max_{z_* \in \bar{\Gamma}_{\text{conj}}} (-\omega_{\text{conj}}(z_*)) = \max_{z_* \in P^\ominus} (-\bar{\omega}(z_*))$, where both maximums are achieved at $-z_{*0}$, and $\forall z_* \in P^\ominus$,

$$-\bar{\omega}(-z_*) = \inf_{u \in \mathcal{X}} (f^{(1)}(x_0)u + \langle z_*, G(x_0) + G^{(1)}(x_0)u \rangle)$$

Then, we have

$$0 = \bar{\mu}_0 = \max_{z_* \stackrel{\geq}{=} \vartheta_{Z^*}} \inf_{u \in \mathcal{X}} (f^{(1)}(x_0)u + \langle z_*, G(x_0) + G^{(1)}(x_0)u \rangle)$$

where the maximum is achieved at z_{*0} . This implies that $\langle z_{*0}, G(x_0) \rangle + \inf_{u \in \mathcal{X}} \langle f^{(1)}(x_0) + (G^{(1)}(x_0))' z_{*0}, u \rangle = 0$ by Proposition 8.37. For the infimum to be finite, we must have $f^{(1)}(x_0) + (G^{(1)}(x_0))' z_{*0} = \vartheta_{\mathcal{X}^*}$. Hence, the Lagrangian L is stationary at x_0 . Then, $0 = \langle z_{*0}, G(x_0) \rangle$.

This completes the proof of the proposition. $\square$

Proposition 10.26 *Let $\mathcal{X}$ be a real Banach space, $\mathcal{Y}$ be a real Banach space with a positive cone $P \subseteq \mathcal{Y}$, and $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be surjective. Then,*

$$(A_{\text{inv}}(P))^{\oplus} = A'(P^{\oplus})$$

Proof $\forall x_* \in A'(P^{\oplus}), \exists y_* \in P^{\oplus}$ such that $x_* = A'y_*$. $\forall x \in A_{\text{inv}}(P)$, we have $Ax \in P$. Then, $\langle\langle x_*, x \rangle\rangle = \langle\langle A'y_*, x \rangle\rangle = \langle\langle y_*, Ax \rangle\rangle \geq 0$. By the arbitrariness of x , $x_* \in (A_{\text{inv}}(P))^{\oplus}$. By the arbitrariness of x_* , we have $A'(P^{\oplus}) \subseteq (A_{\text{inv}}(P))^{\oplus}$.

On the other hand, fix any $x_* \in (A_{\text{inv}}(P))^{\oplus}$. $\forall x \in \mathcal{N}(A)$, $x \in A_{\text{inv}}(P)$, since $\vartheta_{\mathcal{Y}} \in P$. Then, $\langle\langle x_*, x \rangle\rangle \geq 0$. Since $-x \in \mathcal{N}(A)$ as well, then $\langle\langle x_*, x \rangle\rangle = 0$. Hence, $x_* \in (\mathcal{N}(A))^{\perp}$. By Proposition 7.114, $(\mathcal{N}(A))^{\perp} = \mathcal{R}(A')$. Then, $\exists y_* \in \mathcal{Y}^*$ such that $x_* = A'y_*$. $\forall y \in P$, since A is surjective, then $\exists x \in A_{\text{inv}}(P)$ such that $y = Ax$. Then, we have $\langle\langle y_*, y \rangle\rangle = \langle\langle y_*, Ax \rangle\rangle = \langle\langle A'y_*, x \rangle\rangle = \langle\langle x_*, x \rangle\rangle \geq 0$. By the arbitrariness of y , we have $y_* \in P^{\oplus}$. Then, $x_* \in A'(P^{\oplus})$. By the arbitrariness of x_* , we have $(A_{\text{inv}}(P))^{\oplus} \subseteq A'(P^{\oplus})$.

Therefore, we have $A'(P^{\oplus}) = (A_{\text{inv}}(P))^{\oplus}$. This completes the proof of the proposition. $\square$

Next, we present the second-order sufficient condition for a relative strict minimum point in the optimization problem (10.3) with inequality constraints.

Proposition 10.27 *Let $\mathcal{X}$ be a real Banach space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^{\circ}$, $\mathcal{Z}$ be a real Banach space with a positive cone $P \subseteq \mathcal{Z}$, and $f : \Omega \rightarrow \mathbb{R}$ and $G : \Omega \rightarrow \mathcal{Z}$ be $\mathcal{C}_2$ at x_0 . Consider the optimization problem (10.3). Assume that*

- (i) $G(x_0) \leq \vartheta_{\mathcal{Z}}$ and x_0 is regular point of G ;
- (ii) $z_{*0} \in P^{\oplus}$ is the Lagrange multiplier, the Lagrangian $L : \Omega \rightarrow \mathbb{R}$ defined by $L(x) = f(x) + \langle\langle z_{*0}, G(x) \rangle\rangle$, $\forall x \in \Omega$, is stationary at x_0 ;
- (iii) $\langle\langle z_{*0}, G(x_0) \rangle\rangle = 0$;
- (iv) $\exists m \in (0, \infty) \subset \mathbb{R}$ such that $L^{(2)}(x_0)(u)(u) \geq m \|u\|^2$, $\forall u \in M_c := \left\{ u \in \mathcal{X} \mid G(x_0) + G^{(1)}(x_0)u \leq \vartheta_{\mathcal{Z}} \right\}$, that is, $L^{(2)}(x_0)$ is positive definite on the set M_c .

Then, x_0 is a point of relative strict minimum of f on the set $\Omega_c := \left\{ x \in \mathcal{X} \mid G(x) \leq \vartheta_{\mathcal{Z}} \right\}$.

Proof By Propositions 9.34, 9.37, 9.40, 9.41, 9.44, and 9.45, L is $\mathcal{C}_2$ at x_0 . By (ii), $L^{(1)}(x_0) = \vartheta_{\mathcal{B}(\mathcal{X}, \mathbb{R})} = \vartheta_{\mathcal{X}^*}$. Then, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $f^{(2)}(x)$, $G^{(2)}(x)$, and $L^{(2)}(x)$ exist, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, \delta_0) \subseteq \Omega$ and $L^{(2)}$ is continuous at x_0 . Then, $\exists \delta_1 \in (0, \delta_0] \subset \mathbb{R}$ and $\exists c_1 \in [0, \infty) \subset \mathbb{R}$

such that $\|L^{(2)}(x) - L^{(2)}(x_0)\| < m/5$ and $\|G^{(2)}(x)\| \leq c_1$, $\forall x \in D_1 := \mathcal{B}_X(x_0, \delta_1)$. By Propositions 7.114 and 7.98 and (i), $(\mathcal{N}(G^{(1)}(x_0)))^\perp = \mathcal{R}\left((G^{(1)}(x_0))'\right)$ is closed. Then, by Proposition 7.113, $\exists c_2 \in [0, \infty) \subset \mathbb{R}$ such that, $\forall x_* \in \mathcal{R}\left((G^{(1)}(x_0))'\right)$, there exists $z_* \in \mathcal{Z}^*$ such that $x_* = (G^{(1)}(x_0))'z_*$ and $\|z_*\| \leq c_2\|x_*\|$.

Claim 10.27.1 $\forall x \in \Omega_c \cap D_1$, $\exists h_0 \in M_c$ such that $\|x - x_0 - h_0\| \leq c_1c_2\|x - x_0\|^2$ and $\|x - x_0\| - c_1c_2\|x - x_0\|^2 \leq \|h_0\| \leq \|x - x_0\| + c_1c_2\|x - x_0\|^2$.

Proof of claim: Fix any $x \in \Omega_c \cap D_1$. Then, $G(x) \leq \vartheta_Z$. It is easy to see that $M_c \ni \vartheta_X$ is a nonempty closed convex set. Then, by Proposition 8.15, we have $0 \leq \bar{\delta} := \inf_{h \in M_c} \|x - x_0 - h\| = \max_{x_* \in M_{c\text{supp}}, \|x_*\| \leq 1} (\langle x_*, x - x_0 \rangle - g(x_*)) \geq 0$, where $g : M_{c\text{supp}} \rightarrow \mathbb{R}$ is the support functional of M_c , $M_{c\text{supp}} := \{x_* \in \mathcal{X}^* \mid \sup_{u \in M_c} \langle x_*, u \rangle < +\infty\}$, and $g(x_*) = \sup_{u \in M_c} \langle x_*, u \rangle$, $\forall x_* \in M_{c\text{supp}}$.

We will show that $M_{c\text{supp}} \subseteq ((G^{(1)}(x_0))_{\text{inv}}(P))^\oplus$. Fix any $x_* \in M_{c\text{supp}}$. $\forall u \in (G^{(1)}(x_0))_{\text{inv}}(P)$, we have $G^{(1)}(x_0)u \geq \vartheta_Z$. $\forall \alpha \in [0, \infty) \subset \mathbb{R}$, it is easy to show that $-\alpha u \in M_c$. Suppose $\langle x_*, u \rangle < 0$. Then, $\sup_{\bar{u} \in M_c} \langle x_*, \bar{u} \rangle \geq \sup_{\alpha \in [0, \infty) \subset \mathbb{R}} \langle x_*, -\alpha u \rangle = +\infty$. This implies that $x_* \notin M_{c\text{supp}}$, which is a contradiction. Hence, we must have $\langle x_*, u \rangle \geq 0$. By the arbitrariness of u , we have $x_* \in ((G^{(1)}(x_0))_{\text{inv}}(P))^\oplus$. By the arbitrariness of x_* , we have $M_{c\text{supp}} \subseteq ((G^{(1)}(x_0))_{\text{inv}}(P))^\oplus$.

Then, $\bar{\delta} = \langle x_{*0}, x - x_0 \rangle - g(x_{*0})$ for some $x_{*0} \in M_{c\text{supp}} \subseteq ((G^{(1)}(x_0))_{\text{inv}}(P))^\oplus$ with $\|x_{*0}\| \leq 1$. By (i) and Proposition 10.26, we have $x_{*0} \in (G^{(1)}(x_0))'(P^\oplus) \subseteq \mathcal{R}\left((G^{(1)}(x_0))'\right)$. Then, $\exists z_{*1} \in \mathcal{Z}^*$ such that $x_{*0} = (G^{(1)}(x_0))'z_{*1}$ and $\|z_{*1}\| \leq c_2\|x_{*0}\| \leq c_2$. This further implies that

$$\begin{aligned} \bar{\delta} &= \langle z_{*1}, G^{(1)}(x_0)(x - x_0) \rangle - \sup_{u \in M_c} \langle z_{*1}, G^{(1)}(x_0)u \rangle \\ &= \langle z_{*1}, -G(x) + G(x_0) + G^{(1)}(x_0)(x - x_0) \rangle \\ &\quad - \sup_{u \in M_c} \langle z_{*1}, -G(x) + G(x_0) + G^{(1)}(x_0)u \rangle \end{aligned}$$

where the second equality follows from Proposition 8.37. By (i), we have $\mathcal{R}(G^{(1)}(x_0)) = \mathcal{Z}$ and $\exists u_1 \in \mathcal{X}$ such that $G^{(1)}(x_0)u_1 = G(x) - G(x_0)$. It is easy to see that $u_1 \in M_c$ since $x \in \Omega_c$ and $G(x) \leq \vartheta_Z$. Then, $\bar{\delta} \leq \langle z_{*1}, -G(x) + G(x_0) + G^{(1)}(x_0)(x - x_0) \rangle$.

By Taylor's Theorem 9.48, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ such that $\|G(x) - G(x_0) - G^{(1)}(x_0)(x - x_0)\| \leq \frac{1}{2}\|G^{(2)}(t_0x + (1-t_0)x_0)\|\|x - x_0\|^2 \leq c_1\|x - x_0\|^2/2$. Then, we have $\bar{\delta} \leq \|z_{*1}\|\|G(x) - G(x_0) - G^{(1)}(x_0)(x - x_0)\| \leq c_1c_2\|x - x_0\|^2/2$. Since M_c is closed, then $\exists h_0 \in M_c$ such that $\|x - x_0 - h_0\| \leq 2\bar{\delta} \leq c_1c_2\|x - x_0\|^2$.

Then, $\|h_0\| \geq \|x - x_0\| - \|x - x_0 - h_0\| \geq \|x - x_0\| - c_1 c_2 \|x - x_0\|^2$; and $\|h_0\| \leq \|x - x_0\| + \|x - x_0 - h_0\| \leq \|x - x_0\| + c_1 c_2 \|x - x_0\|^2$. This completes the proof of the claim. $\square$

Let $c_3 = \|L^{(2)}(x_0)\|$. Let $\delta \in (0, \delta_1] \subset \mathbb{R}$ such that $c_1 c_2 \delta \leq 1/4$, $c_1^2 c_2^2 (c_3 + m/5) \delta^2 \leq m/5$, and $5c_1 c_2 (c_3 + m/5) \delta/2 \leq m/5$. $\forall x \in \Omega_c \cap \mathcal{B}_X(x_0, \delta)$, by Claim 10.27.1, $\exists h_0 \in M_c$ such that $\|x - x_0 - h_0\| \leq c_1 c_2 \|x - x_0\|^2 \leq \|x - x_0\|/4$ and $3\|x - x_0\|/4 \leq \|h_0\| \leq 5\|x - x_0\|/4$. By Taylor's Theorem 9.48, $\exists t_1 \in (0, 1) \subset \mathbb{R}$ such that

$$\begin{aligned}
f(x) - f(x_0) &\geq L(x) - L(x_0) - L^{(1)}(x_0)(x - x_0) \\
&= \frac{1}{2} L^{(2)}(\underbrace{t_1 x + (1 - t_1)x_0}_{\bar{x}})(x - x_0)(x - x_0) \\
&= \frac{1}{2} (L^{(2)}(\bar{x})(h_0)(h_0) + L^{(2)}(\bar{x})(x - x_0 - h_0)(h_0) \\
&\quad + L^{(2)}(\bar{x})(h_0)(x - x_0 - h_0) + L^{(2)}(\bar{x})(x - x_0 - h_0)(x - x_0 - h_0)) \\
&\geq \frac{1}{2} (L^{(2)}(x_0)(h_0)(h_0) + (L^{(2)}(\bar{x}) - L^{(2)}(x_0))(h_0)(h_0) \\
&\quad - \|L^{(2)}(\bar{x})\| (2\|x - x_0 - h_0\| \|h_0\| + \|x - x_0 - h_0\|^2)) \\
&\geq \frac{1}{2} (m\|h_0\|^2 - m\|h_0\|^2/5 - (\|L^{(2)}(\bar{x}) - L^{(2)}(x_0)\| + \|L^{(2)}(x_0)\|) \\
&\quad \cdot (2\|x - x_0 - h_0\| \|h_0\| + \|x - x_0 - h_0\|^2)) \\
&\geq \frac{1}{2} \left(\frac{4m}{5} \|h_0\|^2 - (c_3 + m/5) \left(\frac{5}{2} c_1 c_2 \|x - x_0\|^3 + c_1^2 c_2^2 \|x - x_0\|^4 \right) \right) \\
&\geq \frac{1}{2} \left(\frac{4m}{5} \frac{9}{16} - \frac{m}{5} - \frac{m}{5} \right) \|x - x_0\|^2 = \frac{m}{40} \|x - x_0\|^2
\end{aligned}$$

Hence, x_0 is a point of relative strict minimum for f on the set Ω_c . This completes the proof of the proposition. $\square$

Proposition 10.28 *Let $\mathcal{X}$ be a real Banach space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^\circ$, $\mathcal{Z}$ be a real Banach space with a positive cone $P \subseteq \mathcal{Z}$, and $f : \Omega \rightarrow \mathbb{R}$ and $G : \Omega \rightarrow \mathcal{Z}$ be $\mathcal{C}_2$ at x_0 . Consider the optimization problem (10.3). Assume that x_0 is a regular point of G and a point of relative minimum for f on the set $\Omega_c := \{x \in \mathcal{X} \mid G(x) \leq \vartheta_{\mathcal{Z}}\}$. Then, there exists a Lagrange multiplier $z_{*0} \in P^\oplus$ such that*

- (i) *the Lagrangian $L : \Omega \rightarrow \mathbb{R}$ defined by $L(x) = f(x) + \langle z_{*0}, G(x) \rangle$, $\forall x \in \Omega$, is stationary at x_0 and $\langle z_{*0}, G(x_0) \rangle = 0$;*
- (ii) *$L^{(2)}(x_0)$ is positive semi-definite on the subspace $\mathcal{N}(G^{(1)}(x_0)) =: M$, that is, $G^{(2)}(x_0)(h)(h) \geq 0$, $\forall h \in \mathcal{N}(G^{(1)}(x_0))$.*

Proof By the Generalized Kuhn-Tucker Theorem 10.25, there exists a Lagrange multiplier $z_{*0} \in \mathcal{Z}^*$ with $z_{*0} \geq \vartheta_{\mathcal{Z}^*}$ such that (i) holds. We

will show that (ii) also holds by an argument of contradiction. Suppose (ii) does not hold. Then, $\exists h_0 \in M$ with $\|h_0\| = 1$ such that $L^{(2)}(x_0)(h_0)(h_0) < -m < 0$ for some $m \in (0, \infty) \subset \mathbb{R}$. Let $z_0 := G(x_0) \in \mathcal{Z}$. By Surjective Mapping Theorem 9.53, $\exists r_1 \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$, and $\exists c_1 \in [0, \infty) \subset \mathbb{R}$ with $c_1 \delta_1 \leq r_1$ such that $\forall \bar{z} \in \mathcal{B}_{\mathcal{Z}}(z_0, \delta_1/2)$, $\forall \bar{x} \in \mathcal{B}_{\mathcal{X}}(x_0, r_1/2)$ with $\bar{z} = G(\bar{x})$, $\forall z \in \mathcal{B}_{\mathcal{Z}}(\bar{z}, \delta_1/2)$, $\exists x \in \mathcal{B}_{\mathcal{X}}(x_0, r_1) \subseteq \Omega$ with $\|x - \bar{x}\| \leq c_1 \|z - \bar{z}\|$, we have $z = G(x)$. By Propositions 9.34, 9.37, 9.40, 9.41, 9.44, and 9.45, L is $\mathcal{C}_2$ at x_0 . By (i), $L^{(1)}(x_0) = \vartheta_{B(X, \mathbb{R})} = \vartheta_{X^*}$. Then, $\exists r_2 \in (0, r_1] \subset \mathbb{R}$ such that $L^{(2)}(x)$ exists and $\|L^{(2)}(x) - L^{(2)}(x_0)\| < m/5$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, r_2)$. Since G is $\mathcal{C}_2$ at x_0 , then $\exists c_2 \in [0, \infty) \subset \mathbb{R}$ and $\exists r_3 \in (0, r_2] \subset \mathbb{R}$ such that $\|G^{(2)}(x)\| \leq c_2$, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_0, r_3)$. Let $c_3 := \|L^{(2)}(x_0)\|$.

$\forall \delta \in (0, r_3/2) \subset \mathbb{R}$ such that $c_2 \delta^2 < \delta_1$, $c_1 c_2 \delta/2 \leq 1/4$, $c_1^2 c_2^2 (c_3 + m/5) \delta^2/4 \leq m/5$, and $c_1 c_2 (c_3 + m/5) \delta \leq m/5$. By Taylor's Theorem 9.48, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ such that

$$\begin{aligned} \|G(x_0 + \delta h_0) - z_0\| &= \|G(x_0 + \delta h_0) - G(x_0) - \delta G^{(1)}(x_0)h_0\| \\ &\leq \frac{1}{2} \|G^{(2)}(x_0 + t_0 \delta h_0)\| \delta^2 \leq c_2 \delta^2/2 < \delta_1/2 \end{aligned}$$

Let $\bar{z}_\delta := G(x_0 + \delta h_0) \in \mathcal{B}_{\mathcal{Z}}(z_0, \delta_1/2)$ and $\bar{x}_\delta := x_0 + \delta h_0 \in \mathcal{B}_{\mathcal{X}}(x_0, r_1/2)$. Note that $\bar{z}_\delta = G(\bar{x}_\delta)$ and $z_0 \in \mathcal{B}_{\mathcal{Z}}(\bar{z}_\delta, \delta_1/2)$. Then, $\exists x_\delta \in \mathcal{B}_{\mathcal{X}}(x_0, r_1)$ with $\|x_\delta - \bar{x}_\delta\| \leq c_1 \|z_0 - \bar{z}_\delta\| \leq c_1 c_2 \delta^2/2$ such that $G(x_\delta) = z_0$. Then, we have $\|x_\delta - x_0\| \geq \|\delta h_0\| - \|x_\delta - \bar{x}_\delta\| \geq \delta - c_1 c_2 \delta^2/2 \geq 3\delta/4$; and $\|x_\delta - x_0\| \leq \|\delta h_0\| + \|x_\delta - \bar{x}_\delta\| \leq \delta + c_1 c_2 \delta^2/2 \leq 5\delta/4 < r_3$. Hence, $x_\delta \in \Omega_c \cap \mathcal{B}_{\mathcal{X}}(x_0, r_3)$. By Taylor's Theorem 9.48, $\exists t_1 \in (0, 1) \subset \mathbb{R}$ such that

$$\begin{aligned} f(x_\delta) - f(x_0) &= L(x_\delta) - L(x_0) - L^{(1)}(x_0)(x_\delta - x_0) - \langle\langle z_{*0}, G(x_\delta) \rangle\rangle \\ &\quad + \langle\langle z_{*0}, G(x_0) \rangle\rangle = L(x_\delta) - L(x_0) - L^{(1)}(x_0)(x_\delta - x_0) \\ &= \frac{1}{2} L^{(2)}(\underbrace{t_1 x_\delta + (1 - t_1) x_0}_{\hat{x}})(x_\delta - x_0)(x_\delta - x_0) \\ &= \frac{1}{2} (L^{(2)}(\hat{x})(\bar{x}_\delta - x_0)(\bar{x}_\delta - x_0) + L^{(2)}(\hat{x})(x_\delta - \bar{x}_\delta)(\bar{x}_\delta - x_0) \\ &\quad + L^{(2)}(\hat{x})(\bar{x}_\delta - x_0)(x_\delta - \bar{x}_\delta) + L^{(2)}(\hat{x})(x_\delta - \bar{x}_\delta)(x_\delta - \bar{x}_\delta)) \\ &\leq \frac{1}{2} (\delta^2 L^{(2)}(x_0)(h_0)(h_0) + \delta^2 (L^{(2)}(\hat{x}) - L^{(2)}(x_0))(h_0)(h_0) \\ &\quad + \|L^{(2)}(\hat{x})\| (2 \|x_\delta - \bar{x}_\delta\| \delta \|h_0\| + \|x_\delta - \bar{x}_\delta\|^2)) \\ &< \frac{1}{2} (-m\delta^2 + m\delta^2/5 + (\|L^{(2)}(\hat{x}) - L^{(2)}(x_0)\| + \|L^{(2)}(x_0)\|) \\ &\quad \cdot (c_1 c_2 \delta^3 + c_1^2 c_2^2 \delta^4/4)) \\ &\leq \frac{1}{2} (-\frac{4m}{5} \delta^2 + (c_3 + m/5) (c_1 c_2 \delta^3 + c_1^2 c_2^2 \delta^4/4)) \end{aligned}$$

$$\leq \frac{1}{2} \left(-\frac{4m}{5}\delta^2 + \frac{m}{5}\delta^2 + \frac{m}{5}\delta^2 \right) = -\frac{m}{5}\delta^2 < 0$$

Thus, $\lim_{\delta \rightarrow 0^+} x_\delta = x_0$, $x_\delta \in \Omega_c$, and $f(x_\delta) < f(x_0)$, for sufficiently small δ . Then, x_0 is not a point of relative minimum for f on the set Ω_c . This contradicts with the assumption. Hence, (ii) must hold. This completes the proof of the proposition. $\square$

Proposition 10.29 *Let $\mathcal{X}$ be a real reflexive Banach space and $L \in \mathcal{S}_{\mathcal{X}}$. Then, the following statements holds.*

(i) $L' = L$.

(ii) *If $L \in \mathcal{S}_{+\mathcal{X}}$, that is, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $L(x)(x) \geq \delta \|x\|_{\mathcal{X}}^2$, $\forall x \in \mathcal{X}$, then, $L \in \mathcal{B}(\mathcal{X}, \mathcal{X}^*)$ is bijective, $L^{-1} \in \mathcal{S}_{+\mathcal{X}^*}$, that is, $L^{-1}(x_*)(x_*) \geq \frac{\delta}{\epsilon + \|L\|_{\mathcal{B}(\mathcal{X}, \mathcal{X}^*)}^2} \|x_*\|_{\mathcal{X}^*}^2$, $\forall x_* \in \mathcal{X}^*$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$.*

Proof (i) $\forall x, y \in \mathcal{X}$, we have

$$L(x)(y) = L(y)(x) = \langle \langle L(x), y \rangle \rangle = \langle \langle \phi_{\mathcal{X}}(y), L(x) \rangle \rangle = \langle \langle L'(\phi_{\mathcal{X}}(y)), x \rangle \rangle$$

where the first equality follows from the fact that $L \in \mathcal{S}_{\mathcal{X}} = \mathcal{B}_{\mathcal{S}2}(\mathcal{X}, \mathbb{R})$; and $\phi_{\mathcal{X}} : \mathcal{X} \rightarrow \mathcal{X}^{**}$ is the natural mapping introduced in Remark 7.88. Since $\mathcal{X}$ is reflexive, then $\mathcal{X}^{**}$ and $\mathcal{X}$ is isometrically isomorphic and $\phi_{\mathcal{X}} = \text{id}_{\mathcal{X}}$. This implies that

$$L(x)(y) = \langle \langle L'(y), x \rangle \rangle = L(y)(x) = \langle \langle L(y), x \rangle \rangle$$

By the arbitrariness of x and y , we have $L = L'$.

(ii) By the assumption, $L \in \mathcal{B}_{\mathcal{S}2}(\mathcal{X}, \mathbb{R}) \subseteq \mathcal{B}(\mathcal{X}, \mathcal{X}^*)$. $\forall x \in \mathcal{X}$ with $x \neq \vartheta_{\mathcal{X}}$, $L(x) \neq \vartheta_{\mathcal{X}^*}$ since $L(x)(x) > 0$. Hence, $\mathcal{N}(L) = \{\vartheta_{\mathcal{X}}\}$ and L is injective. By (i), $L = L'$. Then, $\mathcal{N}(L') = \{\vartheta_{\mathcal{X}}\}$. By Proposition 7.112, $(\mathcal{R}(L))^{\perp} = \mathcal{N}(L')$. This shows that ${}^{\perp}((\mathcal{R}(L))^{\perp}) = \mathcal{X}^*$. By Proposition 7.98, $\overline{\mathcal{R}(L)} = \mathcal{X}^*$. Hence, $\mathcal{R}(L)$ is dense in $\mathcal{X}^*$. Clearly, $\vartheta_{\mathcal{X}^*} \in \mathcal{R}(L)$. $\forall x_* \in \mathcal{X}^*$ with $x_* \neq \vartheta_{\mathcal{X}^*}$, there exists $(x_{*i})_{i=1}^{\infty} \subseteq \mathcal{R}(L)$ such that $x_* = \lim_{i \in \mathbb{N}} x_{*i}$. Without loss of generality, we may assume that $x_{*i} \neq \vartheta_{\mathcal{X}^*}$, $\forall i \in \mathbb{N}$. Let $x_{*i} = L(x_i)$, $\forall i \in \mathbb{N}$, where $x_i \in \mathcal{X}$. Clearly, $x_i \neq \vartheta_{\mathcal{X}}$ since $x_{*i} \neq \vartheta_{\mathcal{X}^*}$, $\forall i \in \mathbb{N}$. Then, we have

$$\|x_{*i}\|_{\mathcal{X}^*} \|x_i\|_{\mathcal{X}} \geq \langle \langle x_{*i}, x_i \rangle \rangle = L(x_i)(x_i) \geq \delta \|x_i\|_{\mathcal{X}}^2; \quad \forall i \in \mathbb{N}$$

This implies that $\|x_{*i}\|_{\mathcal{X}^*} \geq \delta \|x_i\|_{\mathcal{X}}$ and $\|x_i\|_{\mathcal{X}} \leq \|x_{*i}\|_{\mathcal{X}^*} / \delta$, $\forall i \in \mathbb{N}$. Since $(x_{*i})_{i=1}^{\infty}$ is convergent, then there exists $c \in (0, \infty) \subset \mathbb{R}$ such that $\|x_{*i}\|_{\mathcal{X}^*} \leq c$, $\forall i \in \mathbb{N}$. Then, $\|x_i\|_{\mathcal{X}} \leq c/\delta =: c_1 \in (0, \infty) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$. This shows that $(x_i)_{i=1}^{\infty} \subseteq \overline{\mathcal{B}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, c_1)} =: S_1 \subseteq \mathcal{X} = \mathcal{X}^{**}$. By Alaoglu Theorem 7.122, $S_1 \subseteq \mathcal{X}^{**}$ is weak* compact. By Propositions 5.22 and 5.26, $(x_i)_{i=1}^{\infty}$ has a cluster point $x \in \mathcal{X}^{**} = \mathcal{X}$ in weak*

topology. Since $\mathcal{X}$ is reflexive, the weak* topology on $\mathcal{X}^{**} = \mathcal{X}$ is identical to the weak topology on $\mathcal{X}$, and the weak* topology on $\mathcal{X}^*$ is identical to the weak topology on $\mathcal{X}^*$. By Proposition 7.123, $L \in \mathcal{B}(\mathcal{X}, \mathcal{X}^*)$ is a continuous as a mapping $L : \mathcal{X}_{\text{weak}} \rightarrow \mathcal{X}_{\text{weak}^*}^*$. By Proposition 3.66, $(x_{*i})_{i=1}^\infty = (L(x_i))_{i=1}^\infty$ has a cluster point $L(x) \in \mathcal{X}_{\text{weak}^*}^*$ in weak* topology. By Proposition 7.120, $\mathcal{X}_{\text{weak}^*}^*$ is completely regular, then it is Hausdorff by Proposition 3.61. By Proposition 3.65, $L(x) = x_*$, since $x_* = \lim_{i \in \mathbb{N}} x_{*i}$ and $x_* = \lim_{i \in \mathbb{N}} x_{*i}$ weak*. This shows that $x_* \in \mathcal{R}(L)$. By the arbitrariness of x_* , we have $\mathcal{R}(L) = \mathcal{X}^*$ and L is surjective.

Hence, L is bijective. By Open Mapping Theorem 7.103, $L^{-1} \in \mathcal{B}(\mathcal{X}^*, \mathcal{X}) = \mathcal{B}(\mathcal{X}^*, \mathcal{X}^{**}) = \mathcal{B}(\mathcal{X}^*, \mathcal{B}(\mathcal{X}^*, \mathbb{R}))$.

$\forall x_*, y_* \in \mathcal{X}^*$, let $x = L^{-1}(x_*)$ and $y = L^{-1}(y_*)$. We have

$$\begin{aligned} L^{-1}(x_*)(y_*) &= \langle \langle \phi_{\mathcal{X}}(L^{-1}(x_*)), y_* \rangle \rangle = \langle \langle y_*, L^{-1}(x_*) \rangle \rangle = \langle \langle L(y), x \rangle \rangle \\ &= L(y)(x) = L(x)(y) = \langle \langle L(x), y \rangle \rangle = \langle \langle x_*, L^{-1}(y_*) \rangle \rangle = L^{-1}(y_*)(x_*) \end{aligned}$$

where $\phi_{\mathcal{X}} : \mathcal{X} \rightarrow \mathcal{X}^{**}$ is the natural mapping introduced in Remark 7.88; and the fifth equality follows from the fact that $L \in \mathcal{S}_{\mathcal{X}}$. By the arbitrariness of x_* and y_* , we have $L^{-1} \in \mathcal{S}_{\mathcal{X}^*}$.

$\forall x_* \in \mathcal{X}^*$, let $x = L^{-1}(x_*)$. Then, $x_* = L(x)$ and $\|x_*\|_{\mathcal{X}^*} \leq \|L\|_{\mathcal{B}(\mathcal{X}, \mathcal{X}^*)} \|x\|_{\mathcal{X}}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\|L\|_{\mathcal{B}(\mathcal{X}, \mathcal{X}^*)} = 0$; Case 2: $\|L\|_{\mathcal{B}(\mathcal{X}, \mathcal{X}^*)} > 0$.

Case 1: $\|L\|_{\mathcal{B}(\mathcal{X}, \mathcal{X}^*)} = 0$, we have $\|x_*\|_{\mathcal{X}^*} = 0$, $\forall x_* \in \mathcal{X}^*$. Hence, $\mathcal{X}^* = \{\vartheta_{\mathcal{X}^*}\}$. Then, $L^{-1}(x_*)(x_*) = 0 \geq \frac{\delta}{\epsilon + \|L\|_{\mathcal{B}(\mathcal{X}, \mathcal{X}^*)}^2} \|x_*\|_{\mathcal{X}^*}^2 = 0$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$.

Case 2: $\|L\|_{\mathcal{B}(\mathcal{X}, \mathcal{X}^*)} > 0$. Then, we have

$$\begin{aligned} L^{-1}(x_*)(x_*) &= L(x)(x) \geq \delta \|x\|_{\mathcal{X}}^2 \geq \frac{\delta}{\|L\|_{\mathcal{B}(\mathcal{X}, \mathcal{X}^*)}^2} \|x_*\|_{\mathcal{X}^*}^2 \\ &\geq \frac{\delta}{\epsilon + \|L\|_{\mathcal{B}(\mathcal{X}, \mathcal{X}^*)}^2} \|x_*\|_{\mathcal{X}^*}^2, \quad \forall \epsilon \in (0, \infty) \subset \mathbb{R} \end{aligned}$$

Hence, we have shown, in both cases, that $L^{-1}(x_*)(x_*) \geq \frac{\delta}{\epsilon + \|L\|_{\mathcal{B}(\mathcal{X}, \mathcal{X}^*)}^2} \cdot \|x_*\|_{\mathcal{X}^*}^2$, $\forall x_* \in \mathcal{X}^*$ and $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$. This proves that $L^{-1} \in \mathcal{S}_+ \mathcal{X}^*$.

This completes the proof of the proposition. $\square$

Chapter 11

General Measure and Integration

11.1 Measure Spaces

Definition 11.1 A measurable space $(X, \mathcal{B})$ is a set X and a σ -algebra $\mathcal{B}$ (of subsets of X) on X . A subset $A \subseteq X$ is said to be $\mathcal{B}$ -measurable if $A \in \mathcal{B}$, or simply measurable when no confusion may arise from the context.

Definition 11.2 Let $(X, \mathcal{B})$ be a measurable space. A measure μ on $(X, \mathcal{B})$ is a nonnegative extended real-valued function on $\mathcal{B}$, i. e. a mapping $\mu : \mathcal{B} \rightarrow [0, \infty] \subseteq \mathbb{R}_e$, such that

$$(i) \quad \mu(\emptyset) = 0;$$

$$(ii) \quad \mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i), \text{ where the sequence } (E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \text{ are disjoint.}$$

Then, the triple $(X, \mathcal{B}, \mu)$ is called a measure space.

Fact 11.3 Let $(X, \mathcal{B}, \mu)$ be a measure space. Then, μ is finitely additive, i. e., $\mu\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n \mu(E_i)$, where $n \in \mathbb{Z}_+$ and $(E_i)_{i=1}^n \subseteq \mathcal{B}$ is pairwise disjoint.

Proof Consider the sequence of measurable sets $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, with $E_i = \emptyset$, $i = n+1, n+2, \dots$. Clearly, the sets in $(E_i)_{i=1}^{\infty}$ are pairwise disjoint. Therefore, $\mu\left(\bigcup_{i=1}^n E_i\right) = \mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^n \mu(E_i)$. This completes the proof of the fact. $\square$

Proposition 11.4 Let $(X, \mathcal{B}, \mu)$ be a measure space. $\forall A, B \in \mathcal{B}$ such that $A \subseteq B$, then, $\mu(A) \leq \mu(B)$.

Proof The set $B \cap \tilde{A} \in \mathcal{B}$ since $\mathcal{B}$ is a σ -algebra. Note that $A \cap (B \cap \tilde{A}) = \emptyset$, then, $\mu(B) = \mu\left(A \cup (B \cap \tilde{A})\right) = \mu(A) + \mu(B \cap \tilde{A}) \geq \mu(A)$, by Fact 11.3. This completes the proof of the proposition. $\square$

Proposition 11.5 *Let $(X, \mathcal{B}, \mu)$ be a measure space. A sequence $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$ is such that $E_n \supseteq E_{n+1}$, $n = 1, 2, \dots$. If $\mu(E_1) < +\infty$, then, $\lim_{n \in \mathbb{N}} \mu(E_n) = \mu(\bigcap_{n=1}^\infty E_n)$.*

Proof Let $E = \bigcap_{n=1}^\infty E_n \in \mathcal{B}$. Then, $\tilde{E} = \bigcup_{n=1}^\infty \widetilde{E_n} \in \mathcal{B}$. Note that $E_1 = E \cup \left(\bigcup_{n=1}^\infty (E_n \cap \widetilde{E_{n+1}})\right)$. Clearly, the right-hand-side of the above equality is a disjoint union of measurable sets. Therefore,

$$\mu(E_1) = \mu(E) + \sum_{n=1}^\infty \mu(E_n \cap \widetilde{E_{n+1}}) \quad (11.1)$$

On the other hand, note that $E_n = E_{n+1} \cup (E_n \cap \widetilde{E_{n+1}})$, $n = 1, 2, \dots$, which implies that $\mu(E_n) = \mu(E_{n+1}) + \mu(E_n \cap \widetilde{E_{n+1}})$, since the two sets on the right-hand-side of the above equality are disjoint, and all three numbers are finite by Proposition 11.4.

Then, substituting the above equality into (11.1) yields $\mu(E_1) = \mu(E) + \sum_{n=1}^\infty (\mu(E_n) - \mu(E_{n+1})) = \mu(E) + \mu(E_1) - \lim_{n \in \mathbb{N}} \mu(E_n)$. Since $\mu(E_1) < +\infty$, then $\mu(E) < +\infty$ by Proposition 11.4 and $E \subseteq E_1$. Therefore, we have $\lim_{n \in \mathbb{N}} \mu(E_n) = \mu(E)$.

This completes the proof of the proposition. $\square$

Proposition 11.6 *Let $(X, \mathcal{B}, \mu)$ be a measure space. A sequence $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$. Then, $\mu(\bigcup_{n=1}^\infty E_n) \leq \sum_{n=1}^\infty \mu(E_n)$.*

Proof Let $G_1 = E_1$, $G_n = E_n \setminus \left(\bigcup_{i=1}^{n-1} E_i\right)$, $n = 2, 3, \dots$. Then, $(G_n)_{n=1}^\infty$ is a sequence of disjoint and measurable subsets of X . Therefore, $\mu(\bigcup_{i=1}^\infty G_i) = \sum_{i=1}^\infty \mu(G_i)$. Note that, $\bigcup_{i=1}^n G_i = \bigcup_{i=1}^n E_i$, $n = 1, 2, \dots$. Therefore, $\bigcup_{i=1}^\infty G_i = \bigcup_{i=1}^\infty E_i$. Thus,

$$\mu\left(\bigcup_{i=1}^\infty E_i\right) = \sum_{i=1}^\infty \mu(G_i) \quad (11.2)$$

By Proposition 11.4, $G_n \subseteq E_n$ implies that $\mu(G_n) \leq \mu(E_n)$, $n = 1, 2, \dots$. Applying these inequalities in the equality (11.2), yields $\mu(\bigcup_{i=1}^\infty E_i) \leq \sum_{i=1}^\infty \mu(E_i)$.

This completes the proof of the proposition. $\square$

Proposition 11.7 *Let $(X, \mathcal{B}, \mu)$ be a measure space and $(A_i)_{i=1}^\infty \subseteq \mathcal{B}$ with $A_i \subseteq A_{i+1}$, $\forall i \in \mathbb{N}$. Then, $\mu(\bigcup_{i=1}^\infty A_i) = \lim_{i \in \mathbb{N}} \mu(A_i) \in \mathbb{R}_e$.*

Proof Let $A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$. Note that $A = A_1 \cup (\bigcup_{i=1}^{\infty} (A_{i+1} \setminus A_i))$, where the sets in the union are pairwise disjoint. Then, by countable additivity of measure, $\mu(A) = \mu(A_1) + \sum_{i=1}^{\infty} \mu(A_{i+1} \setminus A_i) = \lim_{n \in \mathbb{N}} (\mu(A_1) + \sum_{i=1}^n \mu(A_{i+1} \setminus A_i)) = \lim_{n \in \mathbb{N}} \mu(A_{n+1}) = \lim_{i \in \mathbb{N}} \mu(A_i)$, where the third equality follows from Fact 11.3. This completes the proof of the proposition. $\square$

Proposition 11.8 *Let $(X, \mathcal{B}, \mu)$ be a measure space. Then, $\forall E_1, E_2 \in \mathcal{B}$ with $\mu(E_1 \triangle E_2) = 0$, we have $\mu(E_1) = \mu(E_2)$.*

Proof $0 = \mu(E_1 \triangle E_2) = \mu(E_1 \setminus E_2) + \mu(E_2 \setminus E_1)$. Then, $\mu(E_1 \setminus E_2) = \mu(E_2 \setminus E_1) = 0$. Hence, we have $\mu(E_1) = \mu(E_1 \cap E_2) + \mu(E_1 \cap \bar{E}_2) = \mu(E_1 \cap E_2) + \mu(E_2 \cap \bar{E}_1) = \mu(E_2)$. This completes the proof of the proposition. $\square$

Lemma 11.9 (Borel-Cantelli) *Let $(X, \mathcal{B}, \mu)$ be a measure space and $(E_n)_{n=1}^{\infty} \subseteq \mathcal{B}$. Assume that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Then, $\mu(G) := \mu(\bigcap_{n \in \mathbb{N}} \bigcup_{m=n}^{\infty} E_m) = 0$.*

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $\sum_{n=N}^{\infty} \mu(E_n) < \epsilon$. Then, $0 \leq \mu(G) \leq \mu(\bigcup_{m=N}^{\infty} E_m) \leq \sum_{m=N}^{\infty} \mu(E_m) < \epsilon$, where the second inequality follows from Proposition 11.4; and the third inequality follows from Proposition 11.6. By the arbitrariness of ϵ , we have $\mu(G) = 0$. This completes the proof of the lemma. $\square$

The set $\bigcap_{n \in \mathbb{N}} \bigcup_{m=n}^{\infty} E_m$ is equal to the set of elements that is in E_n , $n \in \mathbb{N}$, infinitely often. The book Williams (1991) uses the notation $\limsup_{n \in \mathbb{N}} E_n$ to denote this set. The book Williams (1991) also uses the notation $\liminf_{n \in \mathbb{N}} E_n$ to denote the set $\bigcup_{n \in \mathbb{N}} \bigcap_{m=n}^{\infty} E_m$, which is equal to the set of elements that is in all of E_m , $m \in \mathbb{N}$ with $n \leq m$, eventually. We will not formally use these notations. But, these sets are very important in the study of probability measure spaces.

Definition 11.10 *Let $(X, \mathcal{B}, \mu)$ be a measure space. μ is said to be finite if $\mu(X) < +\infty$. μ is said to be σ -finite if there exists a sequence of measurable subsets of X , $(X_n)_{n=1}^{\infty}$, such that $\bigcup_{n=1}^{\infty} X_n = X$ and $\mu(X_n) < +\infty$, $n = 1, 2, \dots$*

A set $E \subseteq X$ is said to be of finite measure, if $E \in \mathcal{B}$, and $\mu(E) < +\infty$. A set $E \subseteq X$ is said to be of σ -finite measure, if there exists a sequence of measurable subsets of X , $(E_n)_{n=1}^{\infty}$, such that $\bigcup_{n=1}^{\infty} E_n = E$ and $\mu(E_n) < +\infty$, $n = 1, 2, \dots$

Definition 11.11 *Let $(X, \mathcal{B}, \mu)$ be a measure space. It is said to be complete if $\forall A, E \subseteq X$, with $A \subseteq E$, $E \in \mathcal{B}$, and $\mu(E) = 0$, then, $A \in \mathcal{B}$.*

Proposition 11.12 *Let $(X, \mathcal{B}, \mu)$ be a measure space. Then, there is a unique complete measure space $(X, \mathcal{B}_0, \mu_0)$ such that*

- (i) $\mathcal{B} \subseteq \mathcal{B}_0$;
- (ii) $\forall E \in \mathcal{B}, \mu(E) = \mu_0(E)$;
- (iii) $\forall E \subseteq X, E \in \mathcal{B}_0$ if, and only if, $E = A \cup B$, where $A, B \subseteq X$ with $B \in \mathcal{B}$, and there exists a $C \in \mathcal{B}$, such that $A \subseteq C$ and $\mu(C) = 0$.

Then, $(X, \mathcal{B}_0, \mu_0)$ is said to be the completion of $(X, \mathcal{B}, \mu)$.

Furthermore, let $(X, \mathcal{B}_1, \mu_1)$ be another complete measure space that satisfies (i) and (ii), then, $\mathcal{B}_0 \subseteq \mathcal{B}_1$ and $\mu_0(E) = \mu_1(E), \forall E \in \mathcal{B}_0$.

Proof Define $\mathcal{B}_0$ to be the collection of subsets of X as the following.

$$\mathcal{B}_0 = \{ E \subseteq X \mid \text{there exists } A, B, C \subseteq X \text{ such that } E = A \cup B, \\ B, C \in \mathcal{B}, \mu(C) = 0, \text{ and } A \subseteq C \}$$

$\forall E \in \mathcal{B}$, let $A = \emptyset, B = E$, and $C = \emptyset$. Then, $E = A \cup B, B, C \in \mathcal{B}, \mu(C) = 0$, and $A \subseteq C$. Therefore, $E \in \mathcal{B}_0$. Hence, $\mathcal{B} \subseteq \mathcal{B}_0$.

The collection $\mathcal{B}_0$ form a σ -algebra on X , which is proved in the following.

$\forall E_1, E_2 \in \mathcal{B}_0$, there exist $A_1, B_1, C_1, A_2, B_2, C_2 \subseteq X$ such that $E_1 = A_1 \cup B_1, E_2 = A_2 \cup B_2, B_1, C_1, B_2, C_2 \in \mathcal{B}, \mu(C_1) = \mu(C_2) = 0$, and $A_1 \subseteq C_1, A_2 \subseteq C_2$. Then, $E_1 \cup E_2 = (A_1 \cup A_2) \cup (B_1 \cup B_2)$. The sets $B_1 \cup B_2, C_1 \cup C_2 \in \mathcal{B}$, and $A_1 \cup A_2 \subseteq C_1 \cup C_2$. The sequence of inequalities holds, $0 \leq \mu(C_1 \cup C_2) \leq \mu(C_1) + \mu(C_2) \leq 0$, where the first one follows from Proposition 11.4, and the second one follows from Proposition 11.6. This further implies that $\mu(C_1 \cup C_2) = 0$. By the definition of $\mathcal{B}_0, E_1 \cup E_2 \in \mathcal{B}_0$.

Note that $\widetilde{E}_1 = \widetilde{B}_1 \cap \widetilde{A}_1 = \widetilde{B}_1 \cap ((\widetilde{A}_1 \setminus \widetilde{C}_1) \cup \widetilde{C}_1)$, since $\widetilde{C}_1 \subseteq \widetilde{A}_1$. This implies that $\widetilde{E}_1 = (\widetilde{B}_1 \cap (\widetilde{A}_1 \setminus \widetilde{C}_1)) \cup (\widetilde{B}_1 \cap \widetilde{C}_1)$. Since $\widetilde{B}_1 \cap \widetilde{C}_1 \in \mathcal{B}$ and $\widetilde{B}_1 \cap (\widetilde{A}_1 \setminus \widetilde{C}_1) = \widetilde{B}_1 \cap \widetilde{A}_1 \cap C_1 \subseteq C_1$ and $\mu(C_1) = 0$, then, $\widetilde{E}_1 \in \mathcal{B}_0$.

Therefore, $\mathcal{B}_0$ form an algebra on the set X .

Consider an arbitrary sequence $(E_n)_{n=1}^\infty \subseteq \mathcal{B}_0$. Then, there exists $A_n, B_n, C_n \subseteq X$, such that $E_n = A_n \cup B_n, B_n, C_n \in \mathcal{B}, A_n \subseteq C_n$, and $\mu(C_n) = 0, n = 1, 2, \dots$. Then, $\bigcup_{n=1}^\infty E_n = (\bigcup_{n=1}^\infty A_n) \cup (\bigcup_{n=1}^\infty B_n)$. Since $\mathcal{B}$ is a σ -algebra, then, $\bigcup_{n=1}^\infty B_n \in \mathcal{B}$ and $\bigcup_{n=1}^\infty C_n \in \mathcal{B}$. Furthermore, $\bigcup_{n=1}^\infty A_n \subseteq \bigcup_{n=1}^\infty C_n$ and $0 \leq \mu(\bigcup_{n=1}^\infty C_n) \leq \sum_{n=1}^\infty \mu(C_n) = 0$. Then, $\bigcup_{n=1}^\infty E_n \in \mathcal{B}_0$.

This proves that $\mathcal{B}_0$ is a σ -algebra.

Define a mapping $\mu_0 : \mathcal{B}_0 \rightarrow [0, \infty] \subseteq \mathbb{R}_e$ as following. $\forall E \in \mathcal{B}_0$, there exists $A, B, C \subseteq X$ such that $E = A \cup B, B, C \in \mathcal{B}, A \subseteq C$, and $\mu(C) = 0$. $\mu_0(E) := \mu(B)$.

The mapping μ_0 is well defined because of the following. $\forall E \in \mathcal{B}_0$, let $A_1, A_2, B_1, B_2, C_1, C_2 \subseteq X$ be such that,

$$E = A_1 \cup B_1 = A_2 \cup B_2$$

$$B_1, C_1, B_2, C_2 \in \mathcal{B}$$

$$\mu(C_1) = \mu(C_2) = 0 \quad A_1 \subseteq C_1 \quad A_2 \subseteq C_2$$

Then, $\mu(B_1) = \mu\left(\left(B_1 \cap \widetilde{B_2}\right) \cup (B_1 \cap B_2)\right) = \mu(B_1 \cap \widetilde{B_2}) + \mu(B_1 \cap B_2) \leq \mu(B_1 \cap \widetilde{B_2}) + \mu(B_2)$, where the last inequality follows from Proposition 11.4.

Since $A_1 \cup B_1 = A_2 \cup B_2 = E$ implies that $\widetilde{B_2} \cap (A_1 \cup B_1) = \widetilde{B_2} \cap (A_2 \cup B_2) = \widetilde{B_2} \cap A_2 \subseteq C_2$, then $\widetilde{B_2} \cap B_1 \subseteq (\widetilde{B_2} \cap B_1) \cup (\widetilde{B_2} \cap A_1) \subseteq C_2$. By Proposition 11.4, this implies that $\mu(\widetilde{B_2} \cap B_1) = 0$. Therefore, $\mu(B_1) \leq \mu(B_2)$.

By an argument similar to the above, it can be concluded that $\mu(B_2) \leq \mu(B_1)$. Combining the two inequalities, yields $\mu(B_2) = \mu(B_1)$.

This shows that $\mu_0(E)$ is well defined.

$\forall E \in \mathcal{B} \subseteq \mathcal{B}_0$, $E = E \cup \emptyset$, then, $\mu_0(E) = \mu(E)$. This proves that μ_0 agrees with μ on $\mathcal{B}$.

In the following, it will be shown that the nonnegative, extended real-valued function μ_0 defines a measure on the measurable space $(X, \mathcal{B}_0)$. Since $\emptyset \in \mathcal{B}$, then $\mu_0(\emptyset) = \mu(\emptyset) = 0$.

Consider any sequence of subsets $(E_n)_{n=1}^\infty \subseteq \mathcal{B}_0$, which are pairwise disjoint. There exists $A_n, B_n, C_n \subseteq X$, such that $E_n = A_n \cup B_n$, $B_n, C_n \in \mathcal{B}$, $A_n \subseteq C_n$, and $\mu(C_n) = 0$, $n = 1, 2, \dots$. Since $B_n \subseteq E_n$, $n = 1, 2, \dots$, then, $(B_n)_{n=1}^\infty$ are pairwise disjoint. Hence, $\mu(\bigcup_{n=1}^\infty B_n) = \sum_{n=1}^\infty \mu(B_n)$. Since $\bigcup_{n=1}^\infty E_n = (\bigcup_{n=1}^\infty A_n) \cup (\bigcup_{n=1}^\infty B_n)$, $\bigcup_{n=1}^\infty A_n \subseteq \bigcup_{n=1}^\infty C_n$, $0 \leq \mu(\bigcup_{n=1}^\infty C_n) \leq \sum_{n=1}^\infty \mu(C_n) = 0$, and $\bigcup_{n=1}^\infty B_n, \bigcup_{n=1}^\infty C_n \in \mathcal{B}$ then, $\mu_0(\bigcup_{n=1}^\infty E_n) = \mu(\bigcup_{n=1}^\infty B_n) = \sum_{n=1}^\infty \mu(B_n) = \sum_{n=1}^\infty \mu_0(E_n)$. This proves that μ_0 is countably additive for disjoint $\mathcal{B}_0$ -measurable sets.

Hence, μ_0 is a measure on $(X, \mathcal{B}_0)$.

In the following, it is proved that $(X, \mathcal{B}_0, \mu_0)$ is a complete measure space.

$\forall E_0 \subseteq X$, such that there exists $E \in \mathcal{B}_0$ with $E_0 \subseteq E$ and $\mu_0(E) = 0$. Then, there exists $A, B, C \subseteq X$ such that $E = A \cup B$, $B, C \in \mathcal{B}$, $A \subseteq C$, and $\mu(C) = 0$, by the definition of $\mathcal{B}_0$ and $E \in \mathcal{B}_0$. By the definition of μ_0 , we have $\mu_0(E) = \mu(B) = 0$. Hence, $E_0 = E_0 \cup \emptyset$ and $E_0 \subseteq A \cup B \subseteq C \cup B$. Since $\emptyset, C \cup B \in \mathcal{B}$, and $0 \leq \mu(C \cup B) \leq \mu(C) + \mu(B) = 0$, then $E_0 \in \mathcal{B}_0$. This proves that $(X, \mathcal{B}_0, \mu_0)$ is complete.

Let $(X, \mathcal{B}_1, \mu_1)$ be any complete measure space that satisfies the (i) and (ii). It is shown in the following that $\mathcal{B}_0 \subseteq \mathcal{B}_1$ by an argument of contradiction. Suppose there exist a set $E \in \mathcal{B}_0 \setminus \mathcal{B}_1$, then, $E \notin \mathcal{B}$. Since $E \in \mathcal{B}_0$, then, there exists $A, B, C \subseteq X$, such that $E = A \cup B$, $A \subseteq C$, $B, C \in \mathcal{B}$, and $\mu(C) = 0$. This implies $B, C \in \mathcal{B}_1$ and $\mu_1(C) = \mu(C) = 0$. Hence, $A \notin \mathcal{B}_1$, otherwise, we have $E \in \mathcal{B}_1$. But, $A \subseteq C$, with $\mu_1(C) = 0$, which contradicts the fact that $(X, \mathcal{B}_1, \mu_1)$ being complete.

Therefore, we have $\mathcal{B}_0 \subseteq \mathcal{B}_1$.

$\forall E \in \mathcal{B}_0$, there exists $A, B, C \subseteq X$, such that $E = A \cup B$, $A \subseteq C$, $B, C \in \mathcal{B}$, and $\mu(C) = 0$. Then, $\mu_0(E) = \mu(B) = \mu_1(B) \leq \mu_1(E) \leq \mu_1(B \cup C) \leq \mu_1(B) + \mu_1(C) = \mu(B)$, where Propositions 11.4 and 11.6 have been applied. Therefore, μ_0 and μ_1 agrees on the σ -algebra $\mathcal{B}_0$.

Hence, $(X, \mathcal{B}_0, \mu_0)$ is a complete measure space that satisfies (i), (ii), and (iii). It is the unique such space since $\mathcal{B}_0$ is unique by (iii), then μ_0 is also unique since any complete measure space $(X, \mathcal{B}_0, \mu_1)$ satisfying (i) and (ii) will be such that $\mu_1(E) = \mu_0(E)$, $\forall E \in \mathcal{B}_0$.

This completes the proof of the proposition. $\square$

Proposition 11.13 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $A \in \mathcal{B}$, $\hat{\mathcal{B}} \subseteq \mathcal{B}$ be a σ -algebra on X , and $\mathcal{B}_A := \{C \subseteq A \mid C \in \mathcal{B}\}$. Then, the following statements hold.*

- (i) $(X, \hat{\mathcal{B}}, \mu|_{\hat{\mathcal{B}}})$ and $\mathcal{A} := (A, \mathcal{B}_A, \mu|_{\mathcal{B}_A})$ are measure spaces. The measure space $\mathcal{A}$ is said to be the subspace of the measure space $\mathcal{X}$.
- (ii) If, in addition, $\mathcal{X}$ is complete, then $\mathcal{A}$ is also complete.
- (iii) Let $\bar{\mathcal{X}} := (X, \bar{\mathcal{B}}, \bar{\mu})$ be the completion of $\mathcal{X}$ as defined in Proposition 11.12, and $\bar{\mathcal{A}} := (A, \bar{\mathcal{B}}_A, \bar{\mu}_A)$ be the measure subspace of $\bar{\mathcal{X}}$. Then, $\bar{\mathcal{A}}$ is the completion of $\mathcal{A}$.

Proof (i) Since $\hat{\mathcal{B}}$ is a σ -algebra on X , then $(X, \hat{\mathcal{B}})$ is a measurable space. $\forall B \in \hat{\mathcal{B}} \subseteq \mathcal{B}$, $0 \leq \mu(B) \leq +\infty$ is well-defined. Then, $\mu|_{\hat{\mathcal{B}}} : \hat{\mathcal{B}} \rightarrow [0, \infty] \subseteq \mathbb{R}_e$. Clearly, $\mu|_{\hat{\mathcal{B}}}(\emptyset) = 0$. $\forall$ pairwise disjoint sequence $(E_i)_{i=1}^\infty \subseteq \hat{\mathcal{B}} \subseteq \mathcal{B}$, we have $E := \bigcup_{i=1}^\infty E_i \in \hat{\mathcal{B}}$ and $\mu|_{\hat{\mathcal{B}}}(E) = \mu(E) = \sum_{i=1}^\infty \mu(E_i) = \sum_{i=1}^\infty \mu|_{\hat{\mathcal{B}}}(E_i)$. Hence, $(X, \hat{\mathcal{B}}, \mu|_{\hat{\mathcal{B}}})$ is a measure space.

Note that $\emptyset \in \mathcal{B}_A$ and $A \in \mathcal{B}_A$. $\forall E_A \in \mathcal{B}_A$, then $A \supseteq A \setminus E_A \in \mathcal{B}$ and $A \setminus E_A \in \mathcal{B}_A$. $\forall (E_{Ai})_{i=1}^\infty \subseteq \mathcal{B}_A$, $\forall i \in \mathbb{N}$, $A \supseteq E_{Ai} \in \mathcal{B}$. Then, $A \supseteq \bigcup_{i=1}^\infty E_{Ai} \in \mathcal{B}$ and $\bigcup_{i=1}^\infty E_{Ai} \in \mathcal{B}_A$. Hence, $\mathcal{B}_A$ is a σ -algebra on A . It is straightforward to show that $\mu|_{\mathcal{B}_A}$ is a measure on the measurable space $(A, \mathcal{B}_A)$. Hence $(A, \mathcal{B}_A, \mu|_{\mathcal{B}_A})$ is a measure space.

(ii) If, in addition, $(X, \mathcal{B}, \mu)$ is complete. $\forall E_A \subseteq A$ such that there exists $B \in \mathcal{B}_A$ with $E_A \subseteq B$ and $\mu|_{\mathcal{B}_A}(B) = 0$. Then, $B \in \mathcal{B}$ and $\mu(B) = 0$. Then, $E_A \in \mathcal{B}$ by the completeness of $(X, \mathcal{B}, \mu)$. Thus, $E_A \in \mathcal{B}_A$. By the arbitrariness of E_A , $(A, \mathcal{B}_A, \mu|_{\mathcal{B}_A})$ is complete.

(iii) Note that $\bar{\mathcal{B}}_A = \{C \subseteq A \mid C \in \bar{\mathcal{B}}\}$ and $\bar{\mu}_A = \bar{\mu}|_{\bar{\mathcal{B}}_A}$. By (ii), $\bar{\mathcal{A}}$ is a complete measure space. Since $\mathcal{B} \subseteq \bar{\mathcal{B}}$, then $\mathcal{B}_A \subseteq \bar{\mathcal{B}}_A$. $\forall E \in \mathcal{B}_A$, $\mu|_{\mathcal{B}_A}(E) = \mu(E) = \bar{\mu}(E) = \bar{\mu}_A(E)$. $\forall E \subseteq A$, assume that $E = E_A \cup E_B$ with $E_B, E_C \in \mathcal{B}_A$, $E_A \subseteq E_C$, and $\mu|_{\mathcal{B}_A}(E_C) = 0$. Then, $\bar{\mu}_A(E_C) = 0$ and $E_A \in \bar{\mathcal{B}}_A$ since $\bar{\mathcal{A}}$ is complete. Then, $E \in \bar{\mathcal{B}}_A$. $\forall E \subseteq A$, assume that $E \in \bar{\mathcal{B}}_A$. Then, $E \subseteq A$ and $E \in \bar{\mathcal{B}}$. By Proposition 11.12, $E = E_A \cup E_B$ with $E_B, E_C \in \mathcal{B}$, $E_A \subseteq E_C$, and $\mu(E_C) = 0$. Then, $E_B, E_C \cap A \in \mathcal{B}_A$, $E_A \subseteq E_C \cap A$, and $0 \leq \mu|_{\mathcal{B}_A}(E_C \cap A) = \mu(E_C \cap A) \leq \mu(E_C) = 0$. Hence, (i) — (iii) of Proposition 11.12 are satisfied and $\bar{\mathcal{A}}$ is the completion of $\mathcal{A}$.

This completes the proof of the proposition. $\square$

11.2 Outer Measure and the Extension Theorem

Definition 11.14 Let X be a set. An outer measure is a function $\mu_o : {}^X 2 \rightarrow [0, \infty] \subset \mathbb{R}_e$ satisfying

- (i) $\mu_o(\emptyset) = 0$;
- (ii) if $A \subseteq B \subseteq X$, then $\mu_o(A) \leq \mu_o(B)$; (monotonicity)
- (iii) if $E \subseteq \bigcup_{i=1}^{\infty} E_i \subseteq X$, then $\mu_o(E) \leq \sum_{i=1}^{\infty} \mu_o(E_i)$. (countable subadditivity)

Note that (i) and (iii) above implies finite subadditivity.

Definition 11.15 Let X be a set and $\mu_o : {}^X 2 \rightarrow [0, \infty] \subset \mathbb{R}_e$ be an outer measure. $E \subseteq X$ is said to be measurable with respect to μ_o if $\forall A \subseteq X$, we have $\mu_o(A) = \mu_o(A \cap E) + \mu_o(A \cap \widetilde{E})$.

Lemma 11.16 Let X be a set and $\mu_o : {}^X 2 \rightarrow [0, \infty] \subset \mathbb{R}_e$ be an outer measure. $\forall E \subseteq X$ with $\mu_o(E) = 0$, then E is measurable with respect to μ_o .

Proof Fix any $E \subseteq X$ with $\mu_o(E) = 0$. $\forall A \subseteq X$, $0 \leq \mu_o(A \cap E) \leq \mu_o(E) = 0$ and $\mu_o(A \cap \widetilde{E}) \leq \mu_o(A)$. Then, we have $\mu_o(A) \leq \mu_o(A \cap E) + \mu_o(A \cap \widetilde{E}) \leq \mu_o(A)$. Hence, E is measurable with respect to μ_o . This completes the proof of the lemma. $\square$

Theorem 11.17 Let X be a set, $\mu_o : {}^X 2 \rightarrow [0, \infty] \subset \mathbb{R}_e$ be an outer measure, $\mathcal{B} := \{E \subseteq X \mid E \text{ is measurable with respect to } \mu_o\}$, and $\bar{\mu} := \mu_o|_{\mathcal{B}} : \mathcal{B} \rightarrow [0, \infty] \subset \mathbb{R}_e$. Then, $\mathcal{B}$ is a σ -algebra on X and $(X, \mathcal{B}, \bar{\mu})$ is a complete measure space.

Proof Clearly $\emptyset, X \subseteq \mathcal{B}$. $\forall E \in \mathcal{B}$, we have $\widetilde{E} \in \mathcal{B}$. $\forall E_1, E_2 \in \mathcal{B}$, $\forall A \subseteq X$, by the measurability of E_1 and E_2 , we have $\mu_o(A) = \mu_o(A \cap E_1) + \mu_o(A \cap \widetilde{E_1}) = \mu_o(A \cap E_1) + \mu_o(A \cap \widetilde{E_1} \cap E_2) + \mu_o(A \cap \widetilde{E_1} \cap \widetilde{E_2})$. Note that $A \cap (E_1 \cup E_2) = A \cap (E_2 \cup (E_1 \setminus E_2)) = (A \cap E_2) \cup (A \cap E_1 \cap \widetilde{E_2})$. Then, we have $\mu_o(A \cap (E_1 \cup E_2)) \leq \mu_o(A \cap E_2) + \mu_o(A \cap E_1 \cap \widetilde{E_2})$. Then, we have $\mu_o(A) \geq \mu_o(A \cap (E_1 \cup E_2)) + \mu_o(A \cap \widetilde{E_1} \cap \widetilde{E_2}) = \mu_o(A \cap (E_1 \cup E_2)) + \mu_o(A \cap \widetilde{E_1} \cup \widetilde{E_2}) \geq \mu_o(A)$. Hence, $E_1 \cup E_2 \in \mathcal{B}$. Thus, $\mathcal{B}$ is an algebra on X .

$\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$. Let $G_n := \bigcup_{i=1}^n E_i \in \mathcal{B}$ and $\bar{E}_1 := G_1 \in \mathcal{B}$, $\bar{E}_{n+1} := G_{n+1} \setminus G_n \in \mathcal{B}$, $\forall n \in \mathbb{N}$. Then, $\bar{E}_1, \bar{E}_2, \dots$ are pairwise disjoint. Let $E := \bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} \bar{E}_i$. $\forall A \subseteq X$, we have $\mu_o(A) = \mu_o(A \cap G_n) + \mu_o(A \cap \widetilde{G_n}) \geq \mu_o(A \cap G_n) + \mu_o(A \cap \widetilde{E})$ and $\mu_o(A \cap G_{n+1}) = \mu_o(A \cap G_{n+1} \cap \bar{E}_{n+1}) + \mu_o(A \cap G_{n+1} \cap \widetilde{\bar{E}_{n+1}}) = \mu_o(A \cap \bar{E}_{n+1}) + \mu_o(A \cap G_n)$, $\forall n \in \mathbb{N}$. Hence, we have $\mu_o(A \cap G_n) = \sum_{i=1}^n \mu_o(A \cap \bar{E}_i)$ and $\mu_o(A) \geq \sum_{i=1}^n \mu_o(A \cap \bar{E}_i) + \mu_o(A \cap \widetilde{E})$,

$\forall n \in \mathbb{N}$. Therefore, $\mu_o(A) \geq \sum_{i=1}^{\infty} \mu_o(A \cap \bar{E}_i) + \mu_o(A \cap \tilde{E})$. Note that $A \cap E = \bigcup_{i=1}^{\infty} (A \cap \bar{E}_i)$, which implies that $\mu_o(A \cap E) \leq \sum_{i=1}^{\infty} \mu_o(A \cap \bar{E}_i)$, by countable subadditivity of outer measures. Then, we have $\mu_o(A) \geq \sum_{i=1}^{\infty} \mu_o(A \cap \bar{E}_i) + \mu_o(A \cap \tilde{E}) \geq \mu_o(A \cap E) + \mu_o(A \cap \tilde{E}) \geq \mu_o(A)$. Hence, by the arbitrariness of A , $E \in \mathcal{B}$. This implies that $\mathcal{B}$ is a σ -algebra on X .

Set $A = E$ in the above, we have $\bar{\mu}(E) = \mu_o(E) = \sum_{i=1}^{\infty} \mu_o(E \cap \bar{E}_i) = \sum_{i=1}^{\infty} \bar{\mu}(\bar{E}_i)$. Hence, $(X, \mathcal{B}, \bar{\mu})$ is a measure space.

Fix any $A, E \subseteq X$ with $A \subseteq E$, $E \in \mathcal{B}$, and $\bar{\mu}(E) = 0$. Then, $0 \leq \mu_o(A) \leq \mu_o(E) = \bar{\mu}(E) = 0$. By Lemma 11.16, $A \in \mathcal{B}$. Therefore, $(X, \mathcal{B}, \bar{\mu})$ is complete.

This completes the proof of the theorem. $\square$

Definition 11.18 Let X be a set and $\mathcal{A} \subseteq {}^X 2$ be an algebra on X . A measure on algebra $\mathcal{A}$ is a function $\mu : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R}_e$ that satisfies:

$$(i) \quad \mu(\emptyset) = 0;$$

$$(ii) \quad \forall (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \text{ with } A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{A} \text{ and } A_1, A_2, \dots \text{ being pairwise disjoint, then } \mu(A) = \sum_{i=1}^{\infty} \mu(A_i).$$

μ is said to be finite if $\mu(X) < +\infty$; and it is said to be σ -finite if $\exists (X_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ such that $X = \bigcup_{i=1}^{\infty} X_i$ and $\mu(X_i) < \infty$, $\forall i \in \mathbb{N}$.

Clearly, a measure on an algebra is monotonic.

Theorem 11.19 (Carathéodory Extension Theorem) Let X be a set, $\mathcal{A} \subseteq {}^X 2$ be an algebra on X , and $\mu : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R}_e$ be a measure on algebra $\mathcal{A}$. Define $\mu_o : {}^X 2 \rightarrow [0, \infty] \subset \mathbb{R}_e$ by

$$\mu_o(A) = \inf_{(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}, A \subseteq \bigcup_{i=1}^{\infty} A_i} \sum_{i=1}^{\infty} \mu(A_i), \quad \forall A \subseteq X$$

Then, the following statements hold.

(i) μ_o is an outer measure on X and is said to be the outer measure induced by μ .

(ii) Let $\mathcal{B} := \{E \subseteq X \mid E \text{ is measurable with respect to } \mu_o\}$ and $\bar{\mu} := \mu_o|_{\mathcal{B}} : \mathcal{B} \rightarrow [0, \infty] \subset \mathbb{R}_e$. Then, $(X, \mathcal{B}, \bar{\mu})$ is a complete measure space and $\bar{\mu}(A) = \mu(A)$, $\forall A \in \mathcal{A}$. Hence, $\bar{\mu}$ is an extension of μ to the σ -algebra $\mathcal{B} \supseteq \mathcal{A}$.

(iii) $\bar{\mu}$ is finite if μ is finite; and $\bar{\mu}$ is σ -finite if μ is σ -finite.

(iv) Let $\mathcal{B}_a$ be the σ -algebra on X generated by $\mathcal{A}$. If μ is σ -finite, then $\bar{\mu}|_{\mathcal{B}_a}$ is the unique measure on the measurable space $(X, \mathcal{B}_a)$ that is an extension of μ .

(v) If μ is σ -finite, then the measure space $(X, \mathcal{B}, \bar{\mu})$ is the completion of the measure space $(X, \mathcal{B}_a, \bar{\mu}|_{\mathcal{B}_a})$.

Proof We need the following intermediate results.

Claim 11.19.1 Let $A \in \mathcal{A}$, $(A_i)_{i=1}^\infty \subseteq \mathcal{A}$, and $A \subseteq \bigcup_{i=1}^\infty A_i$. Then, $\mu(A) \leq \sum_{i=1}^\infty \mu(A_i)$.

Proof of claim: Let $G_1 := A \cap A_1 \in \mathcal{A}$ and $G_{n+1} := A \cap A_{n+1} \cap \left(\bigcap_{i=1}^n \widetilde{A_i}\right) \in \mathcal{A}$, $\forall n \in \mathbb{N}$. Then, $A = \bigcup_{i=1}^\infty G_i$ and $G_1, G_2, \dots$ are pairwise disjoint. This implies that $\mu(A) = \sum_{i=1}^\infty \mu(G_i) \leq \sum_{i=1}^\infty \mu(A_i)$, where the inequality follows from the monotonicity of μ . This completes the proof of the claim. $\square$

Hence, $\forall A \in \mathcal{A}$, we have $\mu_o(A) = \mu(A)$.

Claim 11.19.2 μ_o is an outer measure.

Proof of claim: Clearly, $\mu_o(\emptyset) = \mu(\emptyset) = 0$ since $\emptyset \in \mathcal{A}$. $\forall A_1 \subseteq A_2 \subseteq X$, by the definition of μ_o , we have $\mu_o(A_1) \leq \mu_o(A_2)$. We need only to show the countable subadditivity for μ_o . Fix any $E \subseteq \bigcup_{i=1}^\infty E_i \subseteq X$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu_o(E_i) < \infty$, $\forall i \in \mathbb{N}$; Case 2: $\exists i_0 \in \mathbb{N}$ such that $\mu_o(E_{i_0}) = \infty$.

Case 1: $\mu_o(E_i) < \infty$, $\forall i \in \mathbb{N}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, $\exists (A_{i,j})_{j=1}^\infty \subseteq \mathcal{A}$ such that $E_i \subseteq \bigcup_{j=1}^\infty A_{i,j}$ and $\mu_o(E_i) \leq \sum_{j=1}^\infty \mu(A_{i,j}) < \mu_o(E_i) + \epsilon/2^i$. Then, $E \subseteq \bigcup_{i=1}^\infty E_i \subseteq \bigcup_{i=1}^\infty \bigcup_{j=1}^\infty A_{i,j}$. By the definition of μ_o , we have $\mu_o(E) \leq \sum_{i=1}^\infty \sum_{j=1}^\infty \mu(A_{i,j}) \leq \sum_{i=1}^\infty \mu_o(E_i) + \epsilon$. By the arbitrariness of ϵ , we have $\mu_o(E) \leq \sum_{i=1}^\infty \mu_o(E_i)$.

Case 2: $\exists i_0 \in \mathbb{N}$ such that $\mu_o(E_{i_0}) = \infty$. Then, $\mu_o(E) \leq \infty = \sum_{i=1}^\infty \mu_o(E_i)$.

In both cases, we have shown that $\mu_o(E) \leq \sum_{i=1}^\infty \mu_o(E_i)$. Hence, μ_o is an outer measure. This completes the proof of the claim. $\square$

Hence, the statement (i) holds.

(ii) By Theorem 11.17, $(X, \mathcal{B}, \bar{\mu})$ is a complete measure space.

Claim 11.19.3 $\mathcal{A} \subseteq \mathcal{B}$.

Proof of claim: Fix any $A \in \mathcal{A}$. $\forall \bar{A} \subseteq X$, we will show that $\mu_o(\bar{A}) = \mu_o(\bar{A} \cap A) + \mu_o(\bar{A} \cap \widetilde{A})$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\mu_o(\bar{A}) = \infty$; Case 2: $\mu_o(\bar{A}) < \infty$. Case 1: $\mu_o(\bar{A}) = \infty$. Then, $\mu_o(\bar{A}) \geq \mu_o(\bar{A} \cap A) + \mu_o(\bar{A} \cap \widetilde{A}) \geq \mu_o(\bar{A})$, where the second inequality follows from the countable subadditivity of an outer measure. This case is proved.

Case 2: $\mu_o(\bar{A}) < \infty$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists (A_i)_{i=1}^\infty \subseteq \mathcal{A}$ such that $\bar{A} \subseteq \bigcup_{i=1}^\infty A_i$ and $\mu_o(\bar{A}) \leq \sum_{i=1}^\infty \mu(A_i) < \mu_o(\bar{A}) + \epsilon$. By the countable additivity of μ , we have $\mu(A_i) = \mu(A_i \cap A) + \mu(A_i \cap \widetilde{A})$, $\forall i \in \mathbb{N}$. Then,

$\mu_o(\bar{A}) + \epsilon > \sum_{i=1}^{\infty} (\mu(A_i \cap A) + \mu(A_i \cap \tilde{A})) \geq \mu_o(\bar{A} \cap A) + \mu_o(\bar{A} \cap \tilde{A}) \geq \mu_o(\bar{A})$, where the second inequality follows from the definition of μ_o ; and the third inequality follows from the countable subadditivity of μ_o . By the arbitrariness of ϵ , we have $\mu_o(\bar{A}) = \mu_o(\bar{A} \cap A) + \mu_o(\bar{A} \cap \tilde{A})$. This case is proved.

Hence, in both cases, we have $\mu_o(\bar{A}) = \mu_o(\bar{A} \cap A) + \mu_o(\bar{A} \cap \tilde{A})$. By the arbitrariness of $\bar{A}$, we have A is measurable with respect to μ_o . Hence, $A \in \mathcal{B}$ and $\mathcal{A} \subseteq \mathcal{B}$. This completes the proof of the claim. $\square$

$\forall A \in \mathcal{A}$, we have $\bar{\mu}(A) = \mu_o(A) = \mu(A)$. Hence, the statement (ii) holds.

(iii) Note that $X \in \mathcal{A}$. If μ is finite, then $\bar{\mu}(X) = \mu(X) < \infty$. Hence, $\bar{\mu}$ is finite.

If μ is σ -finite, then $\exists (X_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ with $X = \bigcup_{i=1}^{\infty} X_i$ such that $\mu(X_i) < +\infty$, $\forall i \in \mathbb{N}$. Then, $\bar{\mu}(X_i) = \mu(X_i) < +\infty$ and $(X_i)_{i=1}^{\infty} \subseteq \mathcal{B}$. Hence, $\bar{\mu}$ is σ -finite.

(iv) Let μ be σ -finite and $\tilde{\mu} : \mathcal{B}_a \rightarrow [0, \infty] \subset \mathbb{R}_e$ be any measure that is an extension of μ , that is, $\tilde{\mu}(A) = \mu(A)$, $\forall A \in \mathcal{A}$. Clearly, $\mathcal{B}_a \subseteq \mathcal{B}$. Then, by Proposition 11.13, $\bar{\mu}|_{\mathcal{B}_a}$ is a measure on $(X, \mathcal{B}_a)$ and it is an extension of μ . Let $\mathcal{A}_\sigma := \{E \subseteq X \mid \exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \ni E = \bigcup_{i=1}^{\infty} A_i\}$. Since $\mathcal{B} \supseteq \mathcal{A}$ and is a σ -algebra on X , then $\mathcal{A}_\sigma \subseteq \mathcal{B}_a \subseteq \mathcal{B}$.

Claim 11.19.4 $\forall E \in \mathcal{A}_\sigma$, $\tilde{\mu}(E) = \bar{\mu}(E)$.

Proof of claim: Fix any $E \in \mathcal{A}_\sigma$. Then, $\exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \ni E = \bigcup_{i=1}^{\infty} A_i$. Let $\bar{A}_1 := A_1$ and $\bar{A}_{n+1} = (\bigcup_{i=1}^{n+1} A_i) \setminus (\bigcup_{i=1}^n A_i)$, $\forall n \in \mathbb{N}$. Then, $(\bar{A}_i)_{i=1}^{\infty} \subseteq \mathcal{A}$, $E = \bigcup_{i=1}^{\infty} \bar{A}_i$, and $\bar{A}_1, \bar{A}_2, \dots$ are pairwise disjoint. Then, we have $\tilde{\mu}(E) = \sum_{i=1}^{\infty} \tilde{\mu}(\bar{A}_i) = \sum_{i=1}^{\infty} \mu(\bar{A}_i) = \sum_{i=1}^{\infty} \bar{\mu}(\bar{A}_i) = \bar{\mu}(E)$, where the first and last equality follows from the countable additivity of measures. This completes the proof of the claim. $\square$

Claim 11.19.5 $\forall B \in \mathcal{B}_a$ with $\bar{\mu}(B) < \infty$, we have $\bar{\mu}(B) = \tilde{\mu}(B)$.

Proof of claim: Fix any $B \in \mathcal{B}_a$ with $\bar{\mu}(B) < \infty$. Then, $\mu_o(B) = \bar{\mu}(B) < \infty$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the definition of μ_o , $\exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ such that $B \subseteq \bigcup_{i=1}^{\infty} A_i =: E \in \mathcal{A}_\sigma \subseteq \mathcal{B}_a$ and $\bar{\mu}(B) = \mu_o(B) \leq \sum_{i=1}^{\infty} \mu(A_i) < \bar{\mu}(B) + \epsilon$. Then, $\tilde{\mu}(B) \leq \tilde{\mu}(E) = \bar{\mu}(E) \leq \sum_{i=1}^{\infty} \mu(A_i) < \bar{\mu}(B) + \epsilon$, where the first inequality follows from the monotonicity of measure, Proposition 11.4; the first equality follows from Claim 11.19.4; and the second inequality follows from the countable subadditivity of measure, Proposition 11.6. By the arbitrariness of ϵ , we have $\tilde{\mu}(B) \leq \bar{\mu}(B)$.

On the other hand, $\bar{\mu}(E) = \bar{\mu}(B) + \bar{\mu}(E \setminus B) < \bar{\mu}(B) + \epsilon$. Since $\bar{\mu}(B) < \infty$, then $\bar{\mu}(E \setminus B) < \epsilon$. Note that $E \setminus B \in \mathcal{B}_a$. This implies that $\bar{\mu}(B) \leq \bar{\mu}(E) = \tilde{\mu}(E) = \tilde{\mu}(B) + \tilde{\mu}(E \setminus B) \leq \tilde{\mu}(B) + \bar{\mu}(E \setminus B) < \tilde{\mu}(B) + \epsilon$, where the second inequality follows from the result of the previous paragraph. By the arbitrariness of ϵ , we have $\bar{\mu}(B) \leq \tilde{\mu}(B)$.

Hence, $\bar{\mu}(B) = \tilde{\mu}(B)$. This completes the proof of the claim. $\square$

Since μ is σ -finite, then $\exists (X_i)_{i=1}^\infty \subseteq \mathcal{A}$ such that $X = \bigcup_{i=1}^\infty X_i$ and $\mu(X_i) < \infty$, $\forall i \in \mathbb{N}$. Without loss of generality, we may assume that $X_1, X_2, \dots$ are pairwise disjoint. $\forall B \in \mathcal{B}_a$, we have $B = \bigcup_{i=1}^\infty (B \cap X_i)$, $B \cap X_1, B \cap X_2, \dots$ are pairwise disjoint, and $\bar{\mu}(B \cap X_i) \leq \bar{\mu}(X_i) < \infty$ with $B \cap X_i \in \mathcal{B}_a$, $\forall i \in \mathbb{N}$. This implies that $\tilde{\mu}(B) = \sum_{i=1}^\infty \tilde{\mu}(B \cap X_i) = \sum_{i=1}^\infty \bar{\mu}(B \cap X_i) = \bar{\mu}(B)$, where the first and last equality follows from the countable additivity of measures; and the second equality follows from Claim 11.19.5. Hence, $\bar{\mu}|_{\mathcal{B}_a} = \tilde{\mu}$. Therefore, $\bar{\mu}|_{\mathcal{B}_a}$ is the unique measure on $(X, \mathcal{B}_a)$ that is an extension of μ .

(v) Let μ be σ -finite. Consider the measure space $(X, \mathcal{B}_a, \bar{\mu}|_{\mathcal{B}_a})$. By Proposition 11.12, this measure space admits the completion $(X, \hat{\mathcal{B}}, \hat{\mu})$. Since $(X, \mathcal{B}, \bar{\mu})$ is a complete measure space that agrees with $\bar{\mu}|_{\mathcal{B}_a}$ on $\mathcal{B}_a$, by Proposition 11.12, we have $\hat{\mathcal{B}} \subseteq \mathcal{B}$ and $\hat{\mu} = \bar{\mu}|_{\hat{\mathcal{B}}}$.

Claim 11.19.6 $\forall B \in \mathcal{B}$, $\exists U \in \mathcal{B}_a$ with $B \subseteq U$ such that $\bar{\mu}(U \setminus B) = 0$.

Proof of claim: $\forall B \in \mathcal{B}$, $\bar{\mu}(B) = \inf_{(A_i)_{i=1}^\infty \subseteq \mathcal{A}, B \subseteq \bigcup_{i=1}^\infty A_i} \sum_{i=1}^\infty \mu(A_i)$. We first prove the special case $\bar{\mu}(B) < +\infty$. $\forall i \in \mathbb{N}$, $\exists (U_{i,j})_{j=1}^\infty \subseteq \mathcal{A}$ with $B \subseteq \bigcup_{j=1}^\infty U_{i,j} =: U_i \in \mathcal{B}_a$ such that $\bar{\mu}(B) \leq \bar{\mu}(U_i) = \hat{\mu}(U_i) \leq \sum_{j=1}^\infty \mu(U_{i,j}) < \bar{\mu}(B) + 2^{-i} < +\infty$. Then, $\bar{\mu}(U_i) = \bar{\mu}(B) + \bar{\mu}(U_i \setminus B) < \bar{\mu}(B) + 2^{-i} < +\infty$. This implies that $\bar{\mu}(U_i \setminus B) < 2^{-i}$. Let $U := \bigcap_{i=1}^\infty U_i \in \mathcal{B}_a$. Clearly, $B \subseteq U$ and $\bar{\mu}(U \setminus B) \leq \bar{\mu}(U_i \setminus B) < 2^{-i}$, $\forall i \in \mathbb{N}$. Hence, $\bar{\mu}(U \setminus B) = 0$.

Now, we prove the general case $\bar{\mu}(B) \in [0, \infty] \subset \mathbb{R}_e$. $B = \bigcup_{i=1}^\infty (X_i \cap B) =: \bigcup_{i=1}^\infty B_i$. $\forall i \in \mathbb{N}$, $\bar{\mu}(B_i) \leq \bar{\mu}(X_i) = \mu(X_i) < +\infty$. By the special case, $\exists U_i \in \mathcal{B}_a$ with $B_i \subseteq U_i$ such that $\bar{\mu}(U_i \setminus B_i) = 0$. Then, $B \subseteq \bigcup_{i=1}^\infty U_i =: U \in \mathcal{B}_a$ and $0 \leq \bar{\mu}(U \setminus B) \leq \sum_{i=1}^\infty \bar{\mu}(U_i \setminus B) \leq \sum_{i=1}^\infty \bar{\mu}(U_i \setminus B_i) = 0$, where the second inequality follows from the countable subadditivity of measure. This completes the proof of the claim. $\square$

$\forall B \in \mathcal{B}$, by Claim 11.19.6, $\exists U, \tilde{U} \in \mathcal{B}_a$ with $B \subseteq U$ and $\tilde{B} \subseteq \tilde{U}$ such that $\bar{\mu}(U \setminus B) = \bar{\mu}(\tilde{U} \setminus \tilde{B}) = 0$. Let $F := \tilde{U} \in \mathcal{B}_a$. Then, $F \subseteq B$ and $\bar{\mu}(B \setminus F) = \bar{\mu}(\tilde{U} \setminus \tilde{B}) = 0$. Then, $B = F \cup (B \setminus F)$, $B \setminus F \subseteq U \setminus F$, $F, U \setminus F \in \mathcal{B}_a$ and $\hat{\mu}(U \setminus F) = \bar{\mu}|_{\mathcal{B}_a}(U \setminus F) = \bar{\mu}(U \setminus B) + \bar{\mu}(B \setminus F) = 0$. This shows that $B \in \hat{\mathcal{B}}$, by the completeness of $(X, \hat{\mathcal{B}}, \hat{\mu})$. By the arbitrariness of B , we have $\mathcal{B} \subseteq \hat{\mathcal{B}}$. Hence, $\mathcal{B} = \hat{\mathcal{B}}$ and $\bar{\mu} = \hat{\mu}$. Hence $(X, \mathcal{B}, \bar{\mu})$ is the completion of $(X, \mathcal{B}_a, \bar{\mu}|_{\mathcal{B}_a})$.

This completes the proof of the theorem. $\square$

Definition 11.20 An interval I in $\mathbb{R}$ is any of the following sets: (i) $\emptyset$; (ii) $(a, b) \subseteq \mathbb{R}$ with $a, b \in \mathbb{R}_e$ and $a < b$; (iii) $[a, b) \subset \mathbb{R}$ with $a \in \mathbb{R}$, $b \in \mathbb{R}_e$, and $a < b$; (iv) $(a, b] \subset \mathbb{R}$ with $a \in \mathbb{R}_e$, $b \in \mathbb{R}$, and $a < b$; and (v) $[a, b] \subset \mathbb{R}$ with $a, b \in \mathbb{R}$ and $a \leq b$.

Example 11.21 Let $X = \mathbb{R}$,

$$\mathcal{A} := \{E \subseteq \mathbb{R} \mid E \text{ is the union of finitely many intervals in } \mathbb{R}\}$$

It is clear that $\mathcal{A}$ is an algebra on $\mathbb{R}$. $\forall E \in \mathcal{A}$, E may be written as the union of finitely many pairwise disjoint intervals: $E = \bigcup_{i=1}^n I_i$, for some $n \in \mathbb{N}$. Define $\mu(E) = \sum_{i=1}^n \mu(I_i)$, where

$$\mu(I) = \begin{cases} 0 & \text{if } I = \emptyset \\ b - a & \text{if } I \text{ is any other types of interval} \end{cases}$$

Clearly, $\mu(\emptyset) = 0$; $\mu(I_1 \cup I_2) = \mu(I_1) + \mu(I_2)$, $\forall$ intervals $I_1, I_2 \subseteq \mathbb{R}$ with $I_1 \cap I_2 = \emptyset$. Then, $\mu : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R}_e$ is well-defined and $\mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2)$, $\forall E_1, E_2 \in \mathcal{A}$ with $E_1 \cap E_2 = \emptyset$. Hence, μ is finitely additive.

Fix any interval $I = \bigcup_{i=1}^{\infty} E_i$, where $(E_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ and $E_1, E_2, \dots$ are pairwise disjoint. We will show that $\mu(I) = \sum_{i=1}^{\infty} \mu(E_i)$. $\forall n \in \mathbb{N}$, $I \supseteq \bigcup_{i=1}^n E_i$. Then, by finite additivity of μ , $\mu(I) \geq \sum_{i=1}^n \mu(E_i)$. Hence, $\mu(I) \geq \sum_{i=1}^{\infty} \mu(E_i)$. We will distinguish three exhaustive and mutually exclusive cases: Case 1: $\sum_{i=1}^{\infty} \mu(E_i) = \infty$; Case 2: $I = \emptyset$; Case 3: $I \neq \emptyset$ and $\sum_{i=1}^{\infty} \mu(E_i) < \infty$. Case 1: $\sum_{i=1}^{\infty} \mu(E_i) = \infty$. Then, $\infty \geq \mu(I) \geq \sum_{i=1}^{\infty} \mu(E_i) = \infty$. Hence, $\mu(I) = \sum_{i=1}^{\infty} \mu(E_i) = \infty$. Case 2: $I = \emptyset$. Then, $0 = \mu(I) \leq \sum_{i=1}^{\infty} \mu(E_i)$. Hence, we have $0 = \mu(I) = \sum_{i=1}^{\infty} \mu(E_i)$. Case 3: $I \neq \emptyset$ and $\sum_{i=1}^{\infty} \mu(E_i) < \infty$. $\forall i \in \mathbb{N}$, $E_i = \bigcup_{j=1}^{n_i} I_{i,j}$, where $n_i \in \mathbb{N}$ and $I_{i,1}, \dots, I_{i,n_i}$ are pairwise disjoint intervals. Clearly, $\mu(E_i) = \sum_{j=1}^{n_i} \mu(I_{i,j})$. Then, $I = \bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{n_i} I_{i,j}$. Then, we have $\sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \mu(I_{i,j}) < \infty$. Fix any closed interval $[a, b] := \bar{I} \subseteq I$, where $a, b \in \mathbb{R}$ and $a \leq b$. Such $a, b \in \mathbb{R}$ exists since $I \neq \emptyset$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, $\forall j \in \{1, \dots, n_i\}$, we may enlarge $I_{i,j}$ to an open interval $\bar{I}_{i,j} \supseteq I_{i,j}$ with $\mu(\bar{I}_{i,j}) \leq \mu(I_{i,j}) + \frac{2^{-i}\epsilon}{n_i}$. Then, $\bar{I} \subseteq \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{n_i} \bar{I}_{i,j}$. By the compactness of $\bar{I}$, $\exists N \in \mathbb{N}$ such that $\bar{I} \subseteq \bigcup_{i=1}^N \bigcup_{j=1}^{n_i} \bar{I}_{i,j}$. Clearly, we have $b - a = \mu(\bar{I}) \leq \sum_{i=1}^N \sum_{j=1}^{n_i} \mu(\bar{I}_{i,j}) \leq \sum_{i=1}^N \sum_{j=1}^{n_i} (\mu(I_{i,j}) + \frac{2^{-i}\epsilon}{n_i}) \leq \epsilon + \sum_{i=1}^N \sum_{j=1}^{n_i} \mu(I_{i,j})$. By the arbitrariness of ϵ, a, b , we have $\mu(I) \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \mu(I_{i,j})$. Then, $\mu(I) \leq \sum_{i=1}^{\infty} \mu(E_i)$. Thus, we have $\mu(I) = \sum_{i=1}^{\infty} \mu(E_i)$. Hence, in all three cases, we have $\mu(I) = \sum_{i=1}^{\infty} \mu(E_i)$.

$\forall E \in \mathcal{A}$, $\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ with $E = \bigcup_{i=1}^{\infty} E_i$ and $E_1, E_2, \dots$ being pairwise disjoint, we have $E = \bigcup_{i=1}^n I_i$, where $n \in \mathbb{N}$ and $I_1, \dots, I_n$ are pairwise disjoint intervals. Then, $I_j = \bigcup_{i=1}^{\infty} (I_j \cap E_i)$, $\forall j \in \{1, \dots, n\}$. By the above argument, we have $\mu(I_j) = \sum_{i=1}^{\infty} \mu(I_j \cap E_i)$, $\forall j \in \{1, \dots, n\}$. Then, $\mu(E) = \sum_{j=1}^n \mu(I_j) = \sum_{j=1}^n \sum_{i=1}^{\infty} \mu(I_j \cap E_i) = \sum_{i=1}^{\infty} \sum_{j=1}^n \mu(I_j \cap E_i) = \sum_{i=1}^{\infty} \mu(E_i)$, where the last equality follows from the finite additivity of μ .

Hence, μ is a measure on algebra $\mathcal{A}$ and is σ -finite. By Carathéodory Extension Theorem 11.19, μ extends uniquely to a σ -finite complete measure space $(X, \mathcal{B}, \bar{\mu})$. This measure space is called the *Lebesgue measure space*. Since $\mathcal{B}$ is a σ -algebra on $\mathbb{R}$ and the topology $\mathcal{O}_{\mathbb{R}}$ on $\mathbb{R}$ is second countable, then $\mathcal{B} \supseteq \mathcal{O}_{\mathbb{R}}$. We will denote the Lebesgue measure space by $(\mathbb{R}, \mathcal{B}_L, \mu_L)$, where $\mathcal{B}_L$ is the collection of *Lebesgue measurable sets* in $\mathbb{R}$ and μ_L is the *Lebesgue measure*.

Let μ_{Lo} be the *Lebesgue outer measure* induced by μ as defined in Carathéodory Extension Theorem 11.19. We claim that

$$\mu_{Lo}(E) = \inf_{(I_i)_{i=1}^\infty \text{ are open intervals, } E \subseteq \bigcup_{i=1}^\infty I_i} \sum_{i=1}^\infty \mu(I_i); \quad \forall E \subseteq \mathbb{R} \quad (11.3)$$

To prove this, we note that $\mu_{Lo}(E) = \inf_{(E_i)_{i=1}^\infty \subseteq \mathcal{A}, E \subseteq \bigcup_{i=1}^\infty E_i} \sum_{i=1}^\infty \mu(E_i)$. Then, clearly, we have

$$\mu_{Lo}(E) \leq \inf_{(I_i)_{i=1}^\infty \text{ are open intervals, } E \subseteq \bigcup_{i=1}^\infty I_i} \sum_{i=1}^\infty \mu(I_i)$$

On the other hand, $\forall (E_i)_{i=1}^\infty \subseteq \mathcal{A}$ with $E \subseteq \bigcup_{i=1}^\infty E_i$, $\forall i \in \mathbb{N}$, $E_i = \bigcup_{j=1}^{n_i} I_{i,j}$, where $n_i \in \mathbb{N}$ and $I_{i,1}, \dots, I_{i,n_i}$ are pairwise disjoint intervals, and $\mu(E_i) = \sum_{j=1}^{n_i} \mu(I_{i,j})$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall j \in \{1, \dots, n_i\}$, we may enlarge $I_{i,j}$ to an open interval $\bar{I}_{i,j} \supseteq I_{i,j}$ with $\mu(\bar{I}_{i,j}) \leq \mu(I_{i,j}) + \frac{2^{-i}\epsilon}{n_i}$. Then, we have $E \subseteq \bigcup_{i=1}^\infty \bigcup_{j=1}^{n_i} I_{i,j} \subseteq \bigcup_{i=1}^\infty \bigcup_{j=1}^{n_i} \bar{I}_{i,j}$ and $\sum_{i=1}^\infty \mu(E_i) = \sum_{i=1}^\infty \sum_{j=1}^{n_i} \mu(I_{i,j}) \geq \sum_{i=1}^\infty \sum_{j=1}^{n_i} (\mu(\bar{I}_{i,j}) - \frac{2^{-i}\epsilon}{n_i}) = \sum_{i=1}^\infty \sum_{j=1}^{n_i} \mu(\bar{I}_{i,j}) - \epsilon \geq -\epsilon + \inf_{(I_i)_{i=1}^\infty \text{ are open intervals, } E \subseteq \bigcup_{i=1}^\infty I_i} \sum_{i=1}^\infty \mu(I_i)$. This implies that $\mu_{Lo}(E) \geq -\epsilon + \inf_{(I_i)_{i=1}^\infty \text{ are open intervals, } E \subseteq \bigcup_{i=1}^\infty I_i} \sum_{i=1}^\infty \mu(I_i)$. Hence, by the arbitrariness of ϵ , we have

$$\mu_{Lo}(E) \geq \inf_{(I_i)_{i=1}^\infty \text{ are open intervals, } E \subseteq \bigcup_{i=1}^\infty I_i} \sum_{i=1}^\infty \mu(I_i)$$

Therefore, (11.3) is true.

By Carathéodory Extension Theorem 11.19, $\mu_L = \mu_{Lo}|_{\mathcal{B}_L}$ and $(\mathbb{R}, \mathcal{B}_L, \mu_L)$ is the completion of $(\mathbb{R}, \mathcal{B}_a, \mu_L|_{\mathcal{B}_a})$, where $\mathcal{B}_a$ is the σ -algebra on $\mathbb{R}$ generated by $\mathcal{A}$. Clearly, $\mathcal{O}_{\mathbb{R}} \subseteq \mathcal{B}_a$ and $\mathcal{B}_a$ is the σ -algebra on $\mathbb{R}$ generated by $\mathcal{O}_{\mathbb{R}}$. $\diamond$

Definition 11.22 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space. The collection of Borel sets on $\mathcal{X}$ is the smallest σ -algebra on X that contains $\mathcal{O}$, which will be denoted by $\mathcal{B}_{\mathbb{B}}(\mathcal{X})$. The smallest algebra on X that contains $\mathcal{O}$ will be denoted by $\mathcal{A}(\mathcal{X})$.

Clearly, we have $\mathcal{B}_{\mathbb{B}}(\mathbb{R}) \subseteq \mathcal{B}_L$.

Proposition 11.23 Let $E \subseteq \mathbb{R}$. Then, the following statements are equivalent.

(i) $E \in \mathcal{B}_L$.

(ii) $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists O \in \mathcal{O}_{\mathbb{R}}$ with $E \subseteq O$ such that $\mu_{Lo}(O \setminus E) < \epsilon$.

(iii) $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists F \subseteq \mathbb{R}$ with $\tilde{F} \in \mathcal{O}_{\mathbb{R}}$ and $F \subseteq E$ such that $\mu_{Lo}(E \setminus F) < \epsilon$.

If $\mu_{Lo}(E) < +\infty$. Then, (i)–(iii) are also equivalent to

(iv) $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists U \in \mathcal{O}_{\mathbb{R}}$ which is finite union of open intervals, such that $\mu_{Lo}(U \triangle E) < \epsilon$.

Hence, the Lebesgue measure space $(\mathbb{R}, \mathcal{B}_L, \mu_L)$ is the completion of the Borel measure space $(\mathbb{R}, \mathcal{B}_B(\mathbb{R}), \mu_B := \mu_L|_{\mathcal{B}_B(\mathbb{R})})$.

Proof Fix $E \subseteq \mathbb{R}$.

(i) $\Rightarrow$ (ii). First, we show the special case that $\mu_{Lo}(E) < +\infty$. Let $E \in \mathcal{B}_L$. By Example 11.21,

$$\mu_{Lo}(E) = \mu_L(E) = \inf_{(I_i)_{i=1}^{\infty} \text{ are open intervals, } E \subseteq \bigcup_{i=1}^{\infty} I_i} \sum_{i=1}^{\infty} \mu_L(I_i)$$

Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists$ a sequence of open intervals $(I_i)_{i=1}^{\infty} \subseteq \mathcal{O}_{\mathbb{R}}$ with $E \subseteq \bigcup_{i=1}^{\infty} I_i =: O \in \mathcal{O}_{\mathbb{R}}$ such that $\mu_L(E) \leq \mu_L(O) \leq \sum_{i=1}^{\infty} \mu_L(I_i) < \mu_L(E) + \epsilon < +\infty$. Then, $\epsilon > \mu_L(O) - \mu_L(E) = \mu_L(O \setminus E) = \mu_{Lo}(O \setminus E)$, where the first equality follows from the countable additivity of μ_L . Hence, (ii) holds.

Next, we show the general case $\mu_{Lo}(E) \in [0, \infty] \subset \mathbb{R}_e$. Let $E \in \mathcal{B}_L$ and $I_1 := (-1, 1] \subset \mathbb{R}$, $I_n := (-n, -n+1] \cup (n-1, n] \subset \mathbb{R}$, $\forall n = 2, 3, \dots$. Clearly, $E = \bigcup_{i=1}^{\infty} (I_i \cap E) =: \bigcup_{i=1}^{\infty} E_i$. Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. $\forall i \in \mathbb{N}$, $I_i \in \mathcal{B}_L$ and $E_i \in \mathcal{B}_L$ with $\mu_L(E_i) \leq \mu_L(I_i) = 2$. By the special case we have just shown, $\exists O_i \in \mathcal{O}_{\mathbb{R}} \subseteq \mathcal{B}_L$ with $E_i \subseteq O_i$ such that $\mu_{Lo}(O_i \setminus E_i) = \mu_L(O_i \setminus E_i) < 2^{-i}\epsilon$. Let $O := \bigcup_{i=1}^{\infty} O_i \in \mathcal{O}_{\mathbb{R}}$. Then, $\mu_{Lo}(O \setminus E) = \mu_L(O \setminus E) = \mu_L((\bigcup_{i=1}^{\infty} O_i) \cap \widetilde{\bigcup_{i=1}^{\infty} E_i}) = \mu_L((\bigcup_{i=1}^{\infty} O_i) \cap (\bigcap_{i=1}^{\infty} \widetilde{E_i})) = \mu_L(\bigcup_{i=1}^{\infty} (O_i \cap (\bigcap_{j=1}^{\infty} \widetilde{E_j}))) \leq \sum_{i=1}^{\infty} \mu_L(O_i \cap (\bigcap_{j=1}^{\infty} \widetilde{E_j})) \leq \sum_{i=1}^{\infty} \mu_L(O_i \cap \widetilde{E_i}) = \sum_{i=1}^{\infty} \mu_L(O_i \setminus E_i) < \epsilon$. Hence (ii) holds.

(ii) $\Rightarrow$ (i). Fix any $A \subseteq \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by (ii), $\exists O \in \mathcal{O}_{\mathbb{R}}$ with $E \subseteq O$ such that $\mu_{Lo}(O \setminus E) < \epsilon$. Note that $A \cap \widetilde{E} = (A \cap \widetilde{E} \cap O) \cup (A \cap \widetilde{E} \cap \widetilde{O}) \subseteq (O \setminus E) \cup (A \cap \widetilde{O})$ and $A \cap E \subseteq A \cap O$. By Definition 11.14, we have $\mu_{Lo}(A) \leq \mu_{Lo}(A \cap E) + \mu_{Lo}(A \cap \widetilde{E}) \leq \mu_{Lo}(A \cap O) + \mu_{Lo}(O \setminus E) + \mu_{Lo}(A \cap \widetilde{O}) \leq \mu_{Lo}(A) + \epsilon$, where the third inequality follows from the fact that $O \in \mathcal{O}_{\mathbb{R}} \subseteq \mathcal{B}_L$ and Definition 11.15. By the arbitrariness of ϵ , we have $\mu_{Lo}(A) = \mu_{Lo}(A \cap E) + \mu_{Lo}(A \cap \widetilde{E})$. By the arbitrariness of A , we have $E \in \mathcal{B}_L$.

(i) $\Leftrightarrow$ (iii). $E \in \mathcal{B}_L \Leftrightarrow \widetilde{E} \in \mathcal{B}_L \Leftrightarrow \forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists O \in \mathcal{O}_{\mathbb{R}}$ with $\widetilde{E} \subseteq O$ such that $\mu_{Lo}(O \setminus \widetilde{E}) < \epsilon$, where the last $\Leftrightarrow$ follows from (i) $\Leftrightarrow$ (ii). Let $F := \widetilde{O}$, we have $F \subseteq E$ and $\mu_{Lo}(E \setminus F) = \mu_{Lo}(O \setminus \widetilde{E}) < \epsilon$.

Next, we show (i) – (iii) are equivalent to (iv) under the additional assumption that $\mu_{Lo}(E) < +\infty$.

(i) $\Rightarrow$ (iv). Let $E \in \mathcal{B}_L$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Example 11.21, $\exists$ a sequence of open intervals $(I_i)_{i=1}^\infty$ with $E \subseteq \bigcup_{i=1}^\infty I_i =: O \in \mathcal{O}_\mathbb{R}$ such that $\mu_{Lo}(E) = \mu_L(E) \leq \sum_{i=1}^\infty \mu_L(I_i) \leq \mu_L(E) + \epsilon/2 < +\infty$. Then, $\exists N \in \mathbb{N}$ such that $\sum_{i=N+1}^\infty \mu_L(I_i) < \epsilon/2$. Let $U := \bigcup_{i=1}^N I_i \in \mathcal{O}_\mathbb{R}$. We have $\mu_{Lo}(U \triangle E) = \mu_L(U \triangle E) = \mu_L(U \setminus E) + \mu_L(E \setminus U) \leq \mu_L(O \setminus E) + \mu_L(E \cap \tilde{U})$. Note that $\mu_L(O) = \mu_L(E) + \mu_L(O \setminus E) \leq \sum_{i=1}^\infty \mu_L(I_i) \leq \mu_L(E) + \epsilon/2 < +\infty$ and $E \cap \tilde{U} = (E \cap \tilde{U} \cap O) \cup (E \cap \tilde{U} \cap \tilde{O}) = E \cap \tilde{U} \cap O \subseteq O \setminus U \subseteq \bigcup_{i=N+1}^\infty I_i$. Then, $\mu_{Lo}(U \triangle E) \leq \epsilon/2 + \mu_L(\bigcup_{i=N+1}^\infty I_i) \leq \epsilon/2 + \sum_{i=N+1}^\infty \mu_L(I_i) < \epsilon$. Hence, (iv) holds.

(iv) $\Rightarrow$ (i). Fix any set $A \subseteq \mathbb{R}$. By (iv), $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists U \in \mathcal{O}_\mathbb{R}$ which is finite union of open intervals, such that $\mu_{Lo}(U \triangle E) < \epsilon/2$. Note that $A \cap E = (A \cap U \cap E) \cup (A \cap E \cap \tilde{U}) \subseteq (A \cap U) \cup (E \setminus U)$ and $A \cap \tilde{E} = (A \cap \tilde{E} \cap U) \cup (A \cap \tilde{E} \cap \tilde{U}) \subseteq (U \setminus E) \cup (A \cap \tilde{U})$. By Definition 11.14, $\mu_{Lo}(A) \leq \mu_{Lo}(A \cap E) + \mu_{Lo}(A \cap \tilde{E}) \leq \mu_{Lo}((A \cap U) \cup (E \setminus U)) + \mu_{Lo}((U \setminus E) \cup (A \cap \tilde{U})) \leq \mu_{Lo}(A \cap U) + \mu_{Lo}(E \setminus U) + \mu_{Lo}(U \setminus E) + \mu_{Lo}(A \cap \tilde{U}) = \mu_{Lo}(A) + \mu_{Lo}(E \setminus U) + \mu_{Lo}(U \setminus E) \leq \mu_{Lo}(A) + 2\mu_{Lo}(E \triangle U) \leq \mu_{Lo}(A) + \epsilon$, where the equality follows from the fact that $U \in \mathcal{O}_\mathbb{R} \subseteq \mathcal{B}_L$ and Definition 11.15. By the arbitrariness of ϵ , we have $\mu_{Lo}(A) = \mu_{Lo}(A \cap E) + \mu_{Lo}(A \cap \tilde{E})$. By the arbitrariness of A , we have $E \in \mathcal{B}_L$. Hence (i) holds.

Finally, by Example 11.21, the Lebesgue measure space is the completion of the Borel measure space $(\mathbb{R}, \mathcal{B}_B(\mathbb{R}), \mu_L|_{\mathcal{B}_B(\mathbb{R})})$. This completes the proof of the proposition. $\square$

Proposition 11.24 *Let $\mathcal{X}_i := (X_i, \mathcal{O}_i)$ be a second countable topological space, $i = 1, 2$, $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2 := (X_1 \times X_2, \mathcal{O})$ be the product topological space, $\mathcal{E} := \{B_1 \times B_2 \subseteq X_1 \times X_2 \mid B_i \in \mathcal{B}_B(\mathcal{X}_i), i = 1, 2\}$, and $\mathcal{B}$ be the σ -algebra on $X_1 \times X_2$ generated by $\mathcal{E}$. Then, $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$.*

Proof Let $\mathcal{O}_{Bi} \subseteq \mathcal{O}_i$ be a countable basis for $\mathcal{X}_i$, $i = 1, 2$. Without loss of generality, assume that $X_i \in \mathcal{O}_{Bi}$, $i = 1, 2$. By Proposition 3.28, $\mathcal{X}$ is second countable with a countable basis $\mathcal{O}_B := \{O_{B1} \times O_{B2} \subseteq X_1 \times X_2 \mid O_{Bi} \in \mathcal{O}_{Bi}, i = 1, 2\} \subseteq \mathcal{O}$. Clearly, $\mathcal{O}_{Bi} \subseteq \mathcal{O}_i \subseteq \mathcal{B}_B(\mathcal{X}_i)$, $i = 1, 2$. Then, $\mathcal{O}_B \subseteq \mathcal{E}$. $\forall O \in \mathcal{O}$, $O = \bigcup_{i=1}^\infty O_i$, where $O_i \in \mathcal{O}_B$, $\forall i \in \mathbb{N}$. Hence, $O \in \mathcal{B}$. This implies that $\mathcal{O} \subseteq \mathcal{B}$. Since $\mathcal{B}$ is a σ -algebra, then $\mathcal{O} \subseteq \mathcal{B}_B(\mathcal{X}) \subseteq \mathcal{B}$.

On the other hand, note that $\mathcal{B}_B(\mathcal{X})$ is the σ -algebra generated by $\mathcal{O}$. Then, $\forall O_1 \in \mathcal{O}_1, \forall O_2 \in \mathcal{O}_2$, we have $O_1 \times O_2 \in \mathcal{O} \subseteq \mathcal{B}_B(\mathcal{X})$.

Claim 11.24.1 $\forall O_1 \in \mathcal{O}_1, \forall B_2 \in \mathcal{B}_B(\mathcal{X}_2)$, we have $O_1 \times B_2 \in \mathcal{B}_B(\mathcal{X})$.

Proof of claim: Fix any $O_1 \in \mathcal{O}_1$. Define $\mathcal{F} := \{E \subseteq X_2 \mid O_1 \times E \in \mathcal{B}_B(\mathcal{X})\}$. Clearly, $\mathcal{O}_2 \subseteq \mathcal{F}$. Then, $\emptyset, X_2 \in \mathcal{O}_2 \subseteq \mathcal{F}$. $\forall E \in \mathcal{F}$, we have $O_1 \times E \in \mathcal{B}_B(\mathcal{X})$. Note that $O_1 \times X_2 \in \mathcal{B}_B(\mathcal{X})$. This leads to $O_1 \times (X_2 \setminus E) = (O_1 \times X_2) \setminus (O_1 \times E) \in \mathcal{B}_B(\mathcal{X})$. Thus, $X_2 \setminus E \in \mathcal{F}$. $\forall (E_i)_{i=1}^\infty \subseteq \mathcal{F}$, we have $O_1 \times E_i \in \mathcal{B}_B(\mathcal{X})$, $\forall i \in \mathbb{N}$. This implies that

$O_1 \times (\bigcup_{i=1}^{\infty} E_i) = \bigcup_{i=1}^{\infty} (O_1 \times E_i) \in \mathcal{B}_B(\mathcal{X})$ and $\bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$. The above shows that $\mathcal{F}$ is a σ -algebra on X_2 . Then, $\mathcal{O}_2 \subseteq \mathcal{B}_B(\mathcal{X}_2) \subseteq \mathcal{F}$. Hence, $\forall B_2 \in \mathcal{B}_B(\mathcal{X}_2)$, we have $B_2 \in \mathcal{F}$ and $O_1 \times B_2 \in \mathcal{B}_B(\mathcal{X})$. This completes the proof of the claim. $\square$

Claim 11.24.2 $\forall B_1 \in \mathcal{B}_B(\mathcal{X}_1), \forall B_2 \in \mathcal{B}_B(\mathcal{X}_2)$, we have $B_1 \times B_2 \in \mathcal{B}_B(\mathcal{X})$.

Proof of claim: Fix any $B_2 \in \mathcal{B}_B(\mathcal{X}_2)$. Define $\mathcal{F} := \{E \subseteq X_1 \mid E \times B_2 \in \mathcal{B}_B(\mathcal{X})\}$. By Claim 11.24.1, $\mathcal{O}_1 \subseteq \mathcal{F}$. Then, $\emptyset, X_1 \in \mathcal{O}_1 \subseteq \mathcal{F}$. $\forall E \in \mathcal{F}$, we have $E \times B_2 \in \mathcal{B}_B(\mathcal{X})$. Note that $X_1 \times B_2 \in \mathcal{B}_B(\mathcal{X})$. This leads to $(X_1 \setminus E) \times B_2 = (X_1 \times B_2) \setminus (E \times B_2) \in \mathcal{B}_B(\mathcal{X})$. Thus, $X_1 \setminus E \in \mathcal{F}$. $\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{F}$, we have $E_i \times B_2 \in \mathcal{B}_B(\mathcal{X}), \forall i \in \mathbb{N}$. This implies that $(\bigcup_{i=1}^{\infty} E_i) \times B_2 = \bigcup_{i=1}^{\infty} (E_i \times B_2) \in \mathcal{B}_B(\mathcal{X})$ and $\bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$. The above shows that $\mathcal{F}$ is a σ -algebra on X_1 . Then, $\mathcal{O}_1 \subseteq \mathcal{B}_B(\mathcal{X}_1) \subseteq \mathcal{F}$. Hence, $\forall B_1 \in \mathcal{B}_B(\mathcal{X}_1)$, we have $B_1 \in \mathcal{F}$ and $B_1 \times B_2 \in \mathcal{B}_B(\mathcal{X})$. This completes the proof of the claim. $\square$

By Claim 11.24.2, $\mathcal{E} \subseteq \mathcal{B}_B(\mathcal{X})$. By the definition of $\mathcal{B}$, we have $\mathcal{B} \subseteq \mathcal{B}_B(\mathcal{X})$. Hence, $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$. This completes the proof of the proposition. $\square$

Proposition 11.25 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $E \in \mathcal{B}_B(\mathcal{X})$, $\mathcal{E} := (E, \mathcal{O}_E)$ be the topological space with the subset topology of $\mathcal{X}$, and $\mathcal{B}_E := \{C \subseteq E \mid C \in \mathcal{B}_B(\mathcal{X})\}$. Then, $\mathcal{B}_E = \mathcal{B}_B(\mathcal{E})$.

Proof Clearly $\emptyset, E \in \mathcal{B}_E$. $\forall A \in \mathcal{B}_E$, we have $A \subseteq E$ and $A \in \mathcal{B}_B(\mathcal{X})$. Then, $E \supseteq E \setminus A \in \mathcal{B}_B(\mathcal{X})$ and $E \setminus A \in \mathcal{B}_E$. $\forall (A_i)_{i=1}^{\infty} \subseteq \mathcal{B}_E$, we have $E \supseteq A_i \in \mathcal{B}_B(\mathcal{X}), \forall i \in \mathbb{N}$. Then, $E \supseteq \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}_B(\mathcal{X})$ and $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}_E$. This shows that $\mathcal{B}_E$ is a σ -algebra on E .

$\forall O_E \in \mathcal{O}_E, \exists O \in \mathcal{O} \subseteq \mathcal{B}_B(\mathcal{X})$ such that $O_E = O \cap E$. Then, $O_E \in \mathcal{B}_E$ and $\mathcal{O}_E \subseteq \mathcal{B}_E$. Hence, $\mathcal{O}_E \subseteq \mathcal{B}_B(\mathcal{E}) \subseteq \mathcal{B}_E$.

On the other hand, let $\tilde{\mathcal{E}} := (X \setminus E, \mathcal{O}_{X \setminus E})$ be the topological space with the subset topology of $\mathcal{X}$.

Claim 11.25.1 $\forall C \in \mathcal{B}_B(\mathcal{X})$, we have $C_1 := C \cap E \in \mathcal{B}_B(\mathcal{E})$ and $C_2 := C \cap (X \setminus E) \in \mathcal{B}_B(\tilde{\mathcal{E}})$.

Proof of claim: Define $\mathcal{F} := \left\{ A_1 \cup A_2 \subseteq X \mid A_1 \in \mathcal{B}_B(\mathcal{E}), A_2 \in \mathcal{B}_B(\tilde{\mathcal{E}}) \right\}$. $\forall O \in \mathcal{O}, O \cap E \in \mathcal{O}_E \subseteq \mathcal{B}_B(\mathcal{E})$ and $O \cap (X \setminus E) \in \mathcal{O}_{X \setminus E} \subseteq \mathcal{B}_B(\tilde{\mathcal{E}})$. Then, $O = (O \cap E) \cup (O \cap (X \setminus E)) \in \mathcal{F}$ and $\mathcal{O} \subseteq \mathcal{F}$.

Clearly, $\emptyset, E \in \mathcal{B}_B(\mathcal{E})$ and $\emptyset, X \setminus E \in \mathcal{B}_B(\tilde{\mathcal{E}})$. Then, $\emptyset = \emptyset \cup \emptyset \in \mathcal{F}$ and $X = E \cup (X \setminus E) \in \mathcal{F}$. $\forall A \in \mathcal{F}, \exists A_1 \in \mathcal{B}_B(\mathcal{E})$ and $\exists A_2 \in \mathcal{B}_B(\tilde{\mathcal{E}})$ such that $A = A_1 \cup A_2$. Then, $A_1 = A \cap E$ and $A_2 = A \cap (X \setminus E)$. Note that

$X \setminus A = (E \cup (X \setminus E)) \setminus A = (E \setminus A) \cup ((X \setminus E) \setminus A) = (E \setminus A_1) \cup ((X \setminus E) \setminus A_2)$, $E \setminus A_1 \in \mathcal{B}_B(\mathcal{E})$, and $(X \setminus E) \setminus A_2 \in \mathcal{B}_B(\tilde{\mathcal{E}})$. This implies that $X \setminus A \in \mathcal{F}$. $\forall (A_i)_{i=1}^\infty \subseteq \mathcal{F}$, let $A_{i,1} := E \cap A_i \in \mathcal{B}_B(\mathcal{E})$ and $A_{i,2} := (X \setminus E) \cap A_i \in \mathcal{B}_B(\tilde{\mathcal{E}})$, $\forall i \in \mathbb{N}$. Note that $\bigcup_{i=1}^\infty A_i = \bigcup_{i=1}^\infty (A_{i,1} \cup A_{i,2}) = (\bigcup_{i=1}^\infty A_{i,1}) \cup (\bigcup_{i=1}^\infty A_{i,2})$, $\bigcup_{i=1}^\infty A_{i,1} \in \mathcal{B}_B(\mathcal{E})$, and $\bigcup_{i=1}^\infty A_{i,2} \in \mathcal{B}_B(\tilde{\mathcal{E}})$. This implies that $\bigcup_{i=1}^\infty A_i \in \mathcal{F}$. The above shows that $\mathcal{F}$ is a σ -algebra on X . Then, $\mathcal{O} \subseteq \mathcal{B}_B(\mathcal{X}) \subseteq \mathcal{F}$. This completes the proof of the claim. $\square$

$\forall C \in \mathcal{B}_E$, we have $E \supseteq C \in \mathcal{B}_B(\mathcal{X})$. By Claim 11.25.1, $C = C \cap E \in \mathcal{B}_B(\mathcal{E})$. Then, $\mathcal{B}_E \subseteq \mathcal{B}_B(\mathcal{E})$. Therefore, $\mathcal{B}_E = \mathcal{B}_B(\mathcal{E})$. This completes the proof of the proposition. $\square$

Definition 11.26 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $(X, \mathcal{B}, \mu)$ be a measure space on the same set X . The triple $\mathbb{X} := (\mathcal{X}, \mathcal{B}, \mu)$ is said to be a topological measure space if $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$ and $\forall E \in \mathcal{B}_B(\mathcal{X})$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists U \in \mathcal{O}$ with $E \subseteq U$ such that $\mu(U \setminus E) < \epsilon$. We will say that $\mathbb{X}$ is finite or σ -finite if the underlying measure space is so. We will say that $\mathbb{X}$ is Tychonoff, Hausdorff, regular, completely regular, or normal if $\mathcal{X}$ is so. We will say that $\mathbb{X}$ is first countable or second countable if $\mathcal{X}$ is so. We will say that $\mathbb{X}$ is separable, second category everywhere, connected, or locally connected if $\mathcal{X}$ is so. We will say that $\mathbb{X}$ is compact, countably compact, sequentially compact, locally compact, σ -compact, or paracompact if $\mathcal{X}$ is so. We will say that $\mathbb{X}$ is locally finite if $\forall$ compact $K \subseteq \mathcal{X}$, $\mu(K) < \infty$.

Let $\mathcal{X} := (X, \rho)$ be a metric space, $\mathcal{O}$ be the natural topology on $\mathcal{X}$ generated by the metric ρ , and $(X, \mathcal{B}, \mu)$ be a measure space on the same set X . The triple $\mathbb{X} := (\mathcal{X}, \mathcal{B}, \mu)$ is said to be a metric measure space if $((X, \mathcal{O}), \mathcal{B}, \mu)$ is a topological measure space. $\mathbb{X}$ is said to be a complete metric measure space, if $\mathcal{X}$ is a complete metric space. $\mathbb{X}$ is said to be totally bounded if $\mathcal{X}$ is so.

Let $\mathbb{X} := (\mathcal{X}, \mathbb{K}, \|\cdot\|)$ be a normed linear space over the field $\mathbb{K}$, $\mathcal{O}$ be the natural topology on $\mathbb{X}$ generated by the norm $\|\cdot\|$, and $(X, \mathcal{B}, \mu)$ be a measure space on the same set X . The triple $\mathbb{X} := (\mathbb{X}, \mathcal{B}, \mu)$ is said to be a normed linear measure space if $((X, \mathcal{O}), \mathcal{B}, \mu)$ is a topological measure space. Furthermore, if $\mathbb{X}$ is a Banach space, then $\mathbb{X}$ is said to be a Banach measure space. Depending on whether $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, we will say that $\mathbb{X}$ is a real or complex Banach measure space.

Proposition 11.27 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $(X, \mathcal{B}, \mu)$ be a measure space on the same set X , and $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$. The triple $\mathbb{X} := (\mathcal{X}, \mathcal{B}, \mu)$ is topological measure space if, and only if, $\forall E \in \mathcal{B}_B(\mathcal{X})$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists X \setminus F \in \mathcal{O}$ with $F \subseteq E$ such that $\mu(E \setminus F) < \epsilon$.

Proof “Necessity” $\forall E \in \mathcal{B}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $X \setminus E \in \mathcal{B}$. By $\mathbb{X}$ being a topological measure space, $\exists U \in \mathcal{O}$ with $X \setminus E \subseteq U$ such that $\mu(U \setminus (X \setminus E)) < \epsilon$. Then, $X \setminus U \subseteq E$ and $\mu(E \setminus (X \setminus U)) = \mu(E \cap U) = \mu(U \setminus (X \setminus E)) < \epsilon$. Hence, $F := X \setminus U$ is the set we seek.

“Sufficiency” $\forall E \in \mathcal{B}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, X \setminus E \in \mathcal{B}$. By the assumption, $\exists X \setminus F \in \mathcal{O}$ with $F \subseteq X \setminus E$ such that $\mu((X \setminus E) \setminus F) < \epsilon$. Let $U := X \setminus F$. Then, $E \subseteq U$ and $\mu(U \setminus E) = \mu((X \setminus F) \setminus E) = \mu((X \setminus E) \setminus F) < \epsilon$. Hence, $\mathbb{X}$ is a topological measure space. $\square$

Example 11.28 The real line $\mathbb{R}$ is a real Banach space with norm defined to be the absolute value. Then, by Proposition 11.23 and Example 11.21, $\mathbb{R} := ((\mathbb{R}, \mathbb{R}, |\cdot|), \mathcal{B}_B(\mathbb{R}), \mu_B)$ is a σ -finite real Banach measure space. $\diamond$

Proposition 11.29 Let $\mathbb{X} := (\mathcal{X}, \mathcal{B}, \mu)$ be a topological measure space, where $\mathcal{X} := (X, \mathcal{O})$ is a topological space, $E \in \mathcal{B}$, $\mathcal{E} := (E, \mathcal{O}_E)$ be the topological space with the subset topology of $\mathcal{X}$, and $(E, \mathcal{B}_E, \mu_E)$ be the measure subspace of $\mathbb{X}$ as defined in Proposition 11.13. Then, $\mathbb{E} := (\mathcal{E}, \mathcal{B}_E, \mu_E)$ is a topological measure space and is said to be the topological measure subspace of $\mathbb{X}$.

Proof By Definition 11.26, $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$. By Proposition 11.13, $\mathcal{B}_E = \{C \subseteq E \mid C \in \mathcal{B}\}$ and $\mu_E = \mu|_{\mathcal{B}_E}$. By Proposition 11.25, we have $\mathcal{B}_E = \mathcal{B}_B(\mathcal{E})$.

$\forall A \in \mathcal{B}_E, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, E \supseteq A \in \mathcal{B}$. By $\mathbb{X}$ being a topological measure space, $\exists U \in \mathcal{O}$ with $A \subseteq U$ such that $\mu(U \setminus A) < \epsilon$. Let $\bar{U} := U \cap E \in \mathcal{O}_E \subseteq \mathcal{B}_E$. Clearly, $A \subseteq \bar{U}$ and $0 \leq \mu_E(\bar{U} \setminus A) = \mu((U \cap E) \setminus A) \leq \mu(U \setminus A) < \epsilon$. Hence, $\mathbb{E}$ is a topological measure space. This completes the proof of the proposition. $\square$

Definition 11.30 Let X be a set and $\mathcal{C} \subseteq \mathcal{P}(X)$ be a nonempty collection of subsets of X . $\mathcal{C}$ is said to be a semialgebra on X if $\forall C_1, C_2 \in \mathcal{C}, C_1 \cap C_2 \in \mathcal{C}$ and $\widetilde{C_1}$ is a finite disjoint union of sets in $\mathcal{C}$.

Proposition 11.31 Let X be a set and $\mathcal{C} \subseteq \mathcal{P}(X)$ be a semialgebra on X . $\mathcal{A} := \{A \subseteq X \mid A \text{ is the finite disjoint union of sets in } \mathcal{C}\}$. Then, $\mathcal{A}$ is the algebra on X generated by $\mathcal{C}$.

Proof Clearly $\emptyset \in \mathcal{A}$. Fix $C_1 \in \mathcal{C} \neq \emptyset$. Then, $\widetilde{C_1} = \bigcup_{i=2}^n C_i$, where $n \in \mathbb{N}, (C_i)_{i=2}^n \subseteq \mathcal{C}$ is pairwise disjoint. Then, $X = C_1 \cup \widetilde{C_1} = \bigcup_{i=1}^n C_i$, where the sets in the union are pairwise disjoint. Hence, $X \in \mathcal{A}$. $\forall A_1, A_2 \in \mathcal{A}$, then $A_i = \bigcup_{j=1}^{n_i} C_{j,i}$, where $n_i \in \mathbb{Z}_+$ and $(C_{j,i})_{j=1}^{n_i} \subseteq \mathcal{C}$ is pairwise disjoint, $i = 1, 2$. Note that $\widetilde{A_1} = \bigcap_{j=1}^{n_1} \widetilde{C_{j,1}}, \forall j \in \{1, \dots, n_1\}, \widetilde{C_{j,1}} = \bigcup_{l=1}^{m_j} C_{l,j,1}$, where $m_j \in \mathbb{Z}_+$ and $(C_{l,j,1})_{l=1}^{m_j} \subseteq \mathcal{C}$ is pairwise disjoint. Then,

$$\begin{aligned} \widetilde{A_1} &= \bigcap_{j=1}^{n_1} \bigcup_{l=1}^{m_j} C_{l,j,1} = \bigcup_{l_1=1}^{m_1} (C_{l_1,1,1} \cap (\bigcap_{j=2}^{n_1} \bigcup_{l=1}^{m_j} C_{l,j,1})) \\ &= \bigcup_{l_1=1}^{m_1} (C_{l_1,1,1} \cap (\bigcup_{l_2=1}^{m_2} (C_{l_2,2,1} \cap (\bigcap_{j=3}^{n_1} \bigcup_{l=1}^{m_j} C_{l,j,1})))) \end{aligned}$$

$$\begin{aligned}
&= \bigcup_{l_1=1}^{m_1} \bigcup_{l_2=1}^{m_2} (C_{l_1,1,1} \cap C_{l_2,2,1} \cap (\bigcap_{j=3}^{n_1} \bigcup_{l=1}^{m_j} C_{l,j,1})) \\
&= \bigcup_{l_1=1}^{m_1} \cdots \bigcup_{l_{n_1}=1}^{m_{n_1}} \bigcap_{j=1}^{n_1} C_{l_j,j,1}
\end{aligned}$$

When $n_1 = 0$, clearly $\widetilde{A}_1 = X \in \mathcal{A}$. When $n_1 \geq 1$, $\bigcap_{j=1}^{n_1} C_{l_j,j,1} \in \mathcal{C}$ for any admissible choice of indices $l_1, \dots, l_{n_1}$. For different admissible choices of indices $l_1, \dots, l_{n_1}$ and $\bar{l}_1, \dots, \bar{l}_{n_1}$, $(\bigcap_{j=1}^{n_1} C_{l_j,j,1}) \cap (\bigcap_{j=1}^{n_1} C_{\bar{l}_j,j,1}) = \emptyset$. Hence, $\widetilde{A}_1 \in \mathcal{A}$. Note that $A_1 \cap A_2 = \bigcup_{j_1=1}^{m_1} \bigcup_{j_2=1}^{m_2} (C_{j_1,1} \cap C_{j_2,2})$. $C_{j_1,1} \cap C_{j_2,2} \in \mathcal{C}$ for any admissible choice of indices j_1, j_2 . For different admissible choices of indices j_1, j_2 and $\bar{j}_1, \bar{j}_2$, $(C_{j_1,1} \cap C_{j_2,2}) \cap (C_{\bar{j}_1,1} \cap C_{\bar{j}_2,2}) = \emptyset$. Hence, $A_1 \cap A_2 \in \mathcal{A}$.

Therefore, $\mathcal{A}$ is an algebra on X . Clearly, $\mathcal{A}$ is the smallest algebra on X containing $\mathcal{C}$. Hence, $\mathcal{A}$ is the algebra on X generated by $\mathcal{C}$. This completes the proof of the proposition. $\square$

Proposition 11.32 *Let X be a set, $\mathcal{C}$ be a semialgebra on X , and $\mu : \mathcal{C} \rightarrow [0, +\infty] \subset \mathbb{R}_e$. Assume that*

- (i) $\forall C \in \mathcal{C}, \forall n \in \mathbb{Z}_+, \forall$ pairwise disjoint $(C_i)_{i=1}^n \subseteq \mathcal{C}$ with $C = \bigcup_{i=1}^n C_i$, then $\mu(C) = \sum_{i=1}^n \mu(C_i)$;
- (ii) $\forall C \in \mathcal{C}, \forall$ pairwise disjoint $(C_i)_{i=1}^\infty \subseteq \mathcal{C}$ with $C = \bigcup_{i=1}^\infty C_i$, then $\mu(C) \leq \sum_{i=1}^\infty \mu(C_i)$.

Then, μ admits a unique extension to a measure $\bar{\mu}$ on the algebra on X , $\mathcal{A}$, generated by $\mathcal{C}$.

Proof By Proposition 11.31, $\mathcal{A} = \{A \subseteq X \mid A \text{ is the finite disjoint union of sets in } \mathcal{C}\}$. Define $\bar{\mu} : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $\bar{\mu}(A) = \sum_{i=1}^n \mu(C_i)$, $\forall A = \bigcup_{i=1}^n C_i \in \mathcal{A}$ where $n \in \mathbb{Z}_+$ and $(C_i)_{i=1}^n \subseteq \mathcal{C}$ is pairwise disjoint. By (i), $\bar{\mu}$ is well-defined. Clearly $\bar{\mu}(\emptyset) = 0$. $\forall A_1, A_2 \in \mathcal{A}$ with $A_1 \cap A_2 = \emptyset$, $A_l = \bigcup_{i=1}^{n_l} C_{i,l}$, where $n_l \in \mathbb{Z}_+$ and $(C_{i,l})_{i=1}^{n_l} \subseteq \mathcal{C}$ is pairwise disjoint, $l = 1, 2$. $A_1 \cup A_2 = \bigcup_{l=1}^2 \bigcup_{i=1}^{n_l} C_{i,l}$ where $C_{i,l}$, $l = 1, 2$, $i = 1, \dots, n_l$ are pairwise disjoint. Then, $\bar{\mu}(A_1 \cup A_2) = \sum_{l=1}^2 \sum_{i=1}^{n_l} \mu(C_{i,l}) = \bar{\mu}(A_1) + \bar{\mu}(A_2)$. Hence, $\bar{\mu}$ is finitely additive.

$\forall (A_i)_{i=1}^\infty \subseteq \mathcal{A}$ with $A = \bigcup_{i=1}^\infty A_i \in \mathcal{A}$ and $A_1, A_2, \dots$ being pairwise disjoint. By finite additivity $\bar{\mu}(A) = \bar{\mu}((\bigcup_{i=1}^n A_i) \cup (A \setminus \bigcup_{i=1}^n A_i)) \geq \sum_{i=1}^n \bar{\mu}(A_i)$, $\forall n \in \mathbb{N}$. Then, $\bar{\mu}(A) \geq \sum_{i=1}^\infty \bar{\mu}(A_i)$. On the other hand, since $A \in \mathcal{A}$, then $A = \bigcup_{i=1}^{\bar{n}} C_i$, where $\bar{n} \in \mathbb{Z}_+$ and $(C_i)_{i=1}^{\bar{n}} \subseteq \mathcal{C}$ is pairwise disjoint. $\forall i \in \mathbb{N}$, $A_i = \bigcup_{j=1}^{n_i} C_{j,i}$, where $n_i \in \mathbb{Z}_+$ and $(C_{j,i})_{j=1}^{n_i} \subseteq \mathcal{C}$ is pairwise disjoint. Then, $\forall i \in \{1, \dots, \bar{n}\}$, $C_i = C_i \cap (\bigcup_{l=1}^\infty A_l) = \bigcup_{l=1}^\infty \bigcup_{j=1}^{n_l} (C_{j,l} \cap C_i)$. By (ii), we have $\mu(C_i) \leq \sum_{l=1}^\infty \sum_{j=1}^{n_l} \mu(C_{j,l} \cap C_i)$. Therefore, $\bar{\mu}(A) = \sum_{i=1}^{\bar{n}} \mu(C_i) \leq \sum_{i=1}^{\bar{n}} \sum_{l=1}^\infty \sum_{j=1}^{n_l} \mu(C_{j,l} \cap C_i) = \sum_{l=1}^\infty \sum_{j=1}^{n_l} \sum_{i=1}^{\bar{n}} \mu(C_{j,l} \cap C_i)$

$C_i) = \sum_{l=1}^{\infty} \sum_{j=1}^{n_l} \mu(C_{j,l}) = \sum_{l=1}^{\infty} \bar{\mu}(A_l)$, where the third equality follows from (i). Hence, $\bar{\mu}(A) = \sum_{l=1}^{\infty} \bar{\mu}(A_l)$.

This implies that $\bar{\mu}$ is a measure on algebra $\mathcal{A}$. Clearly, it is an extension of μ to $\mathcal{A}$ and it is the unique extension. This completes the proof of the proposition. $\square$

11.3 Measurable Functions

Definition 11.33 Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be a topological space, $D \in \mathcal{B}$, and $f : D \rightarrow \mathcal{Y}$. We will say that f is $\mathcal{B}$ -measurable, if $\forall O \in \mathcal{O}_Y$, $f_{\text{inv}}(O) \in \mathcal{B}$. When it is clear from the context, we will simply say that f is measurable.

Proposition 11.34 Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be a topological space, $D \in \mathcal{B}$, and $f : D \rightarrow \mathcal{Y}$. Then,

- (i) f is $\mathcal{B}$ -measurable if, and only if, $\forall E \in \mathcal{B}_B(\mathcal{Y})$, $f_{\text{inv}}(E) \in \mathcal{B}$;
- (ii) if $\mathcal{Y}$ is second countable with a countable basis $\bar{\mathcal{B}}$, then f is $\mathcal{B}$ -measurable if, and only if, $\forall B_Y \in \bar{\mathcal{B}}$, $f_{\text{inv}}(B_Y) \in \mathcal{B}$.

Proof (i) “Sufficiency” is obvious. “Necessity” Let f be $\mathcal{B}$ -measurable. Define $\mathcal{B}_Y := \{E \subseteq Y \mid f_{\text{inv}}(E) \in \mathcal{B}\}$. Clearly, $\mathcal{O}_Y \subseteq \mathcal{B}_Y$. We will show that $\mathcal{B}_Y$ is a σ -algebra on Y . Then, $\mathcal{B}_B(\mathcal{Y}) \subseteq \mathcal{B}_Y$ and the result is established.

Clearly, $f_{\text{inv}}(\emptyset) = \emptyset \in \mathcal{B}$ and $f_{\text{inv}}(Y) = D \in \mathcal{B}$, then $\emptyset, Y \in \mathcal{B}_Y$. $\forall E \in \mathcal{B}_Y$, by Proposition 2.5, $f_{\text{inv}}(Y \setminus E) = f_{\text{inv}}(Y) \setminus f_{\text{inv}}(E) = D \setminus f_{\text{inv}}(E) \in \mathcal{B}$. Hence, $Y \setminus E \in \mathcal{B}_Y$. $\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{B}_Y$, we have $f_{\text{inv}}(E_i) \in \mathcal{B}$, $\forall i \in \mathbb{N}$. Then, by Proposition 2.5, $f_{\text{inv}}(\bigcup_{i=1}^{\infty} E_i) = \bigcup_{i=1}^{\infty} f_{\text{inv}}(E_i) \in \mathcal{B}$. Then, $\bigcup_{i=1}^{\infty} E_i \in \mathcal{B}_Y$. This shows $\mathcal{B}_Y$ is a σ -algebra on Y .

(ii) “Necessity” is obvious. “Sufficiency” $\forall O \in \mathcal{O}_Y$, since $\mathcal{Y}$ is second countable, then $\exists (B_i)_{i=1}^{\infty} \subseteq \bar{\mathcal{B}}$ such that $O = \bigcup_{i=1}^{\infty} B_i$. This implies that $f_{\text{inv}}(O) = \bigcup_{i=1}^{\infty} f_{\text{inv}}(B_i) \in \mathcal{B}$, where the equality follows from Proposition 2.5. Hence, f is $\mathcal{B}$ -measurable.

This completes the proof of the proposition. $\square$

Proposition 11.35 Let $(X, \mathcal{B})$ be a measurable space, $D \in \mathcal{B}$, $f : D \rightarrow \mathbb{R}_e$. Then, the following statements are equivalent.

- (i) f is $\mathcal{B}$ -measurable.
- (ii) $\forall \alpha \in \mathbb{R}$, the set $\{x \in D \mid f(x) < \alpha\} \in \mathcal{B}$.
- (iii) $\forall \alpha \in \mathbb{R}$, the set $\{x \in D \mid f(x) \geq \alpha\} \in \mathcal{B}$.
- (iv) $\forall \alpha \in \mathbb{R}$, the set $\{x \in D \mid f(x) \leq \alpha\} \in \mathcal{B}$.
- (v) $\forall \alpha \in \mathbb{R}$, the set $\{x \in D \mid f(x) > \alpha\} \in \mathcal{B}$.

These statements imply

(vi) $\forall \alpha \in \mathbb{R}_e$, the set $\{x \in D \mid f(x) = \alpha\} \in \mathcal{B}$.

Proof The proof is straightforward, and is therefore omitted. $\square$

Proposition 11.36 Let $\mathcal{X} := (X, \mathcal{B})$ be a measurable space, $D \in \mathcal{B}$, $f_i : D \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable, $i = 1, 2$, and $c \in (0, \infty) \subset \mathbb{R}$. Then, cf_1 and $f_1 + f_2$ are $\mathcal{B}$ -measurable.

Proof $\forall \alpha \in \mathbb{R}$, $\{x \in D \mid cf_1(x) < \alpha\} = \{x \in D \mid f_1(x) < \alpha/c\} \in \mathcal{B}$ by Proposition 11.35. Then, by Proposition 11.35, cf_1 is $\mathcal{B}$ -measurable. $\{x \in D \mid f_1(x) + f_2(x) < \alpha\} = \bigcup_{r \in \mathbb{Q}} (\{x \in D \mid f_1(x) < r\} \cap \{x \in D \mid f_2(x) < \alpha - r\}) \in \mathcal{B}$ by Proposition 11.35. Then, by Proposition 11.35, $f_1 + f_2$ is $\mathcal{B}$ -measurable. This completes the proof of the proposition. $\square$

Proposition 11.37 Let $\mathbb{X} := (X, \mathcal{B}, \mu)$ be a topological measure space, $\mathcal{Y}$ be a topological space, $D \in \mathcal{B}$, and $f : D \rightarrow \mathcal{Y}$ be continuous. Then, f is $\mathcal{B}$ -measurable.

Proof The proof is straightforward, and is therefore omitted. $\square$

Proposition 11.38 Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ and $\mathcal{Z}$ be topological spaces, $D \in \mathcal{B}$, $f : D \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, and $g : \mathcal{Y} \rightarrow \mathcal{Z}$ be continuous. Then, $h := g \circ f : D \rightarrow \mathcal{Z}$ is $\mathcal{B}$ -measurable.

Proof $\forall O \in \mathcal{O}_{\mathcal{Z}}$, $g_{\text{inv}}(O) \in \mathcal{O}_{\mathcal{Y}}$ since g is continuous. $h_{\text{inv}}(O) = f_{\text{inv}}(g_{\text{inv}}(O)) \in \mathcal{B}$. Hence, h is $\mathcal{B}$ -measurable. $\square$

Proposition 11.39 Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}_i := (Y_i, \mathcal{O}_i)$ be a second countable topological space, $D \in \mathcal{B}$, $f_i : D \rightarrow \mathcal{Y}_i$, $\forall i \in N \subseteq \mathbb{N}$, N is a countable index set, $\mathcal{Y} := (Y, \mathcal{O}) := \prod_{i \in N} \mathcal{Y}_i$ be the product topological space, and $f : D \rightarrow \mathcal{Y}$ be given by $\pi_i(f(x)) = f_i(x)$, $\forall x \in D$, $\forall i \in N$. Then, f is $\mathcal{B}$ -measurable if, and only if, f_i is $\mathcal{B}$ -measurable, $\forall i \in N$.

Proof “Sufficiency” Let f_i ’s be $\mathcal{B}$ -measurable. $\forall i \in N$, let $\mathcal{B}_{Y_i}$ be a countable basis for $\mathcal{Y}_i$. Without loss of generality, assume $\emptyset, Y_i \in \mathcal{B}_{Y_i}$. Then, $\mathcal{Y}$ is second countable with a countable basis $\mathcal{B}_Y$ as defined in Proposition 3.28. $\forall O \in \mathcal{O}$, $O = \bigcup_{l=1}^{\infty} O_{Bl}$ where $O_{Bl} \in \mathcal{B}_Y$, $\forall l \in \mathbb{N}$. $\forall l \in \mathbb{N}$, $O_{Bl} = \prod_{i \in N} B_{li}$, where $B_{li} \in \mathcal{B}_{Y_i}$, $\forall i \in N$, $B_{li} = Y_i$, $\forall i \in N \setminus N_l$, and $N_l \subseteq N$ is a finite set. Then, $f_{\text{inv}}(O_{Bl}) = \bigcap_{i \in N} f_{i \text{ inv}}(B_{li}) \in \mathcal{B}$ since f_i ’s are $\mathcal{B}$ -measurable and $\mathcal{B}$ is a σ -algebra. Then, $f_{\text{inv}}(O) = f_{\text{inv}}(\bigcup_{l=1}^{\infty} O_{Bl}) = \bigcup_{l=1}^{\infty} f_{\text{inv}}(O_{Bl}) \in \mathcal{B}$. Hence, f is $\mathcal{B}$ -measurable.

“Necessity” Let f be $\mathcal{B}$ -measurable. $\forall i_0 \in N$, $\forall O_{i_0} \in \mathcal{O}_{i_0}$, let $O := \prod_{i \in N} O_i \in \mathcal{O}$, where $O_i = Y_i$, $\forall i \in N$ with $i \neq i_0$. By the measurability of f , we have $f_{i_0 \text{ inv}}(O_{i_0}) = f_{\text{inv}}(O) \in \mathcal{B}$. Hence, f_{i_0} is $\mathcal{B}$ -measurable.

This completes the proof of the proposition. $\square$

Proposition 11.40 *Let $(X, \mathcal{B})$ be a measurable space, $D \in \mathcal{B}$, $f_n : D \rightarrow \mathbb{R}_e$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $\bar{f}_S := \limsup_{n \in \mathbb{N}} f_n : D \rightarrow \mathbb{R}_e$ defined by $\bar{f}_S(x) = \limsup_{n \in \mathbb{N}} f_n(x)$, $\forall x \in D$, and $\bar{f}_I := \liminf_{n \in \mathbb{N}} f_n : D \rightarrow \mathbb{R}_e$ defined by $\bar{f}_I(x) = \liminf_{n \in \mathbb{N}} f_n(x)$, $\forall x \in D$. Then, $\forall n \in \mathbb{Z}_+$, $\bigvee_{i=1}^n f_i$, $\bigwedge_{i=1}^n f_i$, $\sup_{i \in \mathbb{N}} f_i$, $\inf_{i \in \mathbb{N}} f_i$, $\limsup_{n \in \mathbb{N}} f_n = \bar{f}_S$, and $\liminf_{n \in \mathbb{N}} f_n = \bar{f}_I$ are $\mathcal{B}$ -measurable.*

Proof $\forall \alpha \in \mathbb{R}$, $\{x \in D \mid \bigvee_{i=1}^n f_i(x) > \alpha\} = \bigcup_{i=1}^n \{x \in D \mid f_i(x) > \alpha\} \in \mathcal{B}$. Hence, by Proposition 11.35, $\bigvee_{i=1}^n f_i$ is $\mathcal{B}$ -measurable.

$\forall \alpha \in \mathbb{R}$, $\{x \in D \mid \sup_{i \in \mathbb{N}} f_i(x) > \alpha\} = \bigcup_{i=1}^{\infty} \{x \in D \mid f_i(x) > \alpha\} \in \mathcal{B}$. Hence, by Proposition 11.35, $\sup_{i \in \mathbb{N}} f_i$ is $\mathcal{B}$ -measurable.

By arguments that are similar to the above, $\bigwedge_{i=1}^n f_i$ and $\inf_{i \in \mathbb{N}} f_i$ are $\mathcal{B}$ -measurable.

By Definition 3.82, $\bar{f}_S = \inf_{i \in \mathbb{N}} \sup_{j \in \mathbb{N}, j \geq i} f_j$ and $\bar{f}_I = \sup_{i \in \mathbb{N}} \inf_{j \in \mathbb{N}, j \geq i} f_j$. Then, $\bar{f}_S$ and $\bar{f}_I$ are $\mathcal{B}$ -measurable. This completes the proof of the proposition. $\square$

Proposition 11.41 *Let $(X, \mathcal{B})$ be a measurable space, $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$ be such that $\bigcup_{i=1}^{\infty} E_i = X$, $\mathcal{Y} := (Y, \mathcal{O})$ be a topological space, $D \in \mathcal{B}$, $\mathcal{B}_{E_i} := \{B \in \mathcal{B} \mid B \subseteq E_i\}$, $\forall i \in \mathbb{N}$, and $f : D \rightarrow \mathcal{Y}$. Then, $f|_{E_i}$ is $\mathcal{B}_{E_i}$ -measurable, $\forall i \in \mathbb{N}$, if, and only if, $f|_{E_i}$ is $\mathcal{B}$ -measurable, $\forall i \in \mathbb{N}$, if, and only if, f is $\mathcal{B}$ -measurable.*

Proof (i) $\Rightarrow$ (ii). Assume that $f|_{E_i}$ is $\mathcal{B}_{E_i}$ -measurable, $\forall i \in \mathbb{N}$. $\forall O \in \mathcal{O}$, by the assumption, we have $(f|_{E_i})_{\text{inv}}(O) \in \mathcal{B}_{E_i} \subseteq \mathcal{B}$, $\forall i \in \mathbb{N}$. Then, $f|_{E_i}$ is $\mathcal{B}$ -measurable.

(ii) $\Rightarrow$ (iii). Assume that $f|_{E_i}$ is $\mathcal{B}$ -measurable, $\forall i \in \mathbb{N}$. $\forall O \in \mathcal{O}$, by the assumption, we have $(f|_{E_i})_{\text{inv}}(O) \in \mathcal{B}$, $\forall i \in \mathbb{N}$. Then, $f_{\text{inv}}(O) = \bigcup_{i=1}^{\infty} (f|_{E_i})_{\text{inv}}(O) \in \mathcal{B}$. Hence, f is $\mathcal{B}$ -measurable.

(iii) $\Rightarrow$ (i). Assume that f is $\mathcal{B}$ -measurable. $\forall O \in \mathcal{O}$, by the assumption, $f_{\text{inv}}(O) \in \mathcal{B}$. Then, $(f|_{E_i})_{\text{inv}}(O) = f_{\text{inv}}(O) \cap E_i \in \mathcal{B}_{E_i}$, $\forall i \in \mathbb{N}$. Hence, $f|_{E_i}$ is $\mathcal{B}_{E_i}$ -measurable, $\forall i \in \mathbb{N}$. This completes the proof of the proposition. $\square$

Definition 11.42 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space. A property P is said to hold almost everywhere in $\mathcal{X}$ (abbreviated a.e.) if the set of points where it fails to hold or does not make sense belongs to $\mathcal{B}$ and have measure 0. We will write P a.e. in $\mathcal{X}$ or $P(x)$ a.e. $x \in \mathcal{X}$.*

Lemma 11.43 *Let $\mathcal{X} := (X, \mathcal{B})$ be a measurable space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_1, D_2 \in \mathcal{B}$, and $f : D_1 \rightarrow \mathcal{Y}$ and $g : D_2 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Then, the function $h : D_1 \cap D_2 \rightarrow \mathbb{R}$, defined by $h(x) = \rho(f(x), g(x))$, $\forall x \in D_1 \cap D_2$, is $\mathcal{B}$ -measurable.*

Proof By Proposition 4.4, $\mathcal{Y}$ is second countable. By Propositions 11.39 and 11.41, the function $h_1 : D_1 \cap D_2 \rightarrow \mathcal{Y} \times \mathcal{Y}$, defined by $h_1(x) = (f(x), g(x))$, $\forall x \in D_1 \cap D_2$, is $\mathcal{B}$ -measurable. By Propositions 11.38 and 4.30, the function h is $\mathcal{B}$ -measurable. $\square$

Lemma 11.44 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_i \in \mathcal{B}$, and $f_i : D_i \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $i = 1, 2, 3$. Assume that $f_1 = f_2$ a.e. in $\mathcal{X}$ and $f_2 = f_3$ a.e. in $\mathcal{X}$. Then, $f_1 = f_3$ a.e. in $\mathcal{X}$.

Proof Let $E_1 := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid f_1(x) \neq f_2(x)\}$, $E_2 := \widetilde{D_2} \cup \widetilde{D_3} \cup \{x \in D_2 \cap D_3 \mid f_2(x) \neq f_3(x)\}$, and $E_3 := \widetilde{D_1} \cup \widetilde{D_3} \cup \{x \in D_1 \cap D_3 \mid f_3(x) \neq f_1(x)\}$. By the assumption, we have $E_1, E_2 \in \mathcal{B}$ and $\mu(E_1) = \mu(E_2) = 0$. Clearly, $E_3 \subseteq E_1 \cup E_2$. Note that $E_3 = \widetilde{D_1} \cup \widetilde{D_3} \cup \{x \in D_1 \cap D_3 \mid \rho(f_1(x), f_3(x)) > 0\} \in \mathcal{B}$ by Lemma 11.43. Then, we have $0 \leq \mu(E_3) \leq \mu(E_1) + \mu(E_2) = 0$. Hence, $f_1 = f_3$ a.e. in $\mathcal{X}$. This completes the proof of the lemma. $\square$

Lemma 11.45 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}_i := (Y_i, \mathcal{O}_i)$ be a second countable metrizable topological space, $\forall i \in N \subseteq \mathbb{N}$, where N is a countable index set, $\mathcal{Y} := \prod_{i \in N} \mathcal{Y}_i$ be the product space, $D_1, D_2 \in \mathcal{B}$, $f_i : D_1 \rightarrow \mathcal{Y}_i$ and $g_i : D_2 \rightarrow \mathcal{Y}_i$ be $\mathcal{B}$ -measurable, $\forall i \in N$, $f : D_1 \rightarrow \mathcal{Y}$ be defined by $\pi_i(f(x)) = f_i(x)$, $\forall i \in N$, $\forall x \in D_1$, and $g : D_2 \rightarrow \mathcal{Y}$ be defined by $\pi_i(g(x)) = g_i(x)$, $\forall i \in N$, $\forall x \in D_2$. Then, $f = g$ a.e. in $\mathcal{X}$ if, and only if, $f_i = g_i$ a.e. in $\mathcal{X}$, $\forall i \in N$.

Proof Define $E_i := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid f_i(x) \neq g_i(x)\}$, $\forall i \in N$, and $E := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid f(x) \neq g(x)\}$.

“Sufficiency” Let $f_i = g_i$ a.e. in $\mathcal{X}$, $\forall i \in N$. Then, $E_i \in \mathcal{B}$ and $\mu(E_i) = 0$, $\forall i \in N$. Note that $E = \bigcup_{i \in N} E_i \in \mathcal{B}$ and $\mu(E) = 0$. Hence, $f = g$ a.e. in $\mathcal{X}$.

“Necessity” Let $f = g$ a.e. in $\mathcal{X}$. Then, $E \in \mathcal{B}$ and $\mu(E) = 0$. $\forall i \in N$, $E_i \subseteq E$. Since $\mathcal{Y}_i$ is a second countable metrizable topological space, let $\rho_i : Y_i \times Y_i \rightarrow [0, \infty) \subset \mathbb{R}$ be the metric on $\mathcal{Y}_i$ whose natural topology is $\mathcal{O}_i$. Then, (Y_i, ρ_i) is a separable metric space, by Proposition 4.4. By Lemma 11.43, we have $E_i = \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid \rho_i(f_i(x), g_i(x)) > 0\} \in \mathcal{B}$. Then, $0 \leq \mu(E_i) \leq \mu(E) = 0$. Hence, $f_i = g_i$ a.e. in $\mathcal{X}$.

This completes the proof of the lemma. $\square$

Lemma 11.46 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a topological space, $\mathcal{Z} := (Z, \rho)$ be a separable metric space, $D_1, D_2 \in \mathcal{B}$, $f_i : D_i \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $i = 1, 2$, $g : \mathcal{Y} \rightarrow \mathcal{Z}$ be continuous. Assume that $f_1 = f_2$ a.e. in $\mathcal{X}$. Then, $g \circ f_1 = g \circ f_2$ a.e. in $\mathcal{X}$.

Proof By the assumption, $E := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid f_1(x) \neq f_2(x)\} \in \mathcal{B}$ and $\mu(E) = 0$. By Proposition 11.38, we have $g \circ f_i$ is $\mathcal{B}$ -measurable, $i = 1, 2$. By Lemma 11.43, $E \supseteq \bar{E} := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid g(f_1(x)) \neq g(f_2(x))\} = \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid \rho(g(f_1(x)), g(f_2(x))) > 0\} \in \mathcal{B}$. Then, $0 \leq \mu(\bar{E}) \leq \mu(E) = 0$. Hence, $g \circ f_1 = g \circ f_2$ a.e. in $\mathcal{X}$. This completes the proof of the lemma. $\square$

Proposition 11.47 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a complete measure space, $\mathcal{Y}$ be a topological space, $D_1, D_2 \in \mathcal{X}^2$, $f : D_1 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $g : D_2 \rightarrow \mathcal{Y}$, and $g = f$ a.e. in $\mathcal{X}$. Then, g is $\mathcal{B}$ -measurable.

Proof Let $E := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid f(x) \neq g(x)\}$. Then, $E \in \mathcal{B}$ and $\mu(E) = 0$. $\forall O \in \mathcal{O}_{\mathcal{Y}}$, $g_{\text{inv}}(O) = g_{\text{inv}}(O) \cap (E \cup (X \setminus E)) = (g_{\text{inv}}(O) \cap E) \cup (g_{\text{inv}}(O) \cap (X \setminus E)) = (g_{\text{inv}}(O) \cap E) \cup (f_{\text{inv}}(O) \cap (X \setminus E))$. Note that $X \setminus E, f_{\text{inv}}(O) \in \mathcal{B}$. By the completeness of $\mathcal{X}$, $E \supseteq g_{\text{inv}}(O) \cap E \in \mathcal{B}$. Then, $g_{\text{inv}}(O) \in \mathcal{B}$. Hence, g is $\mathcal{B}$ -measurable. $\square$

Proposition 11.48 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a metric space, $D \in \mathcal{B}$, $f_n : D \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $f : D \rightarrow \mathcal{Y}$. Assume that $\lim_{n \in \mathbb{N}} f_n(x) = f(x)$, $\forall x \in D$. Then, f is $\mathcal{B}$ -measurable.

Proof Fix any open set $O \subseteq \mathcal{Y}$. $\forall n \in \mathbb{N}$, define $O_n := \{y \in O \mid \mathcal{B}_{\mathcal{Y}}(y, 1/n) \subseteq O\}$. Then, $O_n = O \cap \left(\bigcap_{y \in \mathcal{Y} \setminus O} (\mathcal{Y} \setminus \mathcal{B}_{\mathcal{Y}}(y, 1/n)) \right)$. Note that $\bigcap_{y \in \mathcal{Y} \setminus O} (\mathcal{Y} \setminus \mathcal{B}_{\mathcal{Y}}(y, 1/n))$ is a closed set in $\mathcal{Y}$. Then, $O_n \in \mathcal{B}_{\mathcal{Y}}(\mathcal{Y})$.

Claim 11.48.1 $f_{\text{inv}}(O) = \bigcup_{l=1}^{\infty} \bigcap_{k=l}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} f_{j\text{inv}}(O_k) =: U$.

Proof of claim: $\forall x \in f_{\text{inv}}(O) \subseteq D$, then $y := f(x) \in O$. Since O is open, then $\exists L \in \mathbb{N}$ such that $\mathcal{B}_{\mathcal{Y}}(y, 1/L) \subseteq O$. Then, $\forall l \geq L+1$, $\forall \bar{y} \in \mathcal{B}_{\mathcal{Y}}\left(y, \frac{1}{L(L+1)}\right)$, $\forall \tilde{y} \in \mathcal{B}_{\mathcal{Y}}(\bar{y}, 1/l)$, we have $\rho(y, \tilde{y}) \leq \rho(y, \bar{y}) + \rho(\bar{y}, \tilde{y}) < \frac{1}{L(L+1)} + \frac{1}{l} \leq \frac{1}{L}$. This implies that $\tilde{y} \in \mathcal{B}_{\mathcal{Y}}(y, 1/L) \subseteq O$. Then, $\mathcal{B}_{\mathcal{Y}}(\bar{y}, 1/l) \subseteq O$. Thus, $\bar{y} \in O_l$ and $\mathcal{B}_{\mathcal{Y}}\left(y, \frac{1}{L(L+1)}\right) \subseteq O_l$. Since $\lim_{n \in \mathbb{N}} f_n(x) = f(x) = y$, then $\exists N \in \mathbb{N}$ such that $\forall n \geq N$, $\rho(f_n(x), y) < \frac{1}{L(L+1)}$. Then, $f_n(x) \in \mathcal{B}_{\mathcal{Y}}\left(y, \frac{1}{L(L+1)}\right) \subseteq O_l$ and $x \in f_{n\text{inv}}(O_l)$. By arbitrariness of n , $x \in \bigcap_{j=N}^{\infty} f_{j\text{inv}}(O_l) \subseteq \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} f_{j\text{inv}}(O_l)$. By arbitrariness of l , we have $x \in \bigcap_{k=L+1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} f_{j\text{inv}}(O_k) \subseteq U$. Hence, by the arbitrariness of x , $f_{\text{inv}}(O) \subseteq U$.

On the other hand, $\forall x \in U$, $\exists L \in \mathbb{N}$ such that $x \in \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} f_{j\text{inv}}(O_L)$. Then, $\exists N \in \mathbb{N}$ such that $\forall n \geq N$, $x \in f_{n\text{inv}}(O_L)$ and $f_n(x) \in O_L$. Since $\lim_{n \in \mathbb{N}} f_n(x) = f(x)$, then $\exists \bar{N} \in \mathbb{N}$ with $\bar{N} \geq N$ such that $\rho(f_{\bar{N}}(x), f(x)) < 1/L$. Then, $f(x) \in \mathcal{B}_{\mathcal{Y}}(f_{\bar{N}}(x), 1/L)$. Note that $f_{\bar{N}}(x) \in O_L$ and $\mathcal{B}_{\mathcal{Y}}(f_{\bar{N}}(x), 1/L) \subseteq O$. Then, $f(x) \in O$ and $x \in f_{\text{inv}}(O)$. By the arbitrariness of x , $U \subseteq f_{\text{inv}}(O)$.

Hence, $f_{\text{inv}}(O) = U$. This completes the proof of the claim. $\square$

$\forall k \in \mathbb{N}, \forall j \in \mathbb{N}$, by Proposition 11.34, $f_{j\text{inv}}(O_k) \in \mathcal{B}$. By Claim 11.48.1, $f_{\text{inv}}(O) \in \mathcal{B}$. Hence, f is $\mathcal{B}$ -measurable. This completes the proof of the proposition. $\square$

Proposition 11.49 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a complete measure space, $\mathcal{Y} := (Y, \rho)$ be a metric space, $D_1, D_2 \in \mathcal{X}^2$, $f_n : D_1 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable,

$\forall n \in \mathbb{N}$, and $f : D_2 \rightarrow \mathcal{Y}$. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$. Then, f is $\mathcal{B}$ -measurable.

Proof Let $E := \{x \in D_1 \cap D_2 \mid (f_n(x))_{n=1}^\infty \text{ does not converge to } f(x)\} \cup \widetilde{D_1} \cup \widetilde{D_2}$. By the assumption, $E \in \mathcal{B}$ and $\mu(E) = 0$. Let $g := f|_{X \setminus E}$ and $g_n := f_n|_{X \setminus E}$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} g_n(x) = g(x)$, $\forall x \in X \setminus E \in \mathcal{B}$. By Proposition 11.48, g is $\mathcal{B}$ -measurable. $\forall$ open set $O \subseteq \mathcal{Y}$, $f_{\text{inv}}(O) = g_{\text{inv}}(O) \cup (f_{\text{inv}}(O) \cap E)$. By the completeness of $\mathcal{X}$, we have $E \supseteq f_{\text{inv}}(O) \cap E \in \mathcal{B}$. Then, $f_{\text{inv}}(O) \in \mathcal{B}$. Hence, f is $\mathcal{B}$ -measurable. This completes the proof of the proposition. $\square$

Proposition 11.50 Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_1, D_2 \in \mathcal{B}$, $f_n : D_1 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $f : D_2 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Then, the set

$$E := \{x \in D_1 \cap D_2 \mid (f_n(x))_{n=1}^\infty \text{ converges to } f(x)\} \in \mathcal{B}$$

Proof We will show that

$$E = \bigcap_{m=1}^\infty \bigcup_{i=1}^\infty \bigcap_{k=i}^\infty E_{k,m} =: \bar{E}$$

where $E_{k,m} := \{x \in D_1 \cap D_2 \mid \rho(f_k(x), f(x)) < 1/m\} \in \mathcal{B}$, $\forall m \in \mathbb{N}$, $\forall k \in \mathbb{N}$, by Lemma 11.43. This further implies that $\bar{E} \in \mathcal{B}$.

$\forall x \in E$, $\forall m \in \mathbb{N}$, $\exists k_0 \in \mathbb{N}$, $\forall k \in \mathbb{N}$ with $k \geq k_0$, $x \in D_1 \cap D_2$ and $\rho(f_k(x), f(x)) < 1/m$. Then, $x \in E_{k,m}$ and $x \in \bigcap_{k=k_0}^\infty E_{k,m} \subseteq \bigcup_{i=1}^\infty \bigcap_{k=i}^\infty E_{k,m}$. By the arbitrariness of m , we have $x \in \bar{E}$. This implies that $E \subseteq \bar{E}$. On the other hand, $\forall x \in \bar{E}$, $\forall m \in \mathbb{N}$, $x \in \bigcup_{i=1}^\infty \bigcap_{k=i}^\infty E_{k,m}$. Then, $\exists k_0 \in \mathbb{N}$, such that $x \in \bigcap_{k=k_0}^\infty E_{k,m}$. $\forall k \in \mathbb{N}$ with $k \geq k_0$, $x \in E_{k,m}$. Then, $x \in D_1 \cap D_2$ and $\rho(f_k(x), f(x)) < 1/m$. This leads to $\lim_{n \in \mathbb{N}} f_n(x) = f(x)$ and $x \in E$. Then, we have $\bar{E} \subseteq E$. Hence, $E = \bar{E} \in \mathcal{B}$. This completes the proof of the proposition. $\square$

Proposition 11.51 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a separable complete metric space, $D \in \mathcal{B}$, $f_i : D \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall i \in \mathbb{N}$, $E := \{x \in D \mid (f_i(x))_{i=1}^\infty \text{ converges in } \mathcal{Y}\}$ and $f : E \rightarrow \mathcal{Y}$ defined by $f(x) = \lim_{i \in \mathbb{N}} f_i(x)$, $\forall x \in E$. Then, $E = \bigcap_{j=1}^\infty \bigcup_{m=1}^\infty \bigcap_{l=m}^\infty \bigcap_{n=m}^\infty E_{j,l,n} =: \bar{E} \in \mathcal{B}$, where $E_{j,l,n} := \{x \in D \mid \rho(f_l(x), f_n(x)) < 1/j\}$, $\forall j, l, n \in \mathbb{N}$, and f is $\mathcal{B}$ -measurable.

Proof $\forall j, l, n \in \mathbb{N}$, by Lemma 11.43 and the separability of $\mathcal{Y}$, $E_{j,l,n} \in \mathcal{B}$. Then, $\bar{E} \in \mathcal{B}$.

$\forall x \in \bar{E}$, $\forall j \in \mathbb{N}$, $\exists m_0 \in \mathbb{N}$, $\forall l, n \in \mathbb{N}$ with $l \geq m_0$ and $n \geq m_0$, we have $x \in D$ and $\rho(f_l(x), f_n(x)) < 1/j$. Then, $(f_i(x))_{i=1}^\infty \subseteq \mathcal{Y}$ is a Cauchy sequence. By the completeness of $\mathcal{Y}$, $(f_i(x))_{i=1}^\infty$ converges in $\mathcal{Y}$. Hence, $x \in E$. By the arbitrariness of x , we have $\bar{E} \subseteq E$.

On the other hand, $\forall x \in E$, since $(f_i(x))_{i=1}^\infty$ converges in $\mathcal{Y}$, then it is a Cauchy sequence. $\forall j \in \mathbb{N}$, $\exists m_0 \in \mathbb{N}$, $\forall l, n \in \mathbb{N}$ with $l \geq m_0$ and $n \geq m_0$, we have $\rho(f_l(x), f_n(x)) < 1/j$. Then, $x \in E_{j,l,n}$. Then, $x \in \bar{E}$. By the arbitrariness of x , we have $E \subseteq \bar{E}$. Therefore, $E = \bar{E} \in \mathcal{B}$.

Note that $f(x) = \lim_{i \in \mathbb{N}} f_i|_E(x)$, $\forall x \in E$. By Proposition 11.41, $f_i|_E$ is $\mathcal{B}$ -measurable, $\forall i \in \mathbb{N}$. By Proposition 11.48, f is $\mathcal{B}$ -measurable.

This completes the proof of the proposition. $\square$

Proposition 11.52 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \mathcal{O})$ be a Hausdorff topological space, $\mathcal{Z} := (Z, \rho)$ be a separable metric space, $D_1, D_2 \in \mathcal{B}$, $f_n : D_1 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $f : D_2 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, and $g : \mathcal{Y} \rightarrow \mathcal{Z}$ be continuous. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$. Then, $\lim_{n \in \mathbb{N}} g \circ f_n = g \circ f$ a.e. in $\mathcal{X}$.*

Proof Let $E := \{x \in D_1 \cap D_2 \mid \lim_{n \in \mathbb{N}} f_n(x) = f(x)\}$ and $\bar{E} := \{x \in D_1 \cap D_2 \mid \lim_{n \in \mathbb{N}} g(f_n(x)) = g(f(x))\}$. By Proposition 3.66, we have $E \subseteq \bar{E}$. By the assumption, $E \in \mathcal{B}$ and $\mu(\bar{E}) = 0$. By Proposition 11.38, $g \circ f_n$ and $g \circ f$ are $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. By Proposition 11.50, $\bar{E} \in \mathcal{B}$. Then, $0 \leq \mu(\bar{E}) \leq \mu(\bar{E}) = 0$. Hence, $\lim_{n \in \mathbb{N}} g \circ f_n = g \circ f$ a.e. in $\mathcal{X}$. This completes the proof of the proposition. $\square$

Proposition 11.53 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $D_1, D_2 \in \mathcal{B}$, $\mathcal{Y}_i := (Y_i, \mathcal{O}_i)$ be a second countable metrizable topological space, $f_i : D_1 \rightarrow \mathcal{Y}_i$ be $\mathcal{B}$ -measurable, $g_{n,i} : D_2 \rightarrow \mathcal{Y}_i$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $\forall i \in N \subseteq \mathbb{N}$, N is a countable index set, $\mathcal{Y} := (Y, \mathcal{O}) := \prod_{i \in N} \mathcal{Y}_i$ be the product topological space, $f : D_1 \rightarrow \mathcal{Y}$ be given by $\pi_i(f(x)) = f_i(x)$, $\forall x \in D_1$, $\forall i \in N$, and $g_n : D_2 \rightarrow \mathcal{Y}$ be given by $\pi_i(g_n(x)) = g_{n,i}(x)$, $\forall x \in D_2$, $\forall i \in N$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} g_n = f$ a.e. in $\mathcal{X}$ if, and only if, $\lim_{n \in \mathbb{N}} g_{n,i} = f_i$ a.e. in $\mathcal{X}$, $\forall i \in N$.*

Proof By Proposition 11.39, f and g_n are $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. “Sufficiency” Assume $\lim_{n \in \mathbb{N}} g_{n,i} = f_i$ a.e. in $\mathcal{X}$, $\forall i \in N$. $\forall i \in N$, let $E_i := \widetilde{D}_1 \cup \widetilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid (g_{n,i}(x))_{n=1}^\infty \text{ does not converge to } f_i(x)\}$. Then, $E_i \in \mathcal{B}$ and $\mu(E_i) = 0$. Let $E := \widetilde{D}_1 \cup \widetilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid (g_n(x))_{n=1}^\infty \text{ does not converge to } f(x)\}$. By Proposition 3.67, $E = \bigcup_{i \in N} E_i \in \mathcal{B}$. Then, $0 \leq \mu(E) \leq \sum_{i \in N} \mu(E_i) = 0$. Hence, $\lim_{n \in \mathbb{N}} g_n = f$ a.e. in $\mathcal{X}$. “Necessity” Assume $\lim_{n \in \mathbb{N}} g_n = f$ a.e. in $\mathcal{X}$. Let E and E_i be as defined above, $\forall i \in N$. Then, $E \in \mathcal{B}$ and $\mu(E) = 0$. $\forall i \in N$, by Proposition 3.67, we have $E_i \subseteq E$. Since $\mathcal{Y}_i$ is second countable metrizable topological space, let $\rho_i : Y_i \times Y_i \rightarrow [0, \infty) \subset \mathbb{R}$ be the metric on $\mathcal{Y}_i$ whose natural topology is $\mathcal{O}_i$. Then, by Proposition 4.4, (Y_i, ρ_i) is a separable metric sapce. By Proposition 11.50, we have $E_i \in \mathcal{B}$. Then, $0 \leq \mu(E_i) \leq \mu(E) = 0$. Hence, $\lim_{n \in \mathbb{N}} g_{n,i} = f_i$ a.e. in $\mathcal{X}$. This completes the proof of the proposition. $\square$

Proposition 11.54 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a separable metric space, $D_i \in \mathcal{B}$, $i = 1, 2, 3, 4$, $f_n : D_1 \rightarrow \mathcal{Y}$ and $g_n : D_2 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $f : D_3 \rightarrow \mathcal{Y}$ and $g : D_4 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$, $f_n = g_n$ a.e. in $\mathcal{X}$, $\forall i \in \mathbb{N}$, and $f = g$ a.e. in $\mathcal{X}$. Then, $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$.*

Proof By the assumption, we have $E_1 := \widetilde{D_1} \cup \widetilde{D_3} \cup \{x \in D_1 \cap D_3 \mid (f_n(x))_{n=1}^\infty \text{ does not converge to } f(x)\} \in \mathcal{B}$; $\bar{E}_n := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid f_n(x) \neq g_n(x)\} \in \mathcal{B}$, $\forall n \in \mathbb{N}$; $\bar{E} := \widetilde{D_3} \cup \widetilde{D_4} \cup \{x \in D_3 \cap D_4 \mid f(x) \neq g(x)\} \in \mathcal{B}$; and $\mu(E_1) = \mu(\bar{E}_n) = \mu(\bar{E}) = 0$, $\forall n \in \mathbb{N}$. Then, $E := E_1 \cup (\bigcup_{n=1}^\infty \bar{E}_n) \cup \bar{E} \in \mathcal{B}$ and $\mu(E) = 0$. Let $E_2 := \widetilde{D_2} \cup \widetilde{D_4} \cup \{x \in D_2 \cap D_4 \mid (g_n(x))_{n=1}^\infty \text{ does not converge to } g(x)\}$. By Proposition 11.50, $E_2 \in \mathcal{B}$. Clearly, $\bar{E} \subseteq \widetilde{E_2}$ and $E_2 \subseteq E$. Then, $0 \leq \mu(E_2) \leq \mu(E) = 0$. This shows that $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$. This completes the proof of the proposition. $\square$

Theorem 11.55 (Egoroff's Theorem) *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_1, D_2 \in \mathcal{B}$, $f_n : D_1 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $f : D_2 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$. Then, $\forall \eta \in (0, \infty) \subset \mathbb{R}$, $\exists A \in \mathcal{B}$ with $\mu(A) < \eta$ and $A \supseteq \widetilde{D_1} \cup \widetilde{D_2}$ such that $(f_n|_{X \setminus A})_{n=1}^\infty$ converges uniformly to $f|_{X \setminus A}$.*

Proof Let $E := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid (f_n(x))_{n=1}^\infty \text{ does not converge to } f(x)\}$. By the assumption, $E \in \mathcal{B}$ and $\mu(E) = 0$.

$\forall \eta \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall m \in \mathbb{N}$, let

$$G_{m,n} := \{x \in X \setminus E \mid \rho(f_m(x), f(x)) \geq 2^{-n}\}$$

By Lemma 11.43, $G_{m,n} \in \mathcal{B}$. Let

$$H_{m,n} := \bigcup_{j=m}^\infty G_{j,n} = \{x \in X \setminus E \mid \rho(f_i(x), f(x)) \geq 2^{-n} \text{ for some } i \geq m\}$$

Then, $H_{m+1,n} \subseteq H_{m,n} \in \mathcal{B}$. $\forall x \in X \setminus E$, we have $\lim_{j \in \mathbb{N}} f_j(x) = f(x)$, then $\exists m_x \in \mathbb{N}$ such that $x \notin H_{m_x,n}$. Hence, $\bigcap_{m=1}^\infty H_{m,n} = \emptyset$. Note that $\mu(H_{1,n}) \leq \mu(X) < \infty$. By Proposition 11.5, $\lim_{m \in \mathbb{N}} \mu(H_{m,n}) = \mu(\bigcap_{m=1}^\infty H_{m,n}) = 0$. Then, $\exists m_n \in \mathbb{N}$ such that $\mu(H_{m_n,n}) < 2^{-n}\eta$. Let $A_n := H_{m_n,n} \in \mathcal{B}$. Then, $\mu(A_n) < 2^{-n}\eta$ and $\forall i \geq m_n$, $\forall x \in (X \setminus E) \setminus A_n$, we have $\rho(f_i(x), f(x)) < 2^{-n}$.

Let $A := E \cup (\bigcup_{n=1}^\infty A_n) \in \mathcal{B}$. Clearly, $A \supseteq \widetilde{D_1} \cup \widetilde{D_2}$. By countable subadditivity of measure, we have $\mu(A) \leq \mu(E) + \sum_{n=1}^\infty \mu(A_n) < \eta$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ with $0 < 2^{-n_0} \leq \epsilon$. $\forall j \in \mathbb{N}$ with $m_{n_0} \leq j$, $\forall x \in X \setminus A \subseteq (X \setminus E) \setminus A_{n_0}$, we have $\rho(f_j(x), f(x)) < 2^{-n_0} \leq \epsilon$. Hence,

$\left(f_n|_{X \setminus A}\right)_{n=1}^{\infty}$ converges uniformly to $f|_{X \setminus A}$. This completes the proof of the proposition. $\square$

Definition 11.56 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_n \in \mathcal{B}$, $f_n : D_n \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $D \in \mathcal{B}$, and $f : D \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. We will say that $(f_n)_{n=1}^{\infty}$ converges to f in measure in $\mathcal{X}$ if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n \geq N$, we have $\mu(A_{n, \epsilon}) := \mu(\widetilde{D_n} \cup \widetilde{D} \cup \{x \in D_n \cap D \mid \rho(f_n(x), f(x)) \geq \epsilon\}) < \epsilon$. In this case, we will write $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$.

Note that in the above definition, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $A_{n, \epsilon} \in \mathcal{B}$ by Lemma 11.43. Hence, the definition is well-defined.

Proposition 11.57 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_1, D_2 \in \mathcal{B}$, $f_n : D_1 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $f : D_2 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$. Then, there exists a subsequence $(f_{n_k})_{k=1}^{\infty}$ of $(f_n)_{n=1}^{\infty}$ such that $\lim_{k \in \mathbb{N}} f_{n_k} = f$ a.e. in $\mathcal{X}$.

Proof Define $A_{n, k} := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid \rho(f_n(x), f(x)) \geq 2^{-k}\}$, $\forall n, k \in \mathbb{N}$. By Lemma 11.43, $A_{n, k} \in \mathcal{B}$. By the assumption, $\exists n_1 \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n \geq n_1$, we have $\mu(A_{n, 1}) < 2^{-1}$. $\forall k \in \mathbb{N}$ with $k > 1$, by the assumption, $\exists n_k \in \mathbb{N}$ with $n_k > n_{k-1}$ such that $\forall n \in \mathbb{N}$ with $n \geq n_k$, we have $\mu(A_{n, k}) < 2^{-k}$. This defines a subsequence $(f_{n_k})_{k=1}^{\infty}$.

Let $A := \bigcap_{i=1}^{\infty} \bigcup_{k=i}^{\infty} A_{n_k, k} \in \mathcal{B}$. Then, $\mu(A) \leq \mu(\bigcup_{k=i}^{\infty} A_{n_k, k}) \leq \sum_{k=i}^{\infty} \mu(A_{n_k, k}) < 2^{-i+1}$, $\forall i \in \mathbb{N}$. Hence, $\mu(A) = 0$. $\forall x \in X \setminus A = \bigcup_{i=1}^{\infty} \bigcap_{k=i}^{\infty} \widetilde{A_{n_k, k}}$, $\exists i_0 \in \mathbb{N}$ such that $\forall k \in \mathbb{N}$ with $k \geq i_0$, $x \in \widetilde{A_{n_k, k}}$. Then, we have $x \in D_1 \cap D_2$ and $\rho(f_{n_k}(x), f(x)) < 2^{-k}$. Hence, $\lim_{k \in \mathbb{N}} f_{n_k}(x) = f(x)$. Then, $E := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid (f_{n_k}(x))_{k=1}^{\infty} \text{ does not converge to } f(x)\} \subseteq A$. By Proposition 11.50, we have $E \in \mathcal{B}$ and $\mu(E) = 0$. Therefore, $\lim_{k \in \mathbb{N}} f_{n_k} = f$ a.e. in $\mathcal{X}$. This completes the proof of the proposition. $\square$

Proposition 11.58 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_1, D_2 \in \mathcal{B}$, $f_n : D_1 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $f : D_2 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$. Then, $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$.

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Egoroff's Theorem 11.55, $\exists A \in \mathcal{B}$ with $\mu(A) < \epsilon$ and $A \supseteq \widetilde{D_1} \cup \widetilde{D_2}$ such that $\left(f_n|_{X \setminus A}\right)_{n=1}^{\infty}$ converges uniformly to $f|_{X \setminus A}$. Then, $\exists n_0 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have $\rho(f_n(x), f(x)) < \epsilon$, $\forall x \in X \setminus A$. Then, $A_{n, \epsilon} := \widetilde{D_1} \cup \widetilde{D_2} \cup \{x \in D_1 \cap D_2 \mid \rho(f_n(x), f(x)) \geq \epsilon\} \subseteq A$. By Lemma 11.43, $A_{n, \epsilon} \in \mathcal{B}$ and $\mu(A_{n, \epsilon}) \leq \mu(A) < \epsilon$. This shows that $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$. This completes the proof of the proposition. $\square$

Proposition 11.59 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_n \in \mathcal{B}$, $f_n : D_n \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $D \in \mathcal{B}$, and $f : D \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Then, $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$ if, and only if, every subsequence $(f_{n_k})_{k=1}^\infty$ of $(f_n)_{n=1}^\infty$ admits a subsequence $(f_{n_{k_j}})_{j=1}^\infty$ that converges to f in measure.*

Proof “Necessity” Let $(f_n)_{n=1}^\infty$ converges to f in measure and $(f_{n_k})_{k=1}^\infty$ be a subsequence of $(f_n)_{n=1}^\infty$. Clearly, $(f_{n_k})_{k=1}^\infty$ converges to f in measure, and it is a subsequence of itself. Hence, the result holds.

“Sufficiency” Assume that every subsequence $(f_{n_k})_{k=1}^\infty$ of $(f_n)_{n=1}^\infty$ admits a subsequence $(f_{n_{k_j}})_{j=1}^\infty$ that converges to f in measure. We will prove the result using an argument of contradiction. Suppose $(f_n)_{n=1}^\infty$ does not converge to f in measure. Then, $\exists \epsilon_0 \in (0, \infty) \subset \mathbb{R}$, $\forall n_0 \in \mathbb{N}$, $\exists n \in \mathbb{N}$ with $n \geq n_0$ such that $\mu(A_n) := \mu(\tilde{D}_n \cup \tilde{D} \cup \{x \in D_n \cap D \mid \rho(f_n(x), f(x)) \geq \epsilon_0\}) \geq \epsilon_0$, where $A_n \in \mathcal{B}$ by Lemma 11.43. Then, $\exists n_1 \in \mathbb{N}$ such that $\mu(A_{n_1}) \geq \epsilon_0$. $\forall k \in \mathbb{N}$, $\exists n_{k+1} \in \mathbb{N}$ with $n_{k+1} > n_k$ such that $\mu(A_{n_{k+1}}) \geq \epsilon_0$. This defines a subsequence $(f_{n_k})_{k=1}^\infty$, which satisfies that $\forall k \in \mathbb{N}$, $\mu(A_{n_k}) \geq \epsilon_0$. Clearly, there does not exist a subsequence of $(f_{n_k})_{k=1}^\infty$ that converges to f in measure. This contradicts with the assumption. Therefore, $(f_n)_{n=1}^\infty$ converges to f in measure.

This completes the proof of the proposition. $\square$

Proposition 11.60 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_1, D_2 \in \mathcal{B}$, $f_n : D_1 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $f : D_2 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Then, $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$ if, and only if, every subsequence $(f_{n_k})_{k=1}^\infty$ of $(f_n)_{n=1}^\infty$ admits a subsequence $(f_{n_{k_j}})_{j=1}^\infty$ that converges to f a.e. in $\mathcal{X}$.*

Proof “Necessity” Let $(f_n)_{n=1}^\infty$ converge to f in measure. Let $(f_{n_k})_{k=1}^\infty$ be a subsequence of $(f_n)_{n=1}^\infty$. Clearly, $(f_{n_k})_{k=1}^\infty$ converges to f in measure. By Proposition 11.57, there exists a subsequence $(f_{n_{k_j}})_{j=1}^\infty$ of $(f_{n_k})_{k=1}^\infty$ that converges to f a.e. in $\mathcal{X}$.

“Sufficiency” Assume that every subsequence $(f_{n_k})_{k=1}^\infty$ of $(f_n)_{n=1}^\infty$ admits a subsequence $(f_{n_{k_j}})_{j=1}^\infty$ that converges to f a.e. in $\mathcal{X}$. By Proposition 11.58, $(f_{n_{k_j}})_{j=1}^\infty$ converges to f in measure. By Proposition 11.59, we have $(f_n)_{n=1}^\infty$ converges to f in measure in $\mathcal{X}$.

This completes the proof of the proposition. $\square$

Definition 11.61 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_n \in \mathcal{B}$, and $f_n : D_n \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. We will say that $(f_n)_{n=1}^\infty$ is Cauchy in measure in $\mathcal{X}$ if $\forall \epsilon \in$*

$(0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$, $\forall n, m \in \mathbb{N}$ with $n \geq n_0$ and $m \geq n_0$, we have $\mu(\widetilde{D_n \cup D_m} \cup \{x \in D_n \cap D_m \mid \rho(f_n(x), f_m(x)) \geq \epsilon\}) < \epsilon$.

Clearly, a sequence of measurable functions that converges in measure to a measurable function is Cauchy in measure. The converse is established in the following proposition.

Proposition 11.62 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a separable complete metric space, $D_n \in \mathcal{B}$, and $f_n : D_n \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. Assume that $(f_n)_{n=1}^\infty$ is Cauchy in measure in $\mathcal{X}$. Then, $\exists f : D \rightarrow \mathcal{Y}$, where $D := \bigcup_{n=1}^\infty D_n \in \mathcal{B}$ such that f is $\mathcal{B}$ -measurable and $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$.*

Proof Let $n_0 := 0$. $\forall i \in \mathbb{N}$, by the assumption, $\exists n_i \in \mathbb{N}$ with $n_i > n_{i-1}$ such that $\forall n, m \in \mathbb{N}$ with $n \geq n_i$ and $m \geq n_i$, we have $\mu(\widetilde{D_n \cup D_m} \cup \{x \in D_n \cap D_m \mid \rho(f_n(x), f_m(x)) \geq 2^{-i}\}) < 2^{-i}$. This defines a subsequence $(f_{n_i})_{i=1}^\infty$ of $(f_n)_{n=1}^\infty$. $\forall i \in \mathbb{N}$, let $A_i := \widetilde{D_{n_i} \cup D_{n_{i+1}}} \cup \{x \in D_{n_i} \cap D_{n_{i+1}} \mid \rho(f_{n_i}(x), f_{n_{i+1}}(x)) \geq 2^{-i}\}$. Then, $A_i \in \mathcal{B}$ and $\mu(A_i) < 2^{-i}$. Let $A := \bigcap_{i=1}^\infty \bigcup_{j=i}^\infty A_j \in \mathcal{B}$. Then, $\mu(A) \leq \mu(\bigcup_{j=i}^\infty A_j) \leq \sum_{j=i}^\infty \mu(A_j) < 2^{-i+1}$, $\forall i \in \mathbb{N}$. Hence, $\mu(A) = 0$. $\forall x \in X \setminus A = \bigcup_{i=1}^\infty \bigcap_{j=i}^\infty \widetilde{A_j} \subseteq D$, $\exists i_0 \in \mathbb{N}$ such that $x \in \widetilde{A_i}$ and $\rho(f_{n_i}(x), f_{n_{i+1}}(x)) < 2^{-i}$, $\forall i \in \mathbb{N}$ with $i \geq i_0$. Then, $(f_{n_i}(x))_{i=i_0}^\infty \subseteq \mathcal{Y}$ is a Cauchy sequence. By the completeness of $\mathcal{Y}$, $\exists! f(x) \in \mathcal{Y}$ such that $\lim_{i \in \mathbb{N}} f_{n_i}(x) = f(x)$. $\forall x \in A \cap D$, we assign $f(x) := f_{\bar{n}}(\bar{x}) \in \mathcal{Y}$, where $\bar{n} \in \mathbb{N}$ and $\bar{x} \in D_{\bar{n}} \subseteq D$ are fixed independent of x . This defines a function $f : D \rightarrow \mathcal{Y}$ that satisfies $\lim_{i \in \mathbb{N}} f_{n_i}(x) = f(x)$, $\forall x \in X \setminus A$. $\forall i \in \mathbb{N}$, define $\bar{f}_{n_i} : D \rightarrow \mathcal{Y}$ by $\bar{f}_{n_i}(x) = \begin{cases} f_{n_i}(x) & x \in D_{n_i} \\ f_{\bar{n}}(\bar{x}) & x \in D \setminus D_{n_i} \end{cases}$, $\forall x \in D$. By Proposition 11.41, $\bar{f}_{n_i}$ is $\mathcal{B}$ -measurable. Clearly, $\lim_{i \in \mathbb{N}} \bar{f}_{n_i}(x) = f(x)$, $\forall x \in D$. By Proposition 11.48, f is $\mathcal{B}$ -measurable.

$\forall i \in \mathbb{N}$, $\forall x \in \bigcup_{j=i}^\infty \widetilde{A_j} = \bigcap_{j=i}^\infty \widetilde{A_j} \subseteq X \setminus A$, we have $\rho(f_{n_j}(x), f_{n_{j+1}}(x)) < 2^{-j}$, $\forall j \geq i$. Then, by Propositions 4.30, 3.66, and 3.67, $\rho(f_{n_i}(x), f(x)) = \lim_{l \in \mathbb{N}} \rho(f_{n_i}(x), f_{n_l}(x)) \leq \lim_{l \in \mathbb{N}} \sum_{j=i}^{l-1} \rho(f_{n_j}(x), f_{n_{j+1}}(x)) < 2^{-i+1}$. Then, $E_i := \widetilde{D_{n_i}} \cup \{x \in D_{n_i} \mid \rho(f_{n_i}(x), f(x)) \geq 2^{-i+1}\} \subseteq \bigcup_{j=i}^\infty A_j$. By Lemma 11.43, $E_i \in \mathcal{B}$ and $\mu(E_i) \leq \sum_{j=i}^\infty \mu(A_j) < 2^{-i+1}$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists i_0 \in \mathbb{N}$ such that $2^{-i_0+1} < \epsilon/2$. $\forall n \in \mathbb{N}$ with $n \geq n_{i_0}$, we have $\mu(\widetilde{D_n} \cup \{x \in D_n \mid \rho(f_n(x), f(x)) \geq \epsilon\}) \leq \mu(\widetilde{D_{n_{i_0}}} \cup \{x \in D_n \cap D_{n_{i_0}} \mid \rho(f_n(x), f_{n_{i_0}}(x)) + \rho(f_{n_{i_0}}(x), f(x)) \geq \epsilon\}) \leq \mu(\widetilde{D_n \cup D_{n_{i_0}}} \cup \{x \in D_n \cap D_{n_{i_0}} \mid \rho(f_n(x), f_{n_{i_0}}(x)) \geq \epsilon/2\}) + \mu(\widetilde{D_{n_{i_0}}} \cup \{x \in D_{n_{i_0}} \mid \rho(f_{n_{i_0}}(x), f(x)) \geq \epsilon/2\}) < 2^{-i_0} + 2^{-i_0+1} < \epsilon$. Hence, $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$. This completes the proof of the proposition. $\square$

Definition 11.63 Let X be a set and $A \subseteq X$. The indicator function $\chi_{A,X} : X \rightarrow \{0, 1\} \subset \mathbb{R}$ is defined by $\chi_{A,X}(x) = \begin{cases} 1 & x \in A \\ 0 & x \in X \setminus A \end{cases}, \forall x \in X$.

Definition 11.64 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a normed linear space. $\phi : X \rightarrow \mathcal{Y}$ is said to be a simple function if $\exists n \in \mathbb{Z}_+$, $\exists y_1, \dots, y_n \in \mathcal{Y}$, and $\exists A_1, \dots, A_n \in \mathcal{B}$ with $\mu(A_i) < +\infty, i = 1, \dots, n$, such that $\phi(x) = \sum_{i=1}^n y_i \chi_{A_i,X}(x), \forall x \in X$. We will say that a simple function ϕ is in canonical representation if $y_1, \dots, y_n$ are distinct and none equals to $\vartheta_{\mathcal{Y}}$, and $A_1, \dots, A_n$ are nonempty and pairwise disjoint.

Clearly, every simple function admits a unique canonical representation.

Proposition 11.65 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y} := (\mathcal{Y}, \mathbb{K}, \|\cdot\|)$ be a separable normed linear space, and $f : X \rightarrow U \subseteq \mathcal{Y}$ be $\mathcal{B}$ -measurable. Then,

- (i) $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists$ a simple function $\phi_\epsilon : X \rightarrow U$ such that $\mu(\{x \in X \mid \|f(x) - \phi_\epsilon(x)\| \geq \epsilon\}) < \epsilon$;
- (ii) there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty, \varphi_n : X \rightarrow U, \forall n \in \mathbb{N}$, that converges to f in measure.

Proof (i) By Proposition 4.38, U , considered as a subspace of the metric space $\mathcal{Y}$, is separable. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $U = \emptyset$; Case 2: $U \neq \emptyset$. Case 1: $U = \emptyset$. Then, $X = \emptyset$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, choose $\phi_\epsilon : X \rightarrow U$ to be the simple function with $n = 0$ as in Definition 11.64. Then, $\mu(\{x \in X \mid \|f(x) - \phi_\epsilon(x)\| \geq \epsilon\}) = \mu(\emptyset) = 0 < \epsilon$. Hence, (i) holds. Case 2: $U \neq \emptyset$. Then, $\exists (y_i)_{i=1}^\infty \subseteq U$ such that $\forall y \in U, \forall \delta \in (0, \infty) \subset \mathbb{R}, \exists i \in \mathbb{N} \ni \cdot, \|y - y_i\| < \delta$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, we have $U \subseteq \bigcup_{i=1}^\infty \mathcal{B}_{\mathcal{Y}}(y_i, \epsilon) =: \bigcup_{i=1}^\infty \bar{V}_i$. Define $V_1 := \bar{V}_1 \in \mathcal{B}_{\mathcal{B}}(\mathcal{Y}), V_{i+1} := \bar{V}_{i+1} \setminus (\bigcup_{j=1}^i \bar{V}_j) \in \mathcal{B}_{\mathcal{B}}(\mathcal{Y}), \forall i \in \mathbb{N}$. Clearly, $U \subseteq \bigcup_{i=1}^\infty V_i$ and $V_1, V_2, \dots$ are pairwise disjoint. Then, by Propositions 2.5 and 11.34, $X = f_{\text{inv}}(U) = \bigcup_{i=1}^\infty f_{\text{inv}}(V_i)$ and $(f_{\text{inv}}(V_i))_{i=1}^\infty \subseteq \mathcal{B}$ and are pairwise disjoint. Then, $\mu(X) = \sum_{i=1}^\infty \mu(f_{\text{inv}}(V_i)) < +\infty$. This implies that $\exists n_0 \in \mathbb{N}$ such that $\sum_{i=n_0+1}^\infty \mu(f_{\text{inv}}(V_i)) < \epsilon$. Let $A_i := f_{\text{inv}}(V_i) \in \mathcal{B}, i = 1, \dots, n_0$, and $A_{n_0+1} := X \setminus (\bigcup_{i=1}^{n_0} A_i) \in \mathcal{B}$. Define the simple function $\phi_\epsilon : X \rightarrow U$ by $\phi_\epsilon(x) = \sum_{i=1}^{n_0+1} y_i \chi_{A_i,X}(x), \forall x \in X$. $\forall x \in X, \exists! i_x \in \{1, \dots, n_0+1\}$ such that $x \in A_{i_x}$. Then, $\phi_\epsilon(x) = y_{i_x} \in U$. $\forall x \in X \setminus A_{n_0+1}, i_x \in \{1, \dots, n_0\}$ and $x \in A_{i_x} = f_{\text{inv}}(V_{i_x}) \subseteq f_{\text{inv}}(\bar{V}_{i_x})$. Then, $\phi_\epsilon(x) = y_{i_x}$ and $f(x) \in \bar{V}_{i_x} = \mathcal{B}_{\mathcal{Y}}(y_{i_x}, \epsilon)$. Hence, $\|f(x) - \phi_\epsilon(x)\| < \epsilon$. Then, $A := \{x \in X \mid \|f(x) - \phi_\epsilon(x)\| \geq \epsilon\} \subseteq A_{n_0+1} = \bigcup_{i=n_0+1}^\infty f_{\text{inv}}(V_i)$. By Lemma 11.43, $A \in \mathcal{B}$. Then, $\mu(A) \leq \sum_{i=n_0+1}^\infty \mu(f_{\text{inv}}(V_i)) < \epsilon$. Hence, (i) holds. In both cases, we have proved (i).

(ii) $\forall n \in \mathbb{N}$, let $\varphi_n := \phi_{2^{-n}}$. Then, $(\varphi_n)_{n=1}^\infty$ converges to f in measure. This completes the proof of the proposition. $\square$

Proposition 11.66 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a separable normed linear space, $U \subseteq \mathcal{Y}$ be a conic segment, and $f : X \rightarrow U \subseteq \mathcal{Y}$ be $\mathcal{B}$ -measurable. Then, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : X \rightarrow U$, $\forall n \in \mathbb{N}$, such that $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$ and $\|\varphi_n(x)\| \leq \|f(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$.*

Proof We will first consider the special case $\mathcal{Y} = \mathbb{R}$ and $\mathcal{X}$ is finite. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Proposition 11.5, $\exists M_\epsilon \in \mathbb{N}$ such that $\mu(\{x \in X \mid |f(x)| \geq M_\epsilon \epsilon\}) =: \mu(A_\epsilon) < \epsilon$. Let $I_i := [-(M_\epsilon - i + 1)\epsilon, -(M_\epsilon - i)\epsilon) \subset \mathbb{R}$, $i = 1, \dots, 2M_\epsilon$ and $A_i := f_{\text{inv}}(I_i) \in \mathcal{B}$. Note that $\mu(A_i) \leq \mu(X) < +\infty$, $i = 1, \dots, 2M_\epsilon$ and A_i 's are pairwise disjoint. Define $\phi_\epsilon := \sum_{i=1}^{M_\epsilon} -(M_\epsilon - i)\epsilon \chi_{A_i, X} + \sum_{i=M_\epsilon+1}^{2M_\epsilon} -(M_\epsilon - i + 1)\epsilon \chi_{A_i, X}$. Clearly, we have $\phi_\epsilon : X \rightarrow U$, $|\phi_\epsilon(x)| \leq |f(x)|$, $\forall x \in X$, and $\mu(\{x \in X \mid |f(x) - \phi_\epsilon(x)| > \epsilon\}) \leq \mu(A_\epsilon) < \epsilon$. Then, the sequence of simple functions $(\phi_{2^{-n}})_{n=1}^\infty$ converges to f in measure in $\mathcal{X}$. By Proposition 11.57, there exists a subsequence $(\varphi_k)_{k=1}^\infty = (\phi_{2^{-n_k}})_{k=1}^\infty$ such that $\lim_{k \in \mathbb{N}} \varphi_k = f$ a.e. in $\mathcal{X}$. The result holds in this special case.

Next, we consider the special case $\mathcal{X}$ is finite. By Proposition 11.65, there exists a sequence of simple functions $(\bar{\psi}_i)_{i=1}^\infty$, $\bar{\psi}_i : X \rightarrow U$, $\forall i \in \mathbb{N}$, that converges to f in measure. By Proposition 11.57, there exists a subsequence $(\psi_n)_{n=1}^\infty = (\bar{\psi}_{i_n})_{n=1}^\infty$, such that $\lim_{n \in \mathbb{N}} \psi_n = f$ a.e. in $\mathcal{X}$. By Propositions 7.21 and 11.38, $\mathcal{P} \circ f$ is $\mathcal{B}$ -measurable, where $\mathcal{P} \circ f : X \rightarrow [0, \infty) \subset \mathbb{R}$ is defined by $\mathcal{P} \circ f(x) = \|f(x)\|$, $\forall x \in X$.¹ By the previous special case, there exists a sequence of simple functions $(\phi_n)_{n=1}^\infty$, $\phi_n : X \rightarrow [0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, such that $\lim_{n \in \mathbb{N}} \phi_n = \mathcal{P} \circ f$ a.e. in $\mathcal{X}$ and $0 \leq \phi_n(x) \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Let $E_1 := \{x \in X \mid (\psi_n(x))_{n=1}^\infty \text{ does not converge to } f(x)\}$ and $E_2 := \{x \in X \mid (\phi_n(x))_{n=1}^\infty \text{ does not converge to } \|f(x)\|\}$. Then, $E_1, E_2 \in \mathcal{B}$ and $\mu(E_1) = \mu(E_2) = 0$. Fix any $n \in \mathbb{N}$. Let ψ_n admit the canonical representation $\psi_n = \sum_{j=1}^{\tilde{n}} y_j \chi_{A_j, X}$, where $\tilde{n} \in \mathbb{Z}_+$, $y_1, \dots, y_{\tilde{n}} \in U$ are distinct and none equals to $\vartheta_{\mathcal{Y}}$, $A_1, \dots, A_{\tilde{n}} \in \mathcal{B}$ are pairwise disjoint, nonempty, and of finite measure. Let $y_{\tilde{n}+1} := \vartheta_{\mathcal{Y}} \in U$ and $A_{\tilde{n}+1} := X \setminus (\bigcup_{j=1}^{\tilde{n}} A_j) \in \mathcal{B}$. Then, $\psi_n = \sum_{j=1}^{\tilde{n}+1} y_j \chi_{A_j, X}$. Let ϕ_n admit the canonical representation $\sum_{j=1}^{\bar{n}} a_j \chi_{\bar{A}_j, X}$, where $\bar{n} \in \mathbb{Z}_+$, $a_1, \dots, a_{\bar{n}} \in [0, \infty) \subset \mathbb{R}$ are distinct and none equals to 0, $\bar{A}_1, \dots, \bar{A}_{\bar{n}} \in \mathcal{B}$ are pairwise disjoint, nonempty, and of finite measure. Let $a_{\bar{n}+1} := 0$ and $\bar{A}_{\bar{n}+1} := X \setminus (\bigcup_{j=1}^{\bar{n}} \bar{A}_j) \in \mathcal{B}$. Then, $\phi_n = \sum_{j=1}^{\bar{n}+1} a_j \chi_{\bar{A}_j, X}$. Define a simple function $\varphi_n : X \rightarrow U$ by $\varphi_n(x) = \begin{cases} y_j & \text{if } \|y_j\| \leq a_l \\ (a_l / \|y_j\|) y_j & \text{if } \|y_j\| > a_l \end{cases}, \forall x \in A_j \cap \bar{A}_l, j = 1, \dots, \tilde{n} + 1, l = 1, \dots, \bar{n} + 1$. Clearly, $\varphi_n(x) \in U$, $\forall x \in X$. Alternatively, we

¹This notation will be used throughout the rest of the notes. With this notation, we distinguish that $\|f\|$ is a nonnegative real number and $\mathcal{P} \circ f$ is a nonnegative real valued function.

have $\varphi_n(x) = \begin{cases} \psi_n(x) & \text{if } \|\psi_n(x)\| \leq \phi_n(x) \\ (\phi_n(x)/\|\psi_n(x)\|)\psi_n(x) & \text{if } \|\psi_n(x)\| > \phi_n(x) \end{cases}, \forall x \in X.$

Then, we have $\|\varphi_n(x)\| \leq \phi_n(x) \leq \mathcal{P} \circ f(x), \forall x \in X, \forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n(x) = f(x), \forall x \in X \setminus (E_1 \cup E_2)$. Let $E := \{x \in X \mid (\varphi_n(x))_{n=1}^\infty \text{ does not converge to } f(x)\}$. Then, $E \subseteq E_1 \cup E_2$. By Proposition 11.50, $E \in \mathcal{B}$. This implies that $0 \leq \mu(E) \leq \mu(E_1) + \mu(E_2) = 0$. Therefore, $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$. Hence, the result holds in this special case.

Finally, we consider the general case. Since $\mathcal{X}$ is σ -finite, then $\exists (X_i)_{i=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{i=1}^\infty X_i$ and $\mu(X_i) < +\infty, \forall i \in \mathbb{N}$. Without loss of generality, we may assume that $X_i \subseteq X_{i+1}, \forall i \in \mathbb{N}$. Fix any $i \in \mathbb{N}$, let $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be the finite measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. Then, $f|_{X_i} : X_i \rightarrow U$ is $\mathcal{B}_i$ -measurable. By the second special case, there exists a sequence of simple functions $(\phi_{i,n})_{n=1}^\infty, \phi_{i,n} : X_i \rightarrow U, \forall n \in \mathbb{N}$, such that $\|\phi_{i,n}(x)\| \leq \|f(x)\|, \forall x \in X_i, \forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \phi_{i,n} = f|_{X_i}$ a.e. in $\mathcal{X}_i$. By Proposition 11.58, $\lim_{n \in \mathbb{N}} \phi_{i,n} = f|_{X_i}$ in measure in $\mathcal{X}_i$. Then, $\exists n_i \in \mathbb{N}$ such that $\mu_i(E_i) := \mu_i(\{x \in X_i \mid \|f|_{X_i}(x) - \phi_{i,n_i}(x)\| \geq 2^{-i}\}) < 2^{-i}$. Let ϕ_{i,n_i} admit the canonical representation $\phi_{i,n_i} = \sum_{j=1}^{\bar{n}_i} y_{i,j} \chi_{A_{i,j}, X_i}$, where $\bar{n}_i \in \mathbb{Z}_+, y_{i,1}, \dots, y_{i,\bar{n}_i} \in U$ are distinct and none equals to ϑ_y , and $A_{i,1}, \dots, A_{i,\bar{n}_i} \in \mathcal{B}_i \subseteq \mathcal{B}$ are pairwise disjoint, nonempty, and of finite measure. Then, we may define a simple function $\varphi_i : X \rightarrow U \cup \{\vartheta_y\} = U$ by $\varphi_i = \sum_{j=1}^{\bar{n}_i} y_{i,j} \chi_{A_{i,j}, X}$. Note that $\varphi_i(x) = \phi_{i,n_i}(x)$ and $\|\varphi_i(x)\| \leq \|f(x)\|, \forall x \in X_i, \varphi_i(x) = \vartheta_y \in U$ and $\|\varphi_i(x)\| = 0 \leq \|f(x)\|, \forall x \in X \setminus X_i$. Hence, $\|\varphi_i(x)\| \leq \|f(x)\|, \forall x \in X$. Then, $E_i = \{x \in X \mid \|f(x) - \varphi_i(x)\| \geq 2^{-i}\} \cap X_i \in \mathcal{B}$ and $\mu(E_i) = \mu_i(E_i) < 2^{-i}$ by Proposition 11.13.

Let $E := \bigcap_{i=1}^\infty \bigcup_{j=i}^\infty E_j \in \mathcal{B}$. $\mu(E) \leq \mu(\bigcup_{j=i}^\infty E_j) < 2^{-i+1}, \forall i \in \mathbb{N}$. Then, $\mu(E) = 0$. $\forall x \in X \setminus E, \exists i_0 \in \mathbb{N}$ such that $x \in X_{i_0}$. Then, $x \in X_j, \forall j \in \mathbb{N}$ with $j \geq i_0$. $\exists i_1 \in \mathbb{N}$ such that $x \in X \setminus E_j, \forall j \in \mathbb{N}$ with $j \geq i_1$. Let $k_0 = \max\{i_0, i_1\}$. $\forall k \geq k_0, x \in X_k \setminus E_k$ and $\|\varphi_k(x) - f(x)\| < 2^{-k}$. This implies that $\lim_{k \in \mathbb{N}} \varphi_k(x) = f(x)$. Then, $\bar{E} := \{x \in X \mid (\varphi_k(x))_{k=1}^\infty \text{ does not converge to } f(x)\} \subseteq E$. By Proposition 11.50, we have $\bar{E} \in \mathcal{B}$ and $\mu(\bar{E}) = 0$. Hence, $\lim_{k \in \mathbb{N}} \varphi_k = f$ a.e. in $\mathcal{X}$. This completes the proof of the proposition. $\square$

Proposition 11.67 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space and $f : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable. Then, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty, \varphi_n : X \rightarrow [0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N}$, such that $0 \leq \varphi_n(x) \leq \varphi_{n+1}(x) \leq f(x), \forall x \in X, \forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n(x) = f(x), \forall x \in X$.*

Proof $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty X_n$ and $\mu(X_n) < \infty, \forall n \in \mathbb{N}$. Without loss of generality, we may assume that $X_n \subseteq X_{n+1}, \forall n \in \mathbb{N}$. Fix any $n \in \mathbb{N}$. Let $I_{i,n} := ((i-1)/n, i/n], E_{i,n} := f_{\text{inv}}(I_{i,n}) \cap X_n \in \mathcal{B}, i = 1, \dots, n^2$, and $E_{n^2+1,n} := X_n \setminus (\bigcup_{i=1}^{n^2} E_{i,n})$. Note that $E_{i,n}$'s are

pairwise disjoint and $\mu(E_i) \leq \mu(X_n) < \infty$, $\forall i \in \{1, \dots, n^2 + 1\}$. Define $\phi_n := \sum_{i=1}^{n^2+1} (i-1)/n \chi_{E_{i,n}, X}$. Then, $\phi_n : X \rightarrow [0, \infty) \subset \mathbb{R}$, $0 \leq \phi_n(x) \leq f(x)$, $\forall x \in X$, ϕ_n is a simple function, and $0 \leq f(x) - \phi_n(x) \leq 1/n$, $\forall x \in f_{\text{inv}}([0, n]) \cap X_n$. Define $\varphi_n := \bigvee_{i=1}^n \phi_i$. Then, $\varphi_n : X \rightarrow [0, \infty) \subset \mathbb{R}$, $0 \leq \varphi_n(x) \leq \varphi_{n+1}(x) \leq f(x)$, $\forall x \in X$, φ_n is a simple function, and $0 \leq f(x) - \varphi_n(x) \leq 1/n$, $\forall x \in f_{\text{inv}}([0, n]) \cap X_n$.

Fix any $x \in X$. If $f(x) = \infty$, then $\exists n_0 \in \mathbb{N}$ such that $x \in X_{n_0}$. $\forall n \in \mathbb{N}$ with $n_0 \leq n$, $x \in X_n$, $x \in E_{n^2+1, n}$, and $\phi_n(x) = n$. This further implies that $\varphi_n(x) = n$, $\forall n \geq n_0$. Hence, $\infty = f(x) = \lim_{n \in \mathbb{N}} \varphi_n(x)$. On the other hand, if $f(x) < \infty$, then $\exists n_0 \in \mathbb{N}$ such that $x \in X_{n_0}$ and $0 \leq f(x) \leq n_0$. Then, $\forall n \in \mathbb{N}$ with $n_0 \leq n$, $|f(x) - \varphi_n(x)| \leq 1/n$. This implies that $\lim_{n \in \mathbb{N}} \varphi_n(x) = f(x)$.

Hence, we have $f(x) = \lim_{n \in \mathbb{N}} \varphi_n(x)$, $\forall x \in X$. This completes the proof of the proposition. $\square$

Lemma 11.68 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $U \subseteq \mathcal{Y}$ be a σ -compact conic segment, and $f : X \rightarrow U$ be $\mathcal{B}$ -measurable. Then, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : X \rightarrow U$, $\forall n \in \mathbb{N}$ such that $\lim_{n \in \mathbb{N}} \varphi_n(x) = f(x)$, $\forall x \in X$, and $\|\varphi_n(x)\| \leq \|f(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$.*

Proof Since U is σ -compact, then $U = \bigcup_{n=1}^\infty K_n$, where K_n is compact, $\forall n \in \mathbb{N}$. Since $\mathcal{X}$ is σ -finite, then $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty X_n$, and $\mu(X_n) < +\infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $K_n \subseteq K_{n+1}$ and $X_n \subseteq X_{n+1}$, $\forall n \in \mathbb{N}$.

Fix any $n \in \mathbb{N}$. $K_n \subseteq \bigcup_{y \in K_n} \mathcal{B}_{\mathcal{Y}}(y, 1/n)$. By the compactness of K_n , $\exists D_n \subseteq K_n$ where D_n is a finite set, such that $K_n \subseteq \bigcup_{y \in D_n} \mathcal{B}_{\mathcal{Y}}(y, 1/n)$. Without loss of generality, assume that $D_n = \{y_{n,1}, \dots, y_{n,m_n}\}$, where $m_n \in \mathbb{Z}_+$. Let $E_{n,1} := \mathcal{B}_{\mathcal{Y}}(y_{n,1}, 1/n)$, and $E_{n,i} := \mathcal{B}_{\mathcal{Y}}(y_{n,i}, 1/n) \setminus \left(\bigcup_{j=1}^{i-1} E_{n,j}\right)$, $\forall i = 2, \dots, m_n$. Then, $K_n \subseteq \bigcup_{i=1}^{m_n} E_{n,i}$ and the sets in the union are pairwise disjoint. Let $A_{n,i} := f_{\text{inv}}(E_{n,i}) \cap X_n \in \mathcal{B}$, $i = 1, \dots, m_n$, since $E_{n,i} \in \mathcal{B}_{\mathcal{Y}}(\mathcal{Y})$ and f is $\mathcal{B}$ -measurable. Note that $\mu(A_{n,i}) \leq \mu(X_n) < +\infty$, $i = 1, \dots, m_n$, and $(A_{n,i})_{i=1}^{m_n}$ is pairwise disjoint. Define $\psi_n : X \rightarrow K_n \subseteq U$ by $\psi_n := \sum_{i=1}^{m_n} y_{n,i} \chi_{A_{n,i}, X}$. Clearly, ψ_n is a simple function. $\forall x \in X$, $\exists N \in \mathbb{N}$ such that $x \in X_N$ and $f(x) \in K_N$. Then, $\|f(x) - \psi_n(x)\| < 1/n$, $\forall n \in \mathbb{N}$ with $n \geq N$. Hence, $\lim_{n \in \mathbb{N}} \psi_n(x) = f(x)$, $\forall x \in X$.

By Propositions 7.21 and 11.38, $\mathcal{P} \circ f : X \rightarrow [0, \infty) \subset \mathbb{R}$ is $\mathcal{B}$ -measurable. By Proposition 11.67, there exists a sequence of simple functions $(\phi_n)_{n=1}^\infty$ such that $0 \leq \phi_n(x) \leq \phi_{n+1}(x) \leq \mathcal{P} \circ f(x) = \|f(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \phi_n(x) = \mathcal{P} \circ f(x)$, $\forall x \in X$.

Fix any $n \in \mathbb{N}$, let ϕ_n admit the canonical representation $\phi_n = \sum_{i=1}^{\bar{m}_n} a_j \chi_{\bar{A}_{n,j}, X}$, where $\bar{m}_n \in \mathbb{Z}_+$, $a_1, \dots, a_{\bar{m}_n} \in [0, \infty) \subset \mathbb{R}$ are distinct and none equals to 0, $\bar{A}_{n,1}, \dots, \bar{A}_{n,\bar{m}_n} \in \mathcal{B}$ are pairwise disjoint, nonempty,

and of finite measure. Define $\varphi_n : X \rightarrow U$ by

$$\varphi_n(x) = \begin{cases} y_{n,i} & \text{if } x \in A_{n,i} \cap \bar{A}_{n,j}, \|y_{n,i}\| \leq a_j, i = 1, \dots, m_n, j = 1, \dots, \bar{m}_n \\ \frac{a_j}{\|y_{n,i}\|} y_{n,i} & \text{if } x \in A_{n,i} \cap \bar{A}_{n,j}, \|y_{n,i}\| > a_j, i = 1, \dots, m_n, j = 1, \dots, \bar{m}_n \\ \vartheta_y & x \in X \setminus \left(\left(\bigcup_{i=1}^{m_n} A_{n,i} \right) \cap \left(\bigcup_{j=1}^{\bar{m}_n} \bar{A}_{n,j} \right) \right) \end{cases}$$

Clearly $\varphi_n(x) \in U$, $\forall x \in X$, since U is a conic segment. An equivalent

$$\text{way to express } \varphi_n \text{ is that } \varphi_n(x) = \begin{cases} \psi_n(x) & \text{if } \|\psi_n(x)\| \leq \phi_n(x) \\ \frac{\phi_n(x)}{\|\psi_n(x)\|} \psi_n(x) & \text{if } \|\psi_n(x)\| > \phi_n(x) \end{cases},$$

$\forall x \in X$. Then, we have $\|\varphi_n(x)\| \leq \phi_n(x) \leq \mathcal{P} \circ f(x)$, $\forall x \in X$. Furthermore, $\lim_{n \in \mathbb{N}} \varphi_n(x) = f(x)$, $\forall x \in X$.

This completes the proof of the lemma. $\square$

Thus, we have generalized Littlewood's three principles: "every Lebesgue measurable set in $\mathbb{R}$ is almost a finite union of open intervals" (Proposition 11.23); "every measurable function taking value in a separable normed linear spaces is almost a simple function" (Propositions 11.65 and 11.66); "every convergent sequence of measurable functions taking value in a separable metric space is nearly uniformly convergent" (Egoroff's Theorem 11.55).

11.4 Integration

Proposition 11.69 *Let $\mathcal{X}$ be a normed linear space, $\mathcal{B}_B(\mathcal{X})$ be the collection of Borel sets on $\mathcal{X}$. A representation of $\mathcal{X}$ is the collection $R := \{(x_\alpha, U_\alpha) \mid \alpha \in \Lambda\}$, where Λ is a finite index set, $(U_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{B}_B(\mathcal{X})$ are nonempty and pairwise disjoint, $\bigcup_{\alpha \in \Lambda} U_\alpha = \mathcal{X}$, $x_\alpha \in U_\alpha$, and $\|x_\alpha\| < \inf_{x \in U_\alpha} \|x\| + 1$, $\forall \alpha \in \Lambda$. The set of all representations of $\mathcal{X}$ is denoted $\mathcal{R}(\mathcal{X})$. Introduce a relation $\preceq$ on $\mathcal{R}(\mathcal{X})$ by $\forall R_1 := \{(x_\alpha, U_\alpha) \mid \alpha \in \Lambda\}, R_2 := \{(y_\beta, V_\beta) \mid \beta \in \Gamma\} \in \mathcal{R}(\mathcal{X})$, we will say $R_1 \preceq R_2$ if, $\forall \alpha \in \Lambda$, $\exists \beta \in \Gamma$ such that $x_\alpha = y_\beta$, and, $\forall \beta \in \Gamma$, $\exists \alpha \in \Lambda$ such that $V_\beta \subseteq U_\alpha$. Then, $\mathcal{I}(\mathcal{X}) := (\mathcal{R}(\mathcal{X}), \preceq)$ is a directed system and $\preceq$ is an antisymmetric partial ordering on $\mathcal{R}(\mathcal{X})$. $\mathcal{I}(\mathcal{X})$ is said to be the integration system on $\mathcal{X}$.*

Proof Clearly, $R_0 := \{(\vartheta_{\mathcal{X}}, \mathcal{X})\} \in \mathcal{R}(\mathcal{X}) \neq \emptyset$. Clearly, $\preceq$ is reflexive and transitive. $\forall R_1 := \{(x_\alpha, U_\alpha) \mid \alpha \in \Lambda\}, R_2 := \{(y_\beta, V_\beta) \mid \beta \in \Gamma\} \in \mathcal{R}(\mathcal{X})$ with $R_1 \preceq R_2$ and $R_2 \preceq R_1$, $\forall \alpha \in \Lambda$, $\exists \beta \in \Gamma$ such that $U_\alpha \subseteq V_\beta$, by $R_2 \preceq R_1$. By $R_1 \preceq R_2$, $\exists \bar{\alpha} \in \Lambda$ such that $V_\beta \subseteq U_{\bar{\alpha}}$. Then, we have $U_\alpha \subseteq V_\beta \subseteq U_{\bar{\alpha}}$. By $R_1 \in \mathcal{R}(\mathcal{X})$, $(U_{\bar{\alpha}})_{\bar{\alpha} \in \Lambda}$ are nonempty and pairwise disjoint. Then, we must have $\alpha = \bar{\alpha}$ and $U_\alpha = V_\beta$. By $R_1 \preceq R_2$, $\exists \bar{\beta} \in \Gamma$ such that $x_\alpha = y_{\bar{\beta}}$. Then, $U_\alpha = V_\beta \ni x_\alpha = y_{\bar{\beta}} \in V_{\bar{\beta}}$. By $R_2 \in \mathcal{R}(\mathcal{X})$, $(V_{\bar{\beta}})_{\bar{\beta} \in \Gamma}$ are nonempty and pairwise disjoint. Then, we must have $\bar{\beta} = \beta$ and $x_\alpha = y_\beta$. Then, $(x_\alpha, U_\alpha) \in R_2$. Hence, by the arbitrariness of α ,

we have $R_1 \subseteq R_2$. By an argument that is similar to the above, we have $R_2 \subseteq R_1$. Therefore, $R_1 = R_2$. This shows that $\preceq$ is an antisymmetric partial ordering on $\mathcal{R}(\mathcal{X})$.

$\forall R_1 := \{(x_\alpha, U_\alpha) \mid \alpha \in \Lambda\}, R_2 := \{(y_\beta, V_\beta) \mid \beta \in \Gamma\} \in \mathcal{R}(\mathcal{X})$, define $R_3 \in \mathcal{R}(\mathcal{X})$ as follows. $\forall \alpha \in \Lambda, \forall \beta \in \Gamma$, let $W_{\alpha,\beta} := U_\alpha \cap V_\beta \in \mathcal{B}_B(\mathcal{X})$. We will distinguish six exhaustive and mutually exclusive cases.

Case 1: $W_{\alpha,\beta} = \emptyset$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta}^{(2)} = \emptyset$.

Case 2: $W_{\alpha,\beta} \neq \emptyset, x_\alpha \notin W_{\alpha,\beta}$, and $y_\beta \notin W_{\alpha,\beta}$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta} \in \mathcal{B}_B(\mathcal{X})$, $z_{\alpha,\beta}^{(1)} \in W_{\alpha,\beta}^{(1)}$ such that $\|z_{\alpha,\beta}^{(1)}\| < \inf_{x \in W_{\alpha,\beta}^{(1)}} \|x\| + 1$, and $W_{\alpha,\beta}^{(2)} = \emptyset$.

Case 3: $x_\alpha \in W_{\alpha,\beta}$ and $y_\beta \notin W_{\alpha,\beta}$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta} \in \mathcal{B}_B(\mathcal{X})$, $z_{\alpha,\beta}^{(1)} = x_\alpha \in W_{\alpha,\beta}^{(1)}$, and $W_{\alpha,\beta}^{(2)} = \emptyset$. Then, $\|z_{\alpha,\beta}^{(1)}\| = \|x_\alpha\| < \inf_{x \in U_\alpha} \|x\| + 1 \leq \inf_{x \in W_{\alpha,\beta}^{(1)}} \|x\| + 1$.

Case 4: $x_\alpha \notin W_{\alpha,\beta}$ and $y_\beta \in W_{\alpha,\beta}$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta} \in \mathcal{B}_B(\mathcal{X})$, $z_{\alpha,\beta}^{(1)} = y_\beta \in W_{\alpha,\beta}^{(1)}$, and $W_{\alpha,\beta}^{(2)} = \emptyset$. Then, $\|z_{\alpha,\beta}^{(1)}\| = \|y_\beta\| < \inf_{x \in V_\beta} \|x\| + 1 \leq \inf_{x \in W_{\alpha,\beta}^{(1)}} \|x\| + 1$.

Case 5: $x_\alpha = y_\beta \in W_{\alpha,\beta}$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta} \in \mathcal{B}_B(\mathcal{X})$, $z_{\alpha,\beta}^{(1)} = x_\alpha \in W_{\alpha,\beta}^{(1)}$, and $W_{\alpha,\beta}^{(2)} = \emptyset$. Then, $\|z_{\alpha,\beta}^{(1)}\| = \|x_\alpha\| < \inf_{x \in U_\alpha} \|x\| + 1 \leq \inf_{x \in W_{\alpha,\beta}^{(1)}} \|x\| + 1$.

Case 6: $x_\alpha \in W_{\alpha,\beta}, y_\beta \in W_{\alpha,\beta}$, and $x_\alpha \neq y_\beta$. Let $\delta = \|x_\alpha - y_\beta\| > 0$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta} \cap \mathcal{B}_X(x_\alpha, \delta/2) \in \mathcal{B}_B(\mathcal{X})$, $z_{\alpha,\beta}^{(1)} = x_\alpha \in W_{\alpha,\beta}^{(1)}$, $W_{\alpha,\beta}^{(2)} = W_{\alpha,\beta} \cap (\mathcal{B}_X(x_\alpha, \delta/2))^\sim \in \mathcal{B}_B(\mathcal{X})$, and $z_{\alpha,\beta}^{(2)} = y_\beta \in W_{\alpha,\beta}^{(2)}$. Then, $\|z_{\alpha,\beta}^{(1)}\| = \|x_\alpha\| < \inf_{x \in U_\alpha} \|x\| + 1 \leq \inf_{x \in W_{\alpha,\beta}^{(1)}} \|x\| + 1$ and $\|z_{\alpha,\beta}^{(2)}\| = \|y_\beta\| < \inf_{x \in V_\beta} \|x\| + 1 \leq \inf_{x \in W_{\alpha,\beta}^{(2)}} \|x\| + 1$.

Define $R_3 := \{(z_{\alpha,\beta}^{(i)}, W_{\alpha,\beta}^{(i)}) \mid \alpha \in \Lambda, \beta \in \Gamma, i = 1, 2, W_{\alpha,\beta}^{(i)} \neq \emptyset\}$. Clearly, $R_3 \in \mathcal{R}(\mathcal{X})$, $R_1 \preceq R_3$, and $R_2 \preceq R_3$. Therefore, $\mathcal{J}(\mathcal{X})$ is a directed system. This completes the proof of the proposition. $\square$

Definition 11.70 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a normed linear space with the integration system $\mathcal{J}(\mathcal{Y}) := (\mathcal{R}(\mathcal{Y}), \preceq)$, and $f : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. $\forall R = \{(y_\alpha, U_\alpha) \mid \alpha \in \Lambda\} \in \mathcal{R}(\mathcal{Y})$, define $F_R := \sum_{\alpha \in \Lambda} y_\alpha \mu(f_{\text{inv}}(U_\alpha)) \in \mathcal{Y}$. This defines a net $(F_R)_{R \in \mathcal{J}(\mathcal{Y})}$. f is said to be integrable if the net admits a limit in $\mathcal{Y}$. In this case, $\lim_{R \in \mathcal{J}(\mathcal{Y})} F_R \in \mathcal{Y}$ is said to be the integral of f over $\mathcal{X}$ and denoted by $\int_X f \, d\mu$ or $\int_X f(x) \, d\mu(x)$. When $\mathcal{Y} = \mathbb{R}$, we will denote $\int_X f \, d\mu := \int_X f(x) \, d\mu(x) := \lim_{R \in \mathcal{J}(\mathbb{R})} F_R \in \mathbb{R}_e$ whenever the limit exists, which is said to be the integral of f over $\mathcal{X}$.

Definition 11.71 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space with $\mu(X) = \infty$, $\mathcal{Y}$ be a normed linear space, and $f : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Define the set $\mathcal{M}(\mathcal{X}) := \{A \in \mathcal{B} \mid \mu(A) < +\infty\}$. Clearly, $\mathcal{M}(\mathcal{X}) := (\mathcal{M}(\mathcal{X}), \subseteq)$ is

a directed system. $\forall A \in \mathcal{M}(\mathcal{X})$, let $(A, \mathcal{B}_A, \mu_A)$ be the finite measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. Define a net $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ by $F_A := \int_A f|_A d\mu_A \in \mathcal{Y}$, $\forall A \in \mathcal{M}(\mathcal{X})$. f is said to be integrable if the net is well defined and admits a limit in $\mathcal{Y}$. In this case, $\lim_{A \in \mathcal{M}(\mathcal{X})} F_A \in \mathcal{Y}$ is said to be the integral of f over $\mathcal{X}$ and denoted by $\int_{\mathcal{X}} f d\mu$ or $\int_{\mathcal{X}} f(x) d\mu(x)$. When $\mathcal{Y} = \mathbb{R}$, we will allow the net $(F_A)_{A \in \mathcal{M}(\mathcal{X})} \subseteq \mathbb{R}_e$ and denote $\int_{\mathcal{X}} f d\mu := \int_{\mathcal{X}} f(x) d\mu(x) := \lim_{A \in \mathcal{M}(\mathcal{X})} F_A \in \mathbb{R}_e$ whenever the limit exists, which is said to be the integral of f over $\mathcal{X}$.

Proposition 11.72 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a normed linear space, $f : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, and $\tilde{\mathcal{X}} := (X, \tilde{\mathcal{B}}, \tilde{\mu})$ be a finite measure space that is an extension of $\mathcal{X}$, i. e., $\mathcal{B} \subseteq \tilde{\mathcal{B}}$ and $\mu = \tilde{\mu}|_{\mathcal{B}}$. Then, $\int_{\mathcal{X}} f d\mu = \int_{\tilde{\mathcal{X}}} f d\tilde{\mu}$, whenever one of the integrals exists.

Proof Let $\mathcal{J}(\mathcal{Y})$ be the integration system on $\mathcal{Y}$ as defined in Proposition 11.69, $(F_R)_{R \in \mathcal{J}(\mathcal{Y})}$ be the net for $\int_{\mathcal{X}} f d\mu$, and $(\bar{F}_R)_{R \in \mathcal{J}(\mathcal{Y})}$ be the net for $\int_{\tilde{\mathcal{X}}} f d\tilde{\mu}$, as defined in Definition 11.70. $\forall R := \{(y_\alpha, U_\alpha) \mid \alpha \in \Lambda\} \in \mathcal{J}(\mathcal{Y})$, $\forall \alpha \in \Lambda$, we have $f_{\text{inv}}(U_\alpha) \in \mathcal{B} \subseteq \tilde{\mathcal{B}}$ since f is $\mathcal{B}$ -measurable, and $\mu(f_{\text{inv}}(U_\alpha)) = \tilde{\mu}(f_{\text{inv}}(U_\alpha)) \in \mathbb{R}$ since $\mu = \tilde{\mu}|_{\mathcal{B}}$. Then, $F_R = \sum_{\alpha \in \Lambda} y_\alpha \mu(f_{\text{inv}}(U_\alpha)) = \sum_{\alpha \in \Lambda} y_\alpha \tilde{\mu}(f_{\text{inv}}(U_\alpha)) = \bar{F}_R \in \mathcal{Y}$. Hence, $\int_{\mathcal{X}} f d\mu = \lim_{R \in \mathcal{J}(\mathcal{Y})} F_R = \lim_{R \in \mathcal{J}(\mathcal{Y})} \bar{F}_R = \int_{\tilde{\mathcal{X}}} f d\tilde{\mu}$, whenever one of the integrals exists. This completes the proof of the proposition. $\square$

Lemma 11.73 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a normed linear space, $f : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\hat{X} \in \mathcal{B}$, and $\hat{\mathcal{X}} := (\hat{X}, \hat{\mathcal{B}}, \hat{\mu})$ be the finite measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. Assume that $\mu(X \setminus \hat{X}) = 0$. Then, $\int_{\mathcal{X}} f d\mu = \int_{\hat{\mathcal{X}}} f|_{\hat{X}} d\hat{\mu}$, whenever one of the integrals exists.

Proof Let $\mathcal{J}(\mathcal{Y})$ be the integration system on $\mathcal{Y}$ as defined in Proposition 11.69, $(F_R)_{R \in \mathcal{J}(\mathcal{Y})}$ be the net for $\int_{\mathcal{X}} f d\mu$, and $(\hat{F}_R)_{R \in \mathcal{J}(\mathcal{Y})}$ be the net for $\int_{\hat{\mathcal{X}}} f|_{\hat{X}} d\hat{\mu}$, as defined in Definition 11.70. $\forall R := \{(y_\alpha, U_\alpha) \mid \alpha \in \Lambda\} \in \mathcal{J}(\mathcal{Y})$, $\forall \alpha \in \Lambda$, we have $f_{\text{inv}}(U_\alpha) \cap \hat{X} = f|_{\hat{X}} \text{inv}(U_\alpha) \in \hat{\mathcal{B}} \subseteq \mathcal{B}$ since f is $\mathcal{B}$ -measurable, and $0 \leq \mu(f_{\text{inv}}(U_\alpha) \cap (X \setminus \hat{X})) \leq \mu(X \setminus \hat{X}) = 0$. Then, $F_R = \sum_{\alpha \in \Lambda} y_\alpha \mu(f_{\text{inv}}(U_\alpha)) = \sum_{\alpha \in \Lambda} y_\alpha (\mu(f_{\text{inv}}(U_\alpha) \cap \hat{X}) + \mu(f_{\text{inv}}(U_\alpha) \cap (X \setminus \hat{X}))) = \sum_{\alpha \in \Lambda} y_\alpha \mu(f|_{\hat{X}} \text{inv}(U_\alpha)) = \sum_{\alpha \in \Lambda} y_\alpha \hat{\mu}(f|_{\hat{X}} \text{inv}(U_\alpha)) = \hat{F}_R \in \mathcal{Y}$. Hence, $\int_{\mathcal{X}} f d\mu = \lim_{R \in \mathcal{J}(\mathcal{Y})} F_R = \lim_{R \in \mathcal{J}(\mathcal{Y})} \hat{F}_R = \int_{\hat{\mathcal{X}}} f|_{\hat{X}} d\hat{\mu}$, whenever one of the integrals exists. This completes the proof of the proposition. $\square$

Proposition 11.74 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space with $\mu(X) = +\infty$, $\mathcal{Y}$ be a normed linear space, $f : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, and $\tilde{\mathcal{X}} := (X, \tilde{\mathcal{B}}, \tilde{\mu})$ be the completion of $\mathcal{X}$. Then, $\int_{\mathcal{X}} f d\mu = \int_{\tilde{\mathcal{X}}} f d\tilde{\mu}$, whenever one of the integrals exists.

Proof Let $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X f d\mu$ and $(\bar{F}_{\bar{A}})_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})}$ be the net for $\int_X f d\bar{\mu}$ as defined in Definition 11.71. Clearly, $\mathcal{M}(\mathcal{X}) \subseteq \mathcal{M}(\bar{\mathcal{X}})$, since $\bar{\mathcal{X}}$ is the completion of $\mathcal{X}$. We will distinguish two exhaustive cases: Case 1: $\int_X f d\bar{\mu}$ exists; Case 2: $\int_X f d\mu$ exists.

Case 1: $\int_X f d\bar{\mu}$ exists. Then, the net $(\bar{F}_{\bar{A}})_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})}$ is well defined. $\forall A \in \mathcal{M}(\mathcal{X})$, $A \in \mathcal{M}(\bar{\mathcal{X}})$. By Proposition 11.72, $F_A = \int_A f|_A d\mu_A = \int_A f|_A d\bar{\mu}_A = \bar{F}_{\bar{A}}$, where $(A, \mathcal{B}_A, \mu_A)$ is the finite measure subspace of $\mathcal{X}$ and $(A, \bar{\mathcal{B}}_A, \bar{\mu}_A)$ is the finite complete measure subspace of $\bar{\mathcal{X}}$. Then, $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ is well defined. $\forall \bar{A} \in \mathcal{M}(\bar{\mathcal{X}})$, we have $\bar{A} \in \bar{\mathcal{B}}$ with $\bar{\mu}(\bar{A}) < +\infty$. By Proposition 11.12 and its proof, $\exists \tilde{A}, \tilde{B}, \tilde{C} \subseteq X$ with $\bar{A} = \tilde{A} \cup \tilde{B}$, $\tilde{B}, \tilde{C} \in \mathcal{B}$, $\tilde{A} \subseteq \tilde{C}$, $\mu(\tilde{C}) = 0$, and $\bar{\mu}(\bar{A}) = \mu(\tilde{B})$. Then, $\bar{A} \subseteq \tilde{B} \cup \tilde{C} =: A \in \mathcal{B}$ and $\mu(A) \leq \mu(\tilde{B}) + \mu(\tilde{C}) = \mu(\tilde{B}) = \bar{\mu}(\bar{A}) < +\infty$. This shows that $A \in \mathcal{M}(\mathcal{X})$ and $\bar{A} \subseteq A$. Hence, $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ is a subnet of $(\bar{F}_{\bar{A}})_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})}$. By Proposition 3.70, $\int_X f d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_A = \lim_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})} \bar{F}_{\bar{A}} = \int_X f d\bar{\mu}$.

Case 2: $\int_X f d\mu$ exists. Then, the net $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ is well defined. $\forall \bar{A} \in \mathcal{M}(\bar{\mathcal{X}})$, we have $\bar{A} \in \bar{\mathcal{B}}$ with $\bar{\mu}(\bar{A}) < +\infty$. By Proposition 11.12 and its proof, $\exists \tilde{A}, \tilde{B}, \tilde{C} \subseteq X$ with $\bar{A} = \tilde{A} \cup \tilde{B}$, $\tilde{B}, \tilde{C} \in \mathcal{B}$, $\tilde{A} \subseteq \tilde{C}$, $\mu(\tilde{C}) = 0$, and $\bar{\mu}(\bar{A}) = \mu(\tilde{B})$. Let $\tilde{B} =: A \in \mathcal{B}$. Then, $0 \leq \bar{\mu}(\bar{A} \setminus A) = \bar{\mu}(\tilde{A} \setminus \tilde{B}) \leq \bar{\mu}(\tilde{C}) = \mu(\tilde{C}) = 0$ and $\mu(A) = \bar{\mu}(\bar{A}) < +\infty$. This shows that $A \in \mathcal{M}(\mathcal{X})$ and $A \subseteq \bar{A}$. Let $\mathcal{A} := (A, \mathcal{B}_A, \mu_A)$ be the finite measure subspace of $\mathcal{X}$, $\bar{\mathcal{A}} := (\bar{A}, \bar{\mathcal{B}}_{\bar{A}}, \bar{\mu}_{\bar{A}})$ be the finite complete measure subspace of $\bar{\mathcal{X}}$, and $\hat{\mathcal{A}} := (A, \bar{\mathcal{B}}_A, \bar{\mu}_A)$ be the finite complete measure subspace of $\bar{\mathcal{A}}$. By Lemma 11.73, we have $\bar{F}_{\bar{A}} = \int_{\bar{A}} f|_{\bar{A}} d\bar{\mu}_{\bar{A}} = \int_A f|_A d\bar{\mu}_A$, whenever one of them exists. Since $\bar{\mathcal{X}}$ is the completion of $\mathcal{X}$, then $\hat{\mathcal{A}}$ is an extension of $\mathcal{A}$. By Proposition 11.72, we have $\int_A f|_A d\bar{\mu}_A = \int_A f|_A d\mu_A = F_A$. Hence, $\bar{F}_{\bar{A}} = F_A$. Then, the net $(\bar{F}_{\bar{A}})_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})}$ is well defined.

Fix any open set U ($U \subseteq \mathbb{R}_e$ if $\mathcal{Y} = \mathbb{R}$ or $U \subseteq \mathcal{Y}$ if $\mathcal{Y} \neq \mathbb{R}$) with $\int_X f d\mu \in U$. Since $\lim_{A \in \mathcal{M}(\mathcal{X})} F_A = \int_X f d\mu \in U$, then $\exists A_0 \in \mathcal{M}(\mathcal{X})$ such that $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have $F_A \in U$. Note that $A_0 \in \mathcal{B} \subseteq \bar{\mathcal{B}}$ and $\mu(A_0) = \bar{\mu}(A_0) < +\infty$. Then, $A_0 \in \mathcal{M}(\bar{\mathcal{X}})$. $\forall \bar{A} \in \mathcal{M}(\bar{\mathcal{X}})$ with $A_0 \subseteq \bar{A}$, then $\bar{A} \in \bar{\mathcal{B}}$ and $\bar{\mu}(\bar{A}) < +\infty$. By Proposition 11.12 and its proof, $\exists \tilde{A}, \tilde{B}, \tilde{C} \subseteq X$ with $\bar{A} = \tilde{A} \cup \tilde{B}$, $\tilde{B}, \tilde{C} \in \mathcal{B}$, $\tilde{A} \subseteq \tilde{C}$, $\mu(\tilde{C}) = 0$, and $\bar{\mu}(\bar{A}) = \mu(\tilde{B})$. Let $A := A_0 \cup \tilde{B} \in \mathcal{B}$. Then, $0 \leq \bar{\mu}(\bar{A} \setminus A) = \bar{\mu}((\tilde{A} \cup \tilde{B}) \setminus (A_0 \cup \tilde{B})) \leq \bar{\mu}((\tilde{A} \cup \tilde{B}) \setminus \tilde{B}) \leq \bar{\mu}(\tilde{A} \setminus \tilde{B}) \leq \bar{\mu}(\tilde{C}) = \mu(\tilde{C}) = 0$, $A \subseteq \bar{A}$, and $\mu(A) \leq \mu(\tilde{B}) + \mu(A_0) = \bar{\mu}(\bar{A}) + \mu(A_0) < +\infty$. This shows that $A \in \mathcal{M}(\mathcal{X})$ and $A_0 \subseteq A \subseteq \bar{A}$. Let $\mathcal{A} := (A, \mathcal{B}_A, \mu_A)$ be the finite measure subspace of $\mathcal{X}$, $\bar{\mathcal{A}} := (\bar{A}, \bar{\mathcal{B}}_{\bar{A}}, \bar{\mu}_{\bar{A}})$ be the finite complete measure subspace of $\bar{\mathcal{X}}$, and $\hat{\mathcal{A}} := (A, \bar{\mathcal{B}}_A, \bar{\mu}_A)$ be the finite complete measure subspace of $\bar{\mathcal{A}}$. By Lemma 11.73, we have $\bar{F}_{\bar{A}} = \int_{\bar{A}} f|_{\bar{A}} d\bar{\mu}_{\bar{A}} = \int_A f|_A d\bar{\mu}_A$, whenever one of them exists. Since $\bar{\mathcal{X}}$ is the completion of $\mathcal{X}$, then $\hat{\mathcal{A}}$ is an extension of $\mathcal{A}$. By Proposition 11.72, we have $\int_A f|_A d\bar{\mu}_A = \int_A f|_A d\mu_A = F_A$. Hence, $\bar{F}_{\bar{A}} = F_A \in U$. Therefore, we have $\int_X f d\bar{\mu} = \lim_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})} \bar{F}_{\bar{A}} = \int_X f d\mu$.

Hence, in both cases, we have $\int_X f d\mu = \int_X f d\bar{\mu}$. This completes the proof of the proposition. $\square$

Proposition 11.75 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a normed linear space, $\phi : X \rightarrow \mathcal{Y}$ be a simple function in canonical representation, i. e., $\exists n \in \mathbb{Z}_+$, $\exists y_1, \dots, y_n \in \mathcal{Y}$, which are distinct and none equals to $\vartheta_{\mathcal{Y}}$, $\exists A_1, \dots, A_n \in \mathcal{B}$, which are nonempty, pairwise disjoint, and of finite measure, such that $\phi(x) = \sum_{i=1}^n y_i \chi_{A_i, X}(x)$, $\forall x \in X$. Then, $\int_X \phi d\mu = \sum_{i=1}^n y_i \mu(A_i) =: I \in \mathcal{Y}$.*

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu(X) < +\infty$; Case 2: $\mu(X) = +\infty$.

Case 1: $\mu(X) < +\infty$. Let $y_{n+1} := \vartheta_{\mathcal{Y}}$ and $A_{n+1} := X \setminus (\bigcup_{i=1}^n A_i) \in \mathcal{B}$. Let $\epsilon_0 := \min\{1, \min_{1 \leq i < j \leq n+1} \|y_i - y_j\|\} > 0$, by the assumption. Let $(\Phi_R)_{R \in \mathcal{J}(\mathcal{Y})}$ be the net as defined in Definition 11.70 for the simple function ϕ .

$\forall \epsilon \in (0, \epsilon_0/2) \subset \mathbb{R}$, let $U_i := \mathcal{B}_{\mathcal{Y}}(y_i, \epsilon) \in \mathcal{B}_{\mathcal{B}}(\mathcal{Y})$, $i = 1, \dots, n+1$, which are clearly pairwise disjoint and nonempty. We will distinguish two exhaustive and mutually exclusive subcases: Case 1a: $\mathcal{Y} = \bigcup_{i=1}^{n+1} U_i$; Case 1b: $\mathcal{Y} \supset \bigcup_{i=1}^{n+1} U_i$.

Case 1a: $\mathcal{Y} = \bigcup_{i=1}^{n+1} U_i$. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{Y} = \mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, \delta)$. Then, $\mathcal{Y}$ must be the trivial normed linear space, i. e., $\mathcal{Y}$ is a singleton set containing $\vartheta_{\mathcal{Y}}$. There is a single representation of $\mathcal{Y}$, which is $R_0 := \{(\vartheta_{\mathcal{Y}}, \mathcal{Y})\}$. Then, $\Phi_{R_0} = \vartheta_{\mathcal{Y}}$. Clearly, we must have $n = 0$, since $y_1, \dots, y_n \in \mathcal{Y}$ are distinct and none equals to $\vartheta_{\mathcal{Y}}$. Then, $I = \vartheta_{\mathcal{Y}}$. Hence, $\forall R \in \mathcal{J}(\mathcal{Y})$ with $R_0 \preceq R$, $\|\Phi_R - I\| = \|\Phi_{R_0} - I\| = 0$.

Case 1b: $\mathcal{Y} \supset \bigcup_{i=1}^{n+1} U_i$. Let $U_{n+2} = \bigcap_{i=1}^{n+1} \widetilde{U}_i \in \mathcal{B}_{\mathcal{B}}(\mathcal{Y})$, which is nonempty. Let $y_{n+2} \in U_{n+2}$ be such that $\|y_{n+2}\| < \inf_{y \in U_{n+2}} \|y\| + 1$. Let $R_0 := \{(y_i, U_i) \mid i = 1, \dots, n+2\}$. It is easy to check that $R_0 \in \mathcal{J}(\mathcal{Y})$. $\forall \tilde{R} \in \mathcal{J}(\mathcal{Y})$ with $R_0 \preceq \tilde{R}$, then $\tilde{R} = \{(\tilde{y}_{\alpha}, \tilde{U}_{\alpha}) \mid \alpha \in \Lambda\}$, where Λ is a finite index set, $(\tilde{U}_{\alpha})_{\alpha \in \Lambda} \subseteq \mathcal{B}_{\mathcal{B}}(\mathcal{Y})$ are nonempty and pairwise disjoint, $\bigcup_{\alpha \in \Lambda} \tilde{U}_{\alpha} = \mathcal{Y}$, $\tilde{y}_{\alpha} \in \tilde{U}_{\alpha}$, and $\|\tilde{y}_{\alpha}\| < \inf_{y \in \tilde{U}_{\alpha}} \|y\| + 1$, $\forall \alpha \in \Lambda$. $\forall i \in \{1, \dots, n+2\}$, by $R_0 \preceq \tilde{R}$, $\exists! \alpha_i \in \Lambda$ such that $y_i = \tilde{y}_{\alpha_i}$ and, by $\tilde{R} \in \mathcal{J}(\mathcal{Y})$, $\exists! \bar{\alpha}_i \in \Lambda$ such that $y_i \in \tilde{U}_{\bar{\alpha}_i}$. Since $(\tilde{U}_{\alpha})_{\alpha \in \Lambda} \subseteq \mathcal{B}_{\mathcal{B}}(\mathcal{Y})$ are nonempty and pairwise disjoint and $\tilde{y}_{\alpha} \in \tilde{U}_{\alpha}$, $\forall \alpha \in \Lambda$, then $\alpha_i = \bar{\alpha}_i$. Again by $R_0 \preceq \tilde{R}$, we have $\tilde{U}_{\alpha_i} \subseteq U_i$. Hence, $\alpha_1, \dots, \alpha_{n+2}$ are distinct. Let $\bar{\Lambda} := \{\alpha_1, \dots, \alpha_{n+2}\}$. $\forall \alpha \in \Lambda \setminus \bar{\Lambda}$, $\phi_{\text{inv}}(\tilde{U}_{\alpha}) = \emptyset$. Then, $\Phi_{\tilde{R}} = \sum_{\alpha \in \Lambda} \tilde{y}_{\alpha} \mu(\phi_{\text{inv}}(\tilde{U}_{\alpha})) = \sum_{i=1}^{n+2} \tilde{y}_{\alpha_i} \mu(\phi_{\text{inv}}(\tilde{U}_{\alpha_i})) = \sum_{i=1}^{n+2} y_i \mu(\phi_{\text{inv}}(\tilde{U}_{\alpha_i}))$. Note that $\forall i \in \{1, \dots, n+1\}$, $\phi_{\text{inv}}(\tilde{U}_{\alpha_i}) = A_i$ and $\mu(A_i) \leq \mu(X) < +\infty$. Furthermore, $\phi_{\text{inv}}(\tilde{U}_{\alpha_{n+2}}) = \emptyset$. Hence, $\Phi_{\tilde{R}} = \sum_{i=1}^n y_i \mu(A_i) = I$. Then, we have $\|\Phi_{\tilde{R}} - I\| = 0$.

In both subcases, we have $\exists R_0 \in \mathcal{J}(\mathcal{Y})$, $\forall R \in \mathcal{J}(\mathcal{Y})$ with $R_0 \preceq R$, we have $\|\Phi_R - I\| = 0 < \epsilon$. Hence, $\int_X \phi d\mu = \lim_{R \in \mathcal{J}(\mathcal{Y})} \Phi_R = I$.

Case 2: $\mu(X) = +\infty$. $\forall A \in \mathcal{M}(\mathcal{X})$, $\mu(A) < +\infty$. $\phi|_A : A \rightarrow \mathcal{Y}$ is a simple function. By Case 1, $\Phi_A = \int_A \phi|_A d\mu_A \in \mathcal{Y}$ is well-defined, where $(A, \mathcal{B}_A, \mu_A)$ is the finite measure subspace of $\mathcal{X}$. Hence, the net $(\Phi_A)_{A \in \mathcal{M}(\mathcal{X})}$ is well-defined. Take $A_0 := \bigcup_{i=1}^n A_i \in \mathcal{B}$. Then, $\mu(A_0) = \sum_{i=1}^n \mu(A_i) < +\infty$. Therefore, $A_0 \in \mathcal{M}(\mathcal{X})$. $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, by Case 1, $\Phi_A = \sum_{i=1}^n y_i \mu(A_i) = I$. Then, $\int_X \phi d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} \Phi_A = I$.

This completes the proof of the proposition. $\square$

Clearly, the above result holds for simple functions given in any form, not necessarily in canonical representation. It also shows that integral of simple functions are linear, that is, $\forall$ simple functions ϕ_1 and ϕ_2 , $\forall c \in \mathbb{K}$, we have $\int_X (\phi_1 + \phi_2) d\mu = \int_X \phi_1 d\mu + \int_X \phi_2 d\mu$ and $\int_X (c\phi_1) d\mu = c \int_X \phi_1 d\mu$.

Lemma 11.76 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a separable Banach space, $\phi_n : X \rightarrow \mathcal{Y}$ be a simple function in canonical representation, $\forall n \in \mathbb{N}$, and $f : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Assume that $\lim_{n \in \mathbb{N}} \phi_n = f$ a.e. in $\mathcal{X}$ and $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|\phi_n(x)\| \leq M$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Then, f is integrable over X and $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n d\mu \in \mathcal{Y}$.*

Proof Let $E := \{x \in X \mid (\phi_n(x))_{n=1}^\infty \text{ does not converge to } f(x)\}$. Then, $E \in \mathcal{B}$ and $\mu(E) = 0$. Let $(F_R)_{R \in \mathcal{I}(\mathcal{Y})}$ be the net for $\int_X f d\mu$ as defined in Definition 11.70.

$\forall \epsilon \in (0, 1) \subset \mathbb{R}$, by Egoroff's Theorem 11.55, $\exists E_1 \in \mathcal{B}$ with $\mu(E_1) < \frac{\epsilon}{4M+2}$ such that $\left(\phi_n|_{X \setminus E_1}\right)_{n=1}^\infty$ converges uniformly to $f|_{X \setminus E_1}$. Clearly, $E \subseteq E_1$. Then, $\exists n_0 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq n_0$, $\forall x \in X \setminus E_1$, $\|\phi_n(x) - f(x)\| < \frac{\epsilon}{6\mu(X)+1} =: \bar{\epsilon}$. Fix any $n \geq n_0$. Let ϕ_n admit the canonical representation $\phi_n = \sum_{i=1}^{\bar{n}} y_i \chi_{A_i, X}$, where $\bar{n} \in \mathbb{Z}_+$, $y_1, \dots, y_{\bar{n}} \in \mathcal{Y}$ are distinct and none equals to $\vartheta_{\mathcal{Y}}$, and $A_1, \dots, A_{\bar{n}} \in \mathcal{B}$ are nonempty, pairwise disjoint, and of finite measure. Let $y_{\bar{n}+1} := \vartheta_{\mathcal{Y}}$ and $A_{\bar{n}+1} := X \setminus (\bigcup_{i=1}^{\bar{n}} A_i) \in \mathcal{B}$. Then, $X = \bigcup_{i=1}^{\bar{n}+1} A_i$ and the sets in the union are pairwise disjoint and of finite measure. Define $\bar{V}_i := \mathcal{B}_{\mathcal{Y}}(y_i, \bar{\epsilon})$, $i = 1, \dots, \bar{n} + 1$. Define $V_1 := \bar{V}_1 \in \mathcal{B}_{\mathcal{B}}(\mathcal{Y})$, $V_{i+1} := \bar{V}_{i+1} \setminus (\bigcup_{j=1}^i \bar{V}_j) \in \mathcal{B}_{\mathcal{B}}(\mathcal{Y})$, $i = 1, \dots, \bar{n}$. Let $V_{\bar{n}+2} := \mathcal{Y} \setminus (\bigcup_{j=1}^{\bar{n}+1} \bar{V}_j)$. Clearly, we may form a representation $\bar{R} := \{(\bar{y}_i, V_i) \mid i = 1, \dots, \bar{n} + 2, V_i \neq \emptyset\} \in \mathcal{I}(\mathcal{Y})$, where $\bar{y}_i \in V_i$, $i = 1, \dots, \bar{n} + 2$, are any vectors that satisfy the assumption of Proposition 11.69. $\forall R \in \mathcal{I}(\mathcal{Y})$ with $\bar{R} \preceq R$. Let $R := \{(\hat{y}_\gamma, \hat{U}_\gamma) \mid \gamma \in \Gamma\}$. $\forall \gamma \in \Gamma$, by $\bar{R} \preceq R$, $\exists! i_\gamma \in \{1, \dots, \bar{n} + 2\}$ such that $\hat{U}_\gamma \subseteq V_{i_\gamma}$. Define $\bar{\Gamma} := \{\gamma \in \Gamma \mid i_\gamma = \bar{n} + 2\}$. Let $A_{\gamma, i} := A_i \cap f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1) \in \mathcal{B}$, $i = 1, \dots, \bar{n} + 1$.

Claim 11.76.1 $\forall \gamma \in \bar{\Gamma}$, $f_{\text{inv}}(\hat{U}_\gamma) \subseteq E_1$ and $\left\| \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| \leq (M + 1) \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) = (M + 1) \mu(f_{\text{inv}}(\hat{U}_\gamma))$.

Proof of claim: Fix any $\gamma \in \bar{\Gamma}$. $\hat{U}_\gamma \subseteq V_{\bar{n}+2} = \mathcal{Y} \setminus (\bigcup_{j=1}^{\bar{n}+1} \bar{V}_j)$. Then, $\mathcal{B} \ni f_{\text{inv}}(\hat{U}_\gamma) \subseteq E_1$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $f_{\text{inv}}(\hat{U}_\gamma) \subseteq E$; Case 2: $f_{\text{inv}}(\hat{U}_\gamma) \cap (E_1 \setminus E) \neq \emptyset$.
Case 1: $f_{\text{inv}}(\hat{U}_\gamma) \subseteq E$. Then, $\mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) = 0$ and $\left\| \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| = 0 = (M+1)\mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1)$. Case 2: $f_{\text{inv}}(\hat{U}_\gamma) \cap (E_1 \setminus E) \neq \emptyset$. Then, $\exists x \in E_1 \setminus E$ such that $f(x) \in \hat{U}_\gamma$. Note that $\lim_{m \in \mathbb{N}} \phi_m(x) = f(x)$ and $\|\phi_m(x)\| \leq M$, $\forall m \in \mathbb{N}$. Then, by Propositions 7.21 and 3.66, we have $\|f(x)\| = \lim_{m \in \mathbb{N}} \|\phi_m(x)\| \leq M$. Then, by $R \in \mathcal{I}(\mathcal{Y})$, $\|\hat{y}_\gamma\| < \inf_{y \in \hat{U}_\gamma} \|y\| + 1 \leq M+1$. Therefore, $\left\| \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| \leq (M+1)\mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1)$. This completes the proof of the claim. $\square$

$\forall \gamma \in \Gamma \setminus \bar{\Gamma}$, $\hat{U}_\gamma \subseteq V_{i_\gamma}$, where $i_\gamma \in \{1, \dots, \bar{n}+1\}$. $\forall i \in \{1, \dots, \bar{n}+1\}$ with $A_{\gamma,i} \neq \emptyset$, then $\exists x \in A_{\gamma,i}$, $\phi_n(x) = y_i$ and $f(x) \in \mathcal{B}_\mathcal{Y}(y_i, \bar{\epsilon}) \cap \hat{U}_\gamma \subseteq \mathcal{B}_\mathcal{Y}(y_i, \bar{\epsilon}) \cap V_{i_\gamma} \subseteq \mathcal{B}_\mathcal{Y}(y_i, \bar{\epsilon}) \cap \bar{V}_{i_\gamma} \subseteq \mathcal{B}_\mathcal{Y}(y_i, \bar{\epsilon}) \cap \mathcal{B}_\mathcal{Y}(y_{i_\gamma}, \bar{\epsilon})$. Note that $\hat{y}_\gamma \in \hat{U}_\gamma \subseteq \mathcal{B}_\mathcal{Y}(y_{i_\gamma}, \bar{\epsilon})$. Therefore, we have $\|\hat{y}_\gamma - y_i\| \leq \|\hat{y}_\gamma - y_{i_\gamma}\| + \|y_{i_\gamma} - f(x)\| + \|f(x) - y_i\| < 3\bar{\epsilon}$.

Note that $F_R = \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma))$ and $\int_X \phi_n d\mu = \sum_{i=1}^{\bar{n}} y_i \mu(A_i)$, by Proposition 11.75. Then,

$$\begin{aligned}
\|F_R - \int_X \phi_n d\mu\| &= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma)) - \sum_{i=1}^{\bar{n}} y_i \mu(A_i) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma (\mu(f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1)) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1)) \right. \\
&\quad \left. - \sum_{i=1}^{\bar{n}+1} y_i (\mu(A_i \cap E_1) + \mu(A_i \cap (X \setminus E_1))) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \left(\sum_{i=1}^{\bar{n}+1} \mu(A_{\gamma,i}) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right) - \sum_{i=1}^{\bar{n}+1} y_i (\mu(A_i \cap E_1) \right. \\
&\quad \left. + \mu(A_{\gamma,i})) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} \hat{y}_\gamma \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) - \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i \cap E_1) \right. \\
&\quad \left. - \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} y_i \mu(A_{\gamma,i}) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - y_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\|
\end{aligned}$$

$$\begin{aligned}
& - \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i \cap E_1) \Big\| \\
= & \left\| \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - y_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - y_i) \mu(A_{\gamma,i}) \right. \\
& \left. + \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) - \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i \cap E_1) \right\| \\
= & \left\| \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - y_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right. \\
& \left. - \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i \cap E_1) \right\| \\
\leq & \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} \|\hat{y}_\gamma - y_i\| \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \left\| \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| \\
& + \sum_{i=1}^{\bar{n}+1} \|y_i\| \mu(A_i \cap E_1) \\
\leq & 3\bar{\epsilon} \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \left\| \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| \\
& + \sum_{\gamma \in \bar{\Gamma}} \left\| \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| + \sum_{i=1}^{\bar{n}+1} \|y_i\| \mu(A_i \cap E_1) \\
\leq & \frac{3\mu(X)\epsilon}{6\mu(X)+1} + \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} (M+1)\mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \\
& + \sum_{\gamma \in \bar{\Gamma}} (M+1)\mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) + \sum_{i=1}^{\bar{n}+1} M\mu(A_i \cap E_1) \\
< & \epsilon/2 + \sum_{\gamma \in \Gamma} (M+1)\mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) + \sum_{i=1}^{\bar{n}+1} M\mu(A_i \cap E_1) \\
= & \epsilon/2 + (2M+1)\mu(E_1) < \epsilon
\end{aligned}$$

where the seventh equality follows from the fact that $\forall \gamma \in \bar{\Gamma}$, $A_{\gamma,i} \subseteq f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1) = \emptyset$; and the third inequality follows from the fact that $\forall \gamma \in \Gamma \setminus \bar{\Gamma}$, $\hat{y}_\gamma \in \hat{U}_\gamma \subseteq V_{i_\gamma} \subseteq \bar{V}_{i_\gamma} = \mathcal{B}_y(y_{i_\gamma}, \bar{\epsilon})$ and $\|\hat{y}_\gamma\| \leq \|y_{i_\gamma}\| + \|\hat{y}_\gamma - y_{i_\gamma}\| < M + \bar{\epsilon} < M + 1$.

In summary, we have shown that $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq n_0$, $\exists \bar{R} \in \mathcal{I}(\mathcal{Y})$ such that $\forall R \in \mathcal{I}(\mathcal{Y})$ with $\bar{R} \preceq R$, we have $\|F_R - \int_X \phi_n d\mu\| < \epsilon$. Then, the net $(F_R)_{R \in \mathcal{I}(\mathcal{Y})}$ is a Cauchy net. By Proposi-

tion 4.44, the net admits a limit $\int_X f \, d\mu \in \mathcal{Y}$. By Propositions 4.30, 3.66, and 3.67, we have $\|\int_X f \, d\mu - \int_X \phi_n \, d\mu\| = \lim_{R \in \mathcal{I}(\mathcal{Y})} \|F_R - \int_X \phi_n \, d\mu\| \leq \epsilon$, $\forall n \geq n_0$. Hence, we have $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu$. This completes the proof of the lemma. $\square$

Theorem 11.77 (Bounded Convergence Theorem) *Let $(X, \mathcal{B}, \mu) =: \mathcal{X}$ be a finite measure space, $\mathcal{Y}$ be a separable Banach space, $f_n : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $f : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$ and $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|f_n(x)\| \leq M$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Then, f is integrable over X and $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathcal{Y}$.*

Proof Let $U := \overline{\mathcal{B}_{\mathcal{Y}}(\vartheta_{\mathcal{Y}}, M)}$, which is a conic segment. Then, $\forall n \in \mathbb{N}$, $f_n : X \rightarrow U \subseteq \mathcal{Y}$. By Proposition 11.66, there exists a sequence of simple functions $(\phi_{n,i})_{i=1}^{\infty}$, $\phi_{n,i} : X \rightarrow U$, $\forall i \in \mathbb{N}$, that converges to f_n a.e. in $\mathcal{X}$ and $\|\phi_{n,i}(x)\| \leq \|f_n(x)\| \leq M$, $\forall x \in X$, $\forall i \in \mathbb{N}$. By Lemma 11.76, $\int_X f_n \, d\mu = \lim_{i \in \mathbb{N}} \int_X \phi_{n,i} \, d\mu \in \mathcal{Y}$. By Proposition 11.58, we have $\lim_{i \in \mathbb{N}} \phi_{n,i} = f_n$ in measure in $\mathcal{X}$. Then, $\exists i_n \in \mathbb{N}$ such that $\|\int_X f_n \, d\mu - \int_X \phi_{n,i_n} \, d\mu\| < 2^{-n}$ and $\mu(E_n) := \mu(\{x \in X \mid \|f_n(x) - \phi_{n,i_n}(x)\| \geq 2^{-n}\}) < 2^{-n}$. Denote $\bar{\phi}_n := \phi_{n,i_n}$.

By Proposition 11.58, $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have $2^{-n} < \epsilon/2$ and $\mu(\bar{E}_n) := \mu(\{x \in X \mid \|f_n(x) - f(x)\| \geq \epsilon/2\}) < \epsilon/2$. Note that, by Lemma 11.43, $\mathcal{B} \ni F_n := \{x \in X \mid \|\bar{\phi}_n(x) - f(x)\| \geq \epsilon\} \subseteq \{x \in X \mid \|\bar{\phi}_n(x) - f_n(x)\| + \|f_n(x) - f(x)\| \geq \epsilon\} \subseteq E_n \cup \bar{E}_n$. Then, we have $\mu(F_n) \leq \mu(E_n) + \mu(\bar{E}_n) < \epsilon$. This implies that $\lim_{n \in \mathbb{N}} \bar{\phi}_n = f$ in measure in $\mathcal{X}$. By Proposition 11.57, there exists a subsequence $(\bar{\phi}_{n_k})_{k=1}^{\infty}$ of $(\bar{\phi}_n)_{n=1}^{\infty}$ that converges to f a.e. in $\mathcal{X}$. By Lemma 11.76, f is integrable over X and $\int_X f \, d\mu = \lim_{k \in \mathbb{N}} \int_X \bar{\phi}_{n_k} \, d\mu \in \mathcal{Y}$. By the fact that $\|\int_X f_{n_k} \, d\mu - \int_X \bar{\phi}_{n_k} \, d\mu\| < 2^{-n_k}$, $\forall k \in \mathbb{N}$, then we have $\lim_{k \in \mathbb{N}} \int_X f_{n_k} \, d\mu = \lim_{k \in \mathbb{N}} \int_X \bar{\phi}_{n_k} \, d\mu = \int_X f \, d\mu \in \mathcal{Y}$. So far, we have shown that for any uniformly bounded $\mathcal{B}$ -measurable sequence of functions $(f_n)_{n=1}^{\infty}$ that converges to f a.e. in $\mathcal{X}$, there exists a subsequence $(f_{n_k})_{k=1}^{\infty}$ such that $\lim_{k \in \mathbb{N}} \int_X f_{n_k} \, d\mu = \int_X f \, d\mu \in \mathcal{Y}$. This then holds for any subsequence of $(f_n)_{n=1}^{\infty}$. By Proposition 3.71, we have $\lim_{n \in \mathbb{N}} \int_X f_n \, d\mu = \int_X f \, d\mu \in \mathcal{Y}$ and f is integrable over X .

This completes the proof of the theorem. $\square$

Proposition 11.78 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space and $f : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. Then,*

$$\int_X f \, d\mu = \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu =: S \in \mathbb{R}_e$$

where the supremum is over all simple functions $\phi : X \rightarrow [0, \infty) \subset \mathbb{R}$ that satisfy $0 \leq \phi(x) \leq f(x)$, $\forall x \in X$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu(X) < +\infty$; Case 2: $\mu(X) = +\infty$.

Case 1: $\mu(X) < +\infty$. Let $(F_R)_{R \in \mathcal{J}(\mathbb{R})}$ be the net as defined in Definition 11.70 for $\int_X f d\mu$. Clearly, $S \geq 0$. Fix any simple function $\phi : X \rightarrow [0, \infty) \subset \mathbb{R}$ such that $0 \leq \phi(x) \leq f(x)$, $\forall x \in X$. Without loss of generality, we may assume that ϕ is in canonical representation. Then, $\exists n \in \mathbb{Z}_+$, $\exists y_1, \dots, y_n \in [0, \infty) \subset \mathbb{R}$, which are distinct and none equals to 0, and $\exists A_1, \dots, A_n \in \mathcal{B}$, which are nonempty, pairwise disjoint, and $\mu(A_i) < \infty$, $i = 1, \dots, n$, such that $\phi = \sum_{i=1}^n y_i \chi_{A_i, X}$. Let $y_{n+1} := 0$ and $A_{n+1} := X \setminus (\bigcup_{i=1}^n A_i) \in \mathcal{B}$. Then, $X = \bigcup_{i=1}^{n+1} A_i$ and the sets in the union are pairwise disjoint and of finite measure. Without loss of generality, assume that $0 = y_{n+1} < y_n < \dots < y_1 < +\infty$. Let $I_1 := [y_1, \infty) \subset \mathbb{R}$, $I_i := [y_i, y_{i-1}) \subset \mathbb{R}$, $i = 2, \dots, n+1$, $I_{n+2} := (-\infty, 0) \subset \mathbb{R}$, and $y_{n+2} = -1$. Then, $R_0 := \{(y_i, I_i) \mid i = 1, \dots, n+2\} \in \mathcal{J}(\mathbb{R})$. $\forall R \in \mathcal{J}(\mathbb{R})$ with $R_0 \preceq R$, let $R = \{(\bar{y}_\alpha, \bar{U}_\alpha) \mid \alpha \in \Lambda\}$. $\forall \alpha \in \Lambda$, by $R_0 \preceq R$, $\exists! i_\alpha \in \{1, \dots, n+2\}$ such that $\bar{U}_\alpha \subseteq I_{i_\alpha}$. Define $\bar{\Lambda} := \{\alpha \in \Lambda \mid i_\alpha = n+2\}$. Note that $\forall \alpha \in \bar{\Lambda}$, $\bar{U}_\alpha \subseteq I_{n+2}$ and $f_{\text{inv}}(\bar{U}_\alpha) = \emptyset$. Then,

$$\begin{aligned} F_R &= \sum_{\alpha \in \Lambda} \bar{y}_\alpha \mu(f_{\text{inv}}(\bar{U}_\alpha)) = \sum_{\alpha \in \Lambda} \bar{y}_\alpha \sum_{i=1}^{n+1} \mu(A_i \cap f_{\text{inv}}(\bar{U}_\alpha)) \\ &=: \sum_{\alpha \in \Lambda} \bar{y}_\alpha \sum_{i=1}^{n+1} \mu(E_{i,\alpha}) = \sum_{\alpha \in \Lambda \setminus \bar{\Lambda}} \bar{y}_\alpha \sum_{i=1}^{n+1} \mu(E_{i,\alpha}) \end{aligned}$$

$\forall \alpha \in \Lambda \setminus \bar{\Lambda}$, $\forall i \in \{1, \dots, n+1\}$, if $E_{i,\alpha} \neq \emptyset$, then $\exists x \in E_{i,\alpha}$, $y_i = \phi(x) \leq f(x) \in \bar{U}_\alpha \subseteq I_{i_\alpha}$, which further implies that $y_i \leq y_{i_\alpha}$. Note that $\bar{y}_\alpha \in \bar{U}_\alpha \subseteq I_{i_\alpha}$ and $\bar{y}_\alpha \geq y_{i_\alpha}$. Then, $\bar{y}_\alpha \mu(E_{i,\alpha}) \geq y_i \mu(E_{i,\alpha})$. If $E_{i,\alpha} = \emptyset$, clearly $0 = \bar{y}_\alpha \mu(E_{i,\alpha}) = y_i \mu(E_{i,\alpha}) = 0$. Therefore, we have $\bar{y}_\alpha \mu(E_{i,\alpha}) \geq y_i \mu(E_{i,\alpha})$, $\forall \alpha \in \Lambda \setminus \bar{\Lambda}$, $\forall i \in \{1, \dots, n+1\}$. This leads to

$$\begin{aligned} F_R &= \sum_{\alpha \in \Lambda \setminus \bar{\Lambda}} \sum_{i=1}^{n+1} \bar{y}_\alpha \mu(E_{i,\alpha}) \geq \sum_{\alpha \in \Lambda \setminus \bar{\Lambda}} \sum_{i=1}^{n+1} y_i \mu(E_{i,\alpha}) = \sum_{i=1}^{n+1} \sum_{\alpha \in \Lambda \setminus \bar{\Lambda}} y_i \mu(E_{i,\alpha}) \\ &= \sum_{i=1}^{n+1} \sum_{\alpha \in \Lambda} y_i \mu(E_{i,\alpha}) = \sum_{i=1}^{n+1} y_i \mu(A_i) = \int_X \phi d\mu \end{aligned}$$

where the last equality follows from Proposition 11.75. In summary, we have shown that $\forall R \in \mathcal{J}(\mathbb{R})$ with $R_0 \preceq R$, we have $F_R \geq \int_X \phi d\mu$. Then, $\liminf_{R \in \mathcal{J}(\mathbb{R})} F_R \geq \int_X \phi d\mu$. By the arbitrariness of ϕ , we have $\liminf_{R \in \mathcal{J}(\mathbb{R})} F_R \geq S$.

We will show that $\limsup_{R \in \mathcal{J}(\mathbb{R})} F_R \leq S$ by an argument of contradiction. Suppose $\bar{S} := \limsup_{R \in \mathcal{J}(\mathbb{R})} F_R > S$. Then, $\exists \epsilon_0 \in (0, \infty) \subset \mathbb{R}$, $\forall R_0 \in \mathcal{J}(\mathbb{R})$, $\exists R \in \mathcal{J}(\mathbb{R})$ with $R_0 \preceq R$, such that $F_R > S + \epsilon_0$. Let $\epsilon_1 := \min\{1, \frac{\epsilon_0}{2+2\mu(X)}\} > 0$. By Proposition 11.5, $\exists E_1 \in \mathcal{B}$ with $\mu(E_1) < \epsilon_1$

and $\exists M \in \mathbb{N}$ such that $0 \leq f(x) \leq M - 1$, $\forall x \in X \setminus E_1$. Then, $\exists n \in \mathbb{N}$ such that $M/n < \epsilon_1$. Let $U_i := [(i-1)M/n, iM/n) \subset \mathbb{R}$, $y_i := (i-1)M/n$, $i = 1, \dots, n$, $U_{n+1} := [M, \infty) \subset \mathbb{R}$, $y_{n+1} := M$, $U_{n+2} := (-\infty, 0) \subset \mathbb{R}$, and $y_{n+2} = -1$. Then, $R_0 := \{(y_i, U_i) \mid i = 1, \dots, n+2\} \in \mathcal{I}(\mathbb{R})$. Then, $\exists R \in \mathcal{I}(\mathbb{R})$ with $R_0 \preceq R$ such that $F_R > S + \epsilon_0$. Let $R = \{(\bar{y}_\alpha, \bar{U}_\alpha) \mid \alpha \in \Lambda\}$. $\forall \alpha \in \Lambda$, $\exists! i_\alpha \in \{1, \dots, n+2\}$ such that $\bar{U}_\alpha \subseteq U_{i_\alpha}$. Define $\bar{\Lambda} := \{\alpha \in \Lambda \mid i_\alpha = n+2\}$, $\Lambda_1 := \{\alpha \in \Lambda \mid i_\alpha = 1, \dots, n\}$, and $\Lambda_2 := \{\alpha \in \Lambda \mid i_\alpha = n+1\}$. Clearly, $\Lambda = \bar{\Lambda} \cup \Lambda_1 \cup \Lambda_2$ and the sets in the union are pairwise disjoint. $\forall \alpha \in \bar{\Lambda}$, $\bar{U}_\alpha \subseteq U_{n+2} = (-\infty, 0) \subset \mathbb{R}$. Then, $f_{\text{inv}}(\bar{U}_\alpha) = \emptyset$. $\forall \alpha \in \Lambda_1$, note that $\bar{y}_\alpha \in \bar{U}_\alpha \subseteq U_{i_\alpha}$, then $y_{i_\alpha} \leq \bar{y}_\alpha < y_{i_\alpha+1}$. Let $\tilde{y}_\alpha := \inf_{y \in \bar{U}_\alpha} y \in \bar{U}_\alpha \subseteq U_{i_\alpha}$. Then, $y_{i_\alpha} \leq \tilde{y}_\alpha \leq y_{i_\alpha+1}$. Let $A_\alpha := f_{\text{inv}}(\bar{U}_\alpha)$. $\forall \alpha \in \Lambda_2$, note that $\bar{y}_\alpha \in \bar{U}_\alpha \subseteq U_{n+1} = [M, \infty) \subset \mathbb{R}$ and $\bar{y}_\alpha = |\bar{y}_\alpha| < \inf_{y \in \bar{U}_\alpha} |y| + 1 = \inf_{y \in \bar{U}_\alpha} y + 1$. Let $\tilde{y}_\alpha := \inf_{y \in \bar{U}_\alpha} y$. Then, $\tilde{y}_\alpha \leq \bar{y}_\alpha < \tilde{y}_\alpha + 1$. Let $A_\alpha := f_{\text{inv}}(\bar{U}_\alpha)$. Since $\bar{U}_\alpha \subseteq U_{n+1}$, then $f_{\text{inv}}(\bar{U}_\alpha) \subseteq E_1$.

Define a simple function $\phi : X \rightarrow [0, \infty) \subset \mathbb{R}$ by $\phi := \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} \tilde{y}_\alpha \chi_{A_\alpha, X}$.

Clearly, $0 \leq \phi(x) \leq f(x)$, $\forall x \in X$. Then, by Proposition 11.75,

$$\begin{aligned}
 \int_X \phi \, d\mu &= \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} \tilde{y}_\alpha \mu(A_\alpha) = \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} \tilde{y}_\alpha \mu(f_{\text{inv}}(\bar{U}_\alpha)) - F_R + F_R \\
 &> S + \epsilon_0 + \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} \tilde{y}_\alpha \mu(f_{\text{inv}}(\bar{U}_\alpha)) - F_R \\
 &= S + \epsilon_0 + \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} (\tilde{y}_\alpha - \bar{y}_\alpha) \mu(f_{\text{inv}}(\bar{U}_\alpha)) \\
 &\geq S + \epsilon_0 - \sum_{\alpha \in \Lambda_1} |\tilde{y}_\alpha - \bar{y}_\alpha| \mu(f_{\text{inv}}(\bar{U}_\alpha)) - \sum_{\alpha \in \Lambda_2} |\tilde{y}_\alpha - \bar{y}_\alpha| \mu(f_{\text{inv}}(\bar{U}_\alpha)) \\
 &\geq S + \epsilon_0 - M/n \sum_{\alpha \in \Lambda_1} \mu(f_{\text{inv}}(\bar{U}_\alpha)) - \sum_{\alpha \in \Lambda_2} \mu(f_{\text{inv}}(\bar{U}_\alpha)) \\
 &\geq S + \epsilon_0 - \epsilon_1 \mu(X) - \mu(E_1) > S
 \end{aligned}$$

This contradicts with the definition of S . Therefore, $\limsup_{R \in \mathcal{I}(\mathbb{R})} F_R \leq S$.

Then, we have $S \leq \liminf_{R \in \mathcal{I}(\mathbb{R})} F_R \leq \limsup_{R \in \mathcal{I}(\mathbb{R})} F_R \leq S$. By Proposition 3.83, $S = \lim_{R \in \mathcal{I}(\mathbb{R})} F_R = \int_X f \, d\mu$.

Case 2: $\mu(X) = +\infty$. Let $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ be the net as defined in Definition 11.71. $\forall A \in \mathcal{M}(\mathcal{X})$, $A \in \mathcal{B}$ and $\mu(A) < \infty$. Let $(A, \mathcal{B}_A, \mu_A)$ be the measure subspace of the measure space $\mathcal{X}$ as defined in Proposition 11.13. By Case 1, $F_A = \int_A f|_A \, d\mu_A = \sup_{0 \leq \phi \leq f|_A} \int_A \phi \, d\mu_A \in \mathbb{R}_e$, where supremum is over all simple functions $\phi : A \rightarrow [0, \infty) \subset \mathbb{R}$ with $0 \leq \phi(x) \leq f(x)$, $\forall x \in A$. Then, the net $(F_A)_{A \in \mathcal{M}(\mathcal{X})} \subseteq \mathbb{R}_e$ is well-defined.

Fix any simple function $\phi : X \rightarrow [0, \infty) \subset \mathbb{R}$ with $0 \leq \phi(x) \leq f(x)$, $\forall x \in X$. Then, $\exists n \in \mathbb{Z}_+$, $\exists y_1, \dots, y_n \in [0, \infty) \subset \mathbb{R}$, and $\exists A_1, \dots, A_n \in \mathcal{B}$, which are of finite measure, such that $\phi = \sum_{i=1}^n y_i \chi_{A_i, X}$. Take $A_0 =$

$\bigcup_{i=1}^n A_i \in \mathcal{B}$. Then, $\mu(A_0) \leq \sum_{i=1}^n \mu(A_i) < +\infty$. Then, $A_0 \in \mathcal{M}(\mathcal{X})$. $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have, $F_A = \sup_{0 \leq \phi \leq f|_A} \int_A \phi \, d\mu_A \geq \int_A \phi|_A \, d\mu_A = \sum_{i=1}^n y_i \mu(A_i) = \int_X \phi \, d\mu$, where the last two equalities follows from Proposition 11.75. Then, $\liminf_{A \in \mathcal{M}(\mathcal{X})} F_A \geq \int_X \phi \, d\mu$. By the arbitrariness of ϕ , we have $\liminf_{A \in \mathcal{M}(\mathcal{X})} F_A \geq S$.

On the other hand, $\sup_{A \in \mathcal{M}(\mathcal{X})} F_A =: \bar{S} \geq \liminf_{A \in \mathcal{M}(\mathcal{X})} F_A \geq S$. Suppose $\bar{S} > S$. Then, $\exists A \in \mathcal{M}(\mathcal{X})$ such that $S < F_A = \int_A f|_A \, d\mu_A = \sup_{0 \leq \phi \leq f|_A} \int_A \phi \, d\mu_A$. This implies that $\exists$ a simple function $\phi : A \rightarrow [0, \infty) \subset \mathbb{R}$ with $0 \leq \phi(x) \leq f(x)$, $\forall x \in A$, such that $\int_A \phi \, d\mu_A > S$. We may extend this simple function on A to a simple function on X that is zero on $X \setminus A$. Let the extended simple function be $\tilde{\phi} : X \rightarrow [0, \infty) \subset \mathbb{R}$. Then, clearly $0 \leq \tilde{\phi}(x) \leq f(x)$, $\forall x \in X$, and, by Proposition 11.75, $\int_X \tilde{\phi} \, d\mu = \int_A \phi \, d\mu_A > S$. This contradicts with definition of S . Therefore, we have $\bar{S} \leq S$.

In summary, we have $S \leq \liminf_{A \in \mathcal{M}(\mathcal{X})} F_A \leq \limsup_{A \in \mathcal{M}(\mathcal{X})} F_A \leq \sup_{A \in \mathcal{M}(\mathcal{X})} F_A = \bar{S} \leq S$. Then, by Proposition 3.83, we have $\int_X f \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_A = S \in \mathbb{R}_e$.

This completes the proof of the proposition. $\square$

Definition 11.79 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $f : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable. We will define the integral of f over $\mathcal{X}$ by $\int_X f \, d\mu := \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu =: S \in [0, \infty] \subset \mathbb{R}_e$, where the supremum is over all simple functions $\phi : X \rightarrow [0, \infty) \subset \mathbb{R}$ that satisfy $0 \leq \phi(x) \leq f(x)$, $\forall x \in X$. We will say that f is integrable if $S \in \mathbb{R}$.

This definition generalizes the definition of integral to nonnegative extended real-valued functions. By Proposition 11.78, it is consistent with our earlier definitions of integral. The reason that we introduce this generalization is that this leads to considerable simplification in the integrals of nonnegative functions over product measure spaces.

Theorem 11.80 (Fatou's Lemma) Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $f_n : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $f : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$. Then, $\int_X f \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu$.

Proof By Definition 11.79, $\int_X f \, d\mu = \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu$, where the supremum is over all simple function $\phi : X \rightarrow [0, \infty) \subset \mathbb{R}$ that satisfy $0 \leq \phi(x) \leq f(x)$, $\forall x \in X$. Fix any such simple function ϕ . Let $E := \{x \in X \mid \phi(x) > 0\}$. We must have $E \in \mathcal{B}$ and $\mu(E) < +\infty$. Let $M \in [0, \infty) \subset \mathbb{R}$ be such that $0 \leq \phi(x) \leq M$, $\forall x \in X$. Define $g_n := f_n \wedge \phi$, $\forall n \in \mathbb{N}$. By Proposition 11.40, g_n 's are $\mathcal{B}$ -measurable. Since $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$ and $0 \leq \phi(x) \leq f(x)$, $\forall x \in X$, then $\lim_{n \in \mathbb{N}} g_n = \phi$ a.e. in $\mathcal{X}$. By Bounded Convergence Theorem 11.77 and Proposition 11.75, we have $\int_X \phi \, d\mu = \int_E \phi|_E \, d\mu_E = \lim_{n \in \mathbb{N}} \int_E g_n|_E \, d\mu_E$,

where $(E, \mathcal{B}_E, \mu_E)$ is the finite measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. By Definition 11.79, we have $\int_E g_n|_E d\mu_E \leq \int_X g_n d\mu \leq \int_X f_n d\mu$. Then, $\int_X \phi d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n d\mu$. By the arbitrariness of ϕ , we have $\int_X f d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n d\mu$. This completes the proof of the theorem. $\square$

Theorem 11.81 (Monotone Convergence Theorem) *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $f_n : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $f : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$ and $f_n(x) \leq f_{n+1}(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Then, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu$.*

Proof By the assumption, $f_n \leq f$ a.e. in $\mathcal{X}$, $\forall n \in \mathbb{N}$. By Definition 11.79 and the fact that any two simple functions that equal to each other almost everywhere have the same integral, we have $\mathbb{R}_e \ni \int_X f d\mu \geq \int_X f_n d\mu \in \mathbb{R}_e$, $\forall n \in \mathbb{N}$. By Fatou's Lemma 11.80, we have $\int_X f d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n d\mu \leq \limsup_{n \in \mathbb{N}} \int_X f_n d\mu \leq \int_X f d\mu$. Hence, by Proposition 3.83, we have $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu \in \mathbb{R}_e$. This completes the proof of the proposition. $\square$

Proposition 11.82 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space and $g : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable. Assume that $\int_X g d\mu =: \lambda < \infty$. Then, define $f : X \rightarrow [0, \infty) \subset \mathbb{R}$ by $f(x) = \begin{cases} g(x) & g(x) < \infty \\ 0 & g(x) = \infty \end{cases}$, $\forall x \in X$. Then, f is $\mathcal{B}$ -measurable, $f = g$ a.e. in $\mathcal{X}$ and $\int_X f d\mu = \int_X g d\mu \in \mathbb{R}$.*

Proof Let $E := \{x \in X \mid g(x) < \infty\} \in \mathcal{B}$. We will show that $\mu(X \setminus E) = 0$ by an argument of contradiction. Suppose $\mu(X \setminus E) > 0$. Since $\mathcal{X}$ is σ -finite, $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty X_n$ and $\mu(X_n) < +\infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume $X_n \subseteq X_{n+1}$, $\forall n \in \mathbb{N}$. Let $B_n := X_n \setminus E \in \mathcal{B}$, $\forall n \in \mathbb{N}$. By Proposition 11.7, we have $\lim_{n \in \mathbb{N}} \mu(B_n) = \mu(X \setminus E) > 0$. Then, $\exists n \in \mathbb{N}$ such that $\mu(B_n) \in (0, +\infty) \subset \mathbb{R}$. Clearly $\lambda \in [0, \infty) \subset \mathbb{R}$. Let $M := \frac{\lambda+1}{\mu(B_n)} \in (0, \infty) \subset \mathbb{R}$. Let $\tilde{g} := g \wedge M$. By Proposition 11.40, $\tilde{g}$ is $\mathcal{B}$ -measurable. Then, $\tilde{g} : X \rightarrow [0, M] \subset \mathbb{R}$ and $\lambda = \int_X g d\mu \geq \int_X \tilde{g} d\mu \geq M\mu(B_n) = \lambda + 1 > \lambda$, where the first inequality follows from Definition 11.79; and the second inequality follows from Definition 11.79 and Proposition 11.75. This is a contradiction. Hence, $\mu(X \setminus E) = 0$.

Clearly, $f : X \rightarrow [0, \infty) \subset \mathbb{R}$ satisfies $f(x) = \begin{cases} g(x) & x \in E \\ 0 & x \in X \setminus E \end{cases}$. By Proposition 11.41, f is $\mathcal{B}$ -measurable and $f = g$ a.e. in $\mathcal{X}$. By Monotone Convergence Theorem 11.81, $\int_X f d\mu = \int_X g d\mu = \lambda \in \mathbb{R}$. This completes the proof of the proposition. $\square$

The above proposition says that for a measurable extended real-valued function, if its integral is finite then it can be modified to be a real-valued function by replacing its ∞ value by zero on a set of measure zero, and the integral remains the same for modified and the unmodified function.

Proposition 11.83 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $f_i : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable, $i = 1, 2$, $g_n : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $g : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable. Then, the following statements hold.*

- (i) $\forall c \in (0, \infty) \subset \mathbb{R}$, $\int_X (cg) d\mu = c \int_X g d\mu$; $\forall c \in [0, \infty) \subset \mathbb{R}$, if $g : X \rightarrow [0, \infty)$ and $\int_X g d\mu < \infty$, then $\int_X (cg) d\mu = c \int_X g d\mu$.
- (ii) if $f_1 \leq f_2$ a.e. in $\mathcal{X}$, then $\int_X f_1 d\mu \leq \int_X f_2 d\mu$.
- (iii) $\forall E \in \mathcal{B}$, we have $\int_E g|_E d\mu_E = \int_X (g\chi_{E,X}) d\mu =: \int_E g d\mu \leq \int_X g d\mu$, where $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$ is the measure subspace of $\mathcal{X}$.
- (iv) $\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu \in \mathbb{R}_e$.
- (v) $\forall E_1, E_2 \in \mathcal{B}$ with $E_1 \cap E_2 = \emptyset$, we have $\int_{E_1 \cup E_2} g d\mu = \int_{E_1} g d\mu + \int_{E_2} g d\mu \in \mathbb{R}_e$.
- (vi) $\forall$ pairwise disjoint $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$, we have $\int_{\bigcup_{n=1}^\infty E_n} g d\mu = \sum_{n=1}^\infty \int_{E_n} g d\mu \in \mathbb{R}_e$.
- (vii) If g and g_n 's are integrable over $\mathcal{X}$, $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$, and $\lim_{n \in \mathbb{N}} \int_X g_n d\mu = \int_X g d\mu \in \mathbb{R}$. Then, $\forall E \in \mathcal{B}$, we have $\int_E g d\mu = \lim_{n \in \mathbb{N}} \int_E g_n d\mu \in \mathbb{R}$.

Proof (i) and (ii) follows directly from Propositions 11.75 and 3.81 and Definition 11.79.

(iii) Note that $g|_E$ is $\mathcal{B}_E$ -measurable by Proposition 11.41. By Propositions 7.23, 11.38, and 11.39, $g\chi_{E,X}$ is $\mathcal{B}$ -measurable. Then, the result follows directly from Definition 11.79 and Proposition 11.75.

(iv) By Propositions 7.23, 11.38, and 11.39, $f_1 + f_2$ is $\mathcal{B}$ -measurable. By Definition 11.79, we have $\int_X f_1 d\mu = \sup_{0 \leq \phi \leq f_1} \int_X \phi d\mu$, $\int_X f_2 d\mu = \sup_{0 \leq \phi \leq f_2} \int_X \phi d\mu$, and $\int_X (f_1 + f_2) d\mu = \sup_{0 \leq \phi \leq f_1 + f_2} \int_X \phi d\mu$, where the suprema are over all simple functions $\phi : X \rightarrow [0, \infty) \subset \mathbb{R}$ satisfying the stated inequalities. Any simple function $\phi_1 : X \rightarrow [0, \infty) \subset \mathbb{R}$ and $\phi_2 : X \rightarrow [0, \infty) \subset \mathbb{R}$ satisfying $0 \leq \phi_1 \leq f_1$ and $0 \leq \phi_2 \leq f_2$, we have $\phi_1 + \phi_2$ is a simple function satisfying $0 \leq \phi_1 + \phi_2 \leq f_1 + f_2$. Then, $\int_X \phi_1 d\mu + \int_X \phi_2 d\mu = \int_X (\phi_1 + \phi_2) d\mu \leq \int_X (f_1 + f_2) d\mu$, where the equality follows from Proposition 11.75. Hence, we have $\int_X f_1 d\mu + \int_X f_2 d\mu \leq \int_X (f_1 + f_2) d\mu$.

On the other hand, fix any simple function ϕ satisfying $0 \leq \phi \leq f_1 + f_2$. Let $\bar{f}_1 := \phi \wedge f_1$ and $\bar{f}_2 := \phi - \bar{f}_1$. By Propositions 11.40, 11.38, 11.39, and 7.23, $\bar{f}_1$ and $\bar{f}_2$ are $\mathcal{B}$ -measurable. Then, we have $0 \leq \bar{f}_1 \leq f_1$, $0 \leq \bar{f}_2 \leq f_2$, $\bar{f}_1(x) = \bar{f}_2(x) = 0$, $\forall x \in X \setminus E$, where $E := \{x \in X \mid \phi(x) > 0\} \in \mathcal{B}$, and $0 \leq \bar{f}_1(x) \leq \phi(x) \leq M$ and $0 \leq \bar{f}_2(x) \leq \phi(x) \leq M$, $\forall x \in E$, for some $M \in [0, \infty) \subset \mathbb{R}$. Note that $\mu(E) < +\infty$ since ϕ is a simple function. Let $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$ be the finite measure subspace of $\mathcal{X}$ as

defined in Proposition 11.13. By Proposition 11.66, there exists sequences of simple functions $(\phi_{1,n})_{n=1}^\infty$ and $(\phi_{2,n})_{n=1}^\infty$, $\phi_{i,n} : E \rightarrow [0, M] \subset \mathbb{R}$, $\forall i \in \{1, 2\}$, $\forall n \in \mathbb{N}$, that converges to $\bar{f}_1|_E$ and $\bar{f}_2|_E$ a.e. in $\mathcal{E}$, respectively. By Bounded Convergence Theorem 11.77, we have $\int_E \bar{f}_1|_E d\mu_E = \lim_{n \in \mathbb{N}} \int_E \phi_{1,n} d\mu_E$ and $\int_E \bar{f}_2|_E d\mu_E = \lim_{n \in \mathbb{N}} \int_E \phi_{2,n} d\mu_E$. By Propositions 11.52, 7.23, 11.53, $\lim_{n \in \mathbb{N}} (\phi_{1,n} + \phi_{2,n}) = \bar{f}_1|_E + \bar{f}_2|_E = \phi|_E$ a.e. in $\mathcal{E}$. Again, by Bounded Convergence Theorem 11.77 and Proposition 11.75, we have $\int_E \phi|_E d\mu_E = \lim_{n \in \mathbb{N}} \int_E (\phi_{1,n} + \phi_{2,n}) d\mu_E = \lim_{n \in \mathbb{N}} (\int_E \phi_{1,n} d\mu_E + \int_E \phi_{2,n} d\mu_E) = \int_E \bar{f}_1|_E d\mu_E + \int_E \bar{f}_2|_E d\mu_E$. By Proposition 11.75, (ii), and (iii), we have $\int_X \phi d\mu = \int_E \phi|_E d\mu_E = \int_E \bar{f}_1|_E d\mu_E + \int_E \bar{f}_2|_E d\mu_E \leq \int_X \bar{f}_1 d\mu + \int_X \bar{f}_2 d\mu \leq \int_X f_1 d\mu + \int_X f_2 d\mu$. By the arbitrariness of ϕ , we have $\int_X (f_1 + f_2) d\mu \leq \int_X f_1 d\mu + \int_X f_2 d\mu$.

Therefore, we have $\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu$.

(v) By (iii) and (iv), we have $\int_{E_1 \cup E_2} g d\mu = \int_X (g\chi_{E_1 \cup E_2, X}) d\mu = \int_X (g\chi_{E_1, X} + g\chi_{E_2, X}) d\mu = \int_X (g\chi_{E_1, X}) d\mu + \int_X (g\chi_{E_2, X}) d\mu = \int_{E_1} g d\mu + \int_{E_2} g d\mu \in \mathbb{R}_e$.

(vi) Fix any pairwise disjoint $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$. Let $E := \bigcup_{n=1}^\infty E_n \in \mathcal{B}$ and $\bar{E}_n := \bigcup_{i=1}^n E_i \in \mathcal{B}$, $\forall n \in \mathbb{N}$. By Propositions 7.23, 11.38, and 11.39, $g\chi_{\bar{E}_n, X}$ and $g\chi_{E, X}$ are $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. Clearly, $\lim_{n \in \mathbb{N}} g(x)\chi_{\bar{E}_n, X}(x) = g(x)\chi_{E, X}(x)$, $\forall x \in X$, and $0 \leq g(x)\chi_{\bar{E}_n, X}(x) \leq g(x)\chi_{\bar{E}_{n+1}, X}(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Monotone Convergence Theorem 11.81, $\int_E g d\mu = \int_X (g\chi_{E, X}) d\mu = \lim_{n \in \mathbb{N}} \int_X (g\chi_{\bar{E}_n, X}) d\mu = \lim_{n \in \mathbb{N}} \int_{\bar{E}_n} g d\mu = \lim_{n \in \mathbb{N}} \sum_{i=1}^n \int_{E_i} g d\mu = \sum_{n=1}^\infty \int_{E_n} g d\mu$, where the fourth equality follows from (v).

(vii) By (iii), we have $0 \leq \int_E g d\mu \leq \int_X g d\mu < +\infty$, $0 \leq \int_{X \setminus E} g d\mu \leq \int_X g d\mu < +\infty$, $0 \leq \int_E g_n d\mu \leq \int_X g_n d\mu < +\infty$, and $0 \leq \int_{X \setminus E} g_n d\mu \leq \int_X g_n d\mu < +\infty$, $\forall n \in \mathbb{N}$. By Fatou's Lemma 11.80, $\int_E g d\mu = \int_E g|_E d\mu_E \leq \liminf_{n \in \mathbb{N}} \int_E g_n|_E d\mu_E = \liminf_{n \in \mathbb{N}} \int_E g_n d\mu$. By a similar argument, $\int_{X \setminus E} g d\mu \leq \liminf_{n \in \mathbb{N}} \int_{X \setminus E} g_n d\mu$. By (v) and Proposition 3.83, we have $\int_E g d\mu \leq \liminf_{n \in \mathbb{N}} \int_E g_n d\mu \leq \limsup_{n \in \mathbb{N}} \int_E g_n d\mu = \limsup_{n \in \mathbb{N}} \left(\int_X g_n d\mu - \int_{X \setminus E} g_n d\mu \right) = \lim_{n \in \mathbb{N}} \int_X g_n d\mu - \liminf_{n \in \mathbb{N}} \int_{X \setminus E} g_n d\mu \leq \int_X g d\mu - \int_{X \setminus E} g d\mu = \int_E g d\mu$. By Proposition 3.83, we have $\int_E g d\mu = \lim_{n \in \mathbb{N}} \int_E g_n d\mu \in \mathbb{R}$.

This completes the proof of the proposition. $\square$

Proposition 11.84 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space and $f : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable. Assume that f is integrable over $\mathcal{X}$, that is, $\int_X f d\mu \in \mathbb{R}$. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\forall A \in \mathcal{B}$ with $\mu(A) < \delta$, we have $\int_A f|_A d\mu_A < \epsilon$, where $\mathcal{A} := (A, \mathcal{B}_A, \mu_A)$ is the measure subspace of $\mathcal{X}$ as defined in Proposition 11.13.*

Proof $\forall n \in \mathbb{N}$, let $f_n := f \wedge n : X \rightarrow [0, n] \subset \mathbb{R}$. Then, $f_n(x) \leq f_{n+1}(x)$, $\forall x \in X$ and $\lim_{n \in \mathbb{N}} f_n(x) = f(x)$, $\forall x \in X$. By

Proposition 11.40, f_n is $\mathcal{B}$ -measurable. By Propositions 7.23, 11.38, and 11.39, $f - f_n$ is $\mathcal{B}$ -measurable. By Monotone Convergence Theorem 11.81, $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, then, $\exists n_0 \in \mathbb{N}$ such that $|\int_X f \, d\mu - \int_X f_{n_0} \, d\mu| < \epsilon/2$. Choose $\delta := \frac{\epsilon}{2n_0} \in (0, \infty) \subset \mathbb{R}$. $\forall A \in \mathcal{B}$ with $\mu(A) < \delta$, by Proposition 11.83 and Definition 11.79, we have

$$\begin{aligned} \int_A f|_A \, d\mu_A &= \int_A (f - f_{n_0})|_A \, d\mu_A + \int_A f_{n_0}|_A \, d\mu_A \\ &\leq \int_X (f - f_{n_0}) \, d\mu + n_0 \mu(A) = \int_X f \, d\mu - \int_X f_{n_0} \, d\mu + n_0 \mu(A) \\ &< \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

This completes the proof of the proposition. $\square$

11.5 General Convergence Theorems

Lemma 11.85 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $f_n : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable, $f : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$ and $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathbb{R}$. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ and $\exists n_0 \in \mathbb{N}$ such that $\forall A \in \mathcal{B}$ with $\mu(A) < \delta$, $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have $0 \leq \int_A f|_A \, d\mu_A < \epsilon$ and $0 \leq \int_A f_n|_A \, d\mu_A < \epsilon$, where $(A, \mathcal{B}_A, \mu_A)$ is the finite measure subspace of $\mathcal{X}$ as defined in Proposition 11.13.*

Proof Let $\bar{f}_n := f_n \wedge f$, $\forall n \in \mathbb{N}$. By Proposition 11.40, $\bar{f}_n$ is $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. Then, $0 \leq \bar{f}_n(x) \leq f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\int_X \bar{f}_n \, d\mu \leq \int_X f \, d\mu$, $\forall n \in \mathbb{N}$, by Definition 11.79. By Proposition 11.50, $\lim_{n \in \mathbb{N}} \bar{f}_n = f$ a.e. in $\mathcal{X}$. By Fatou's Lemma 11.80, $\int_X f \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X \bar{f}_n \, d\mu \leq \limsup_{n \in \mathbb{N}} \int_X \bar{f}_n \, d\mu \leq \int_X f \, d\mu$, where the last two inequalities follow from Proposition 3.83 and Definition 3.82, respectively. By Proposition 3.83, $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \bar{f}_n \, d\mu \in \mathbb{R}$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Proposition 11.84, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ and $\exists n_0 \in \mathbb{N}$ such that $\forall A \in \mathcal{B}$ with $\mu(A) < \delta$, $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have $0 \leq \int_A f|_A \, d\mu_A < \epsilon/3$, $|\int_X f \, d\mu - \int_X \bar{f}_n \, d\mu| < \epsilon/3$, and $|\int_X f \, d\mu - \int_X f_n \, d\mu| < \epsilon/3$. Then,

$$\begin{aligned} 0 &\leq \int_A f_n|_A \, d\mu_A = \int_A (f_n - \bar{f}_n)|_A \, d\mu_A + \int_A \bar{f}_n|_A \, d\mu_A \\ &\leq \int_A (f_n - \bar{f}_n)|_A \, d\mu_A + \int_A f|_A \, d\mu_A < \int_X (f_n - \bar{f}_n) \, d\mu + \epsilon/3 \\ &= \int_X f_n \, d\mu - \int_X \bar{f}_n \, d\mu + \epsilon/3 \\ &\leq \epsilon/3 + |\int_X f \, d\mu - \int_X \bar{f}_n \, d\mu| + |\int_X f \, d\mu - \int_X f_n \, d\mu| < \epsilon \end{aligned}$$

where the first inequality follows from Definition 11.79; the first equality, the second inequality, and the third inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.83 and the fact that $\int_X f_n d\mu < +\infty$. This completes the proof of the lemma. $\square$

Lemma 11.86 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $\phi_n : X \rightarrow \mathcal{W}$ be a simple function, $\forall n \in \mathbb{N}$, $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $g_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $g : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. Assume that*

(i) $\lim_{n \in \mathbb{N}} \phi_n = f$ a.e. in $\mathcal{X}$ and $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$;

(ii) $\|\phi_n(x)\| \leq g_n(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$;

(iii) g_n 's and g are integrable over X and $\lim_{n \in \mathbb{N}} \int_X g_n d\mu = \int_X g d\mu \in \mathbb{R}$.

Then, f is integrable over X and $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n d\mu \in \mathcal{Y}$.

Proof Let $E := \{x \in X \mid (\phi_n(x))_{n=1}^\infty \text{ does not converge to } f(x) \text{ or } (g_n(x))_{n=1}^\infty \text{ does not converge to } g(x)\}$. Then, by (i), $E \in \mathcal{B}$ and $\mu(E) = 0$. Let $(F_R)_{R \in \mathcal{J}(\mathcal{Y})}$ be the net for $\int_X f d\mu$ as defined in Definition 11.70.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by (i), (iii), and Lemma 11.85, $\exists \delta \in (0, \epsilon/6] \subset \mathbb{R}$ and $\exists n_0 \in \mathbb{N}$ such that $\forall A \in \mathcal{B}$ with $\mu(A) < \delta$, $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have $0 \leq \int_A g|_A d\mu_A < \epsilon/6$ and $0 \leq \int_A g_n|_A d\mu_A < \epsilon/6$, where $(A, \mathcal{B}_A, \mu_A)$ is the finite measure subspace of $\mathcal{X}$. By Egoroff's Theorem 11.55, $\exists \bar{E}_1 \in \mathcal{B}$ with $\mu(\bar{E}_1) < \delta$ such that $(\phi_n|_{X \setminus \bar{E}_1})_{n=1}^\infty$ converges uniformly to $f|_{X \setminus \bar{E}_1}$. Let $E_1 := E \cup \bar{E}_1$. Then, $E_1 \in \mathcal{B}$ with $\mu(E_1) = \mu(\bar{E}_1) < \delta$ such that $(\phi_n|_{X \setminus E_1})_{n=1}^\infty$ converges uniformly to $f|_{X \setminus E_1}$. Then, $\exists n_1 \in \mathbb{N}$ with $n_1 \geq n_0$, $\forall n \in \mathbb{N}$ with $n \geq n_1$, $\forall x \in X \setminus E_1$, $\|\phi_n(x) - f(x)\| < \frac{\epsilon}{6\mu(X)+1} =: \bar{\epsilon}$. Fix any $n \geq n_1$. Let ϕ_n admit the canonical representation $\phi_n = \sum_{i=1}^{\bar{n}} w_i \chi_{A_i, X}$, where $\bar{n} \in \mathbb{Z}_+$, $w_1, \dots, w_{\bar{n}} \in \mathcal{W}$ are distinct and none equals to $\vartheta_{\mathcal{Y}}$, and $A_1, \dots, A_{\bar{n}} \in \mathcal{B}$ are nonempty, pairwise disjoint, and of finite measure. Let $w_{\bar{n}+1} := \vartheta_{\mathcal{Y}}$ and $A_{\bar{n}+1} := X \setminus (\bigcup_{i=1}^{\bar{n}} A_i) \in \mathcal{B}$. Then, $X = \bigcup_{i=1}^{\bar{n}+1} A_i$ and the sets in the union are pairwise disjoint and of finite measure. Define $\bar{V}_i := \mathcal{B}_{\mathcal{Y}}(w_i, \bar{\epsilon})$, $i = 1, \dots, \bar{n}+1$. Define $V_1 := \bar{V}_1 \in \mathcal{B}_{\mathcal{B}}(\mathcal{Y})$, $V_{i+1} := \bar{V}_{i+1} \setminus (\bigcup_{j=1}^i \bar{V}_j) \in \mathcal{B}_{\mathcal{B}}(\mathcal{Y})$, $i = 1, \dots, \bar{n}$. Let $V_{\bar{n}+2} := \mathcal{Y} \setminus (\bigcup_{j=1}^{\bar{n}+1} \bar{V}_j) \in \mathcal{B}_{\mathcal{B}}(\mathcal{Y})$. Clearly, we may form a representation $\bar{R} := \{(\bar{y}_i, V_i) \mid i = 1, \dots, \bar{n}+2, V_i \neq \emptyset\} \in \mathcal{J}(\mathcal{Y})$, where $\bar{y}_i \in V_i$, $i = 1, \dots, \bar{n}+2$, are any vectors that satisfy the assumption of Proposition 11.69. $\forall R \in \mathcal{J}(\mathcal{Y})$ with $\bar{R} \preceq R$. Let $R := \{(\hat{y}_\gamma, \hat{U}_\gamma) \mid \gamma \in \Gamma\}$. $\forall \gamma \in \Gamma$, by $\bar{R} \preceq R$, $\exists! i_\gamma \in \{1, \dots, \bar{n}+2\}$ such that $\hat{U}_\gamma \subseteq V_{i_\gamma}$. Define $\bar{\Gamma} := \{\gamma \in \Gamma \mid i_\gamma = \bar{n}+2\}$. Let $A_{\gamma, i} := A_i \cap f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1) \in \mathcal{B}$, $i = 1, \dots, \bar{n}+1$. Let $\bar{A}_\gamma := f_{\text{inv}}(\hat{U}_\gamma) \cap (E_1 \setminus E) \in \mathcal{B}$.

Claim 11.86.1 $\forall \gamma \in \bar{\Gamma}, f_{\text{inv}}(\hat{U}_\gamma) \subseteq E_1; \forall \gamma \in \Gamma_s := \{\gamma \in \Gamma \mid \bar{A}_\gamma \neq \emptyset\},$
 $\|\hat{y}_\gamma \mu(\bar{A}_\gamma)\| \leq (1 + \inf_{x \in \bar{A}_\gamma} g(x))\mu(\bar{A}_\gamma);$ and $\left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| < \epsilon/3.$

Proof of claim: Fix any $\gamma \in \bar{\Gamma}$. $\hat{U}_\gamma \subseteq V_{\bar{n}+2} = \mathcal{Y} \setminus (\bigcup_{j=1}^{\bar{n}+1} \bar{V}_j)$. Then, $\mathcal{B} \ni f_{\text{inv}}(\hat{U}_\gamma) \subseteq E_1$.

$\forall \gamma \in \Gamma_s, \bar{A}_\gamma = f_{\text{inv}}(\hat{U}_\gamma) \cap (E_1 \setminus E) \neq \emptyset$. $\forall x \in \bar{A}_\gamma$, we have $f(x) \in \hat{U}_\gamma$, $\lim_{m \in \mathbb{N}} \phi_m(x) = f(x)$, $\|\phi_m(x)\| \leq g_m(x)$, $\forall m \in \mathbb{N}$, and $\lim_{m \in \mathbb{N}} g_m(x) = g(x)$. Then, by Propositions 4.30, 3.66, and 3.67, we have $\|f(x)\| = \lim_{m \in \mathbb{N}} \|\phi_m(x)\| \leq \lim_{m \in \mathbb{N}} g_m(x) = g(x)$. Then, by $R \in \mathcal{I}(\mathcal{Y})$, $\|\hat{y}_\gamma\| < \inf_{y \in \hat{U}_\gamma} \|y\| + 1 \leq \|f(x)\| + 1 \leq g(x) + 1$. By the arbitrariness of x , we have $\|\hat{y}_\gamma\| \leq \inf_{x \in \bar{A}_\gamma} g(x) + 1$. Therefore, $\|\hat{y}_\gamma \mu(\bar{A}_\gamma)\| \leq (1 + \inf_{x \in \bar{A}_\gamma} g(x))\mu(\bar{A}_\gamma)$.

Note that

$$\begin{aligned} \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| &= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma (\mu(\bar{A}_\gamma) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1 \cap E)) \right\| \\ &= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(\bar{A}_\gamma) \right\| = \left\| \sum_{\gamma \in \Gamma_s} \hat{y}_\gamma \mu(\bar{A}_\gamma) \right\| \leq \sum_{\gamma \in \Gamma_s} \|\hat{y}_\gamma \mu(\bar{A}_\gamma)\| \\ &\leq \sum_{\gamma \in \Gamma_s} (1 + \inf_{x \in \bar{A}_\gamma} g(x))\mu(\bar{A}_\gamma) = \sum_{\gamma \in \Gamma_s} \mu(\bar{A}_\gamma) + \sum_{\gamma \in \Gamma_s} (\inf_{x \in \bar{A}_\gamma} g(x))\mu(\bar{A}_\gamma) \end{aligned}$$

The simple function $\psi : X \rightarrow [0, \infty) \subset \mathbb{R}$, defined by $\psi(x) = \sum_{\gamma \in \Gamma_s} (\inf_{\bar{x} \in \bar{A}_\gamma} g(\bar{x}))\chi_{\bar{A}_\gamma, X}(x)$, $\forall x \in X$, satisfies $0 \leq \psi(x) \leq g(x)$, $\forall x \in X$ and $\psi(x) = 0$, $\forall x \in X \setminus E_1$. Then, $\sum_{\gamma \in \Gamma_s} (\inf_{x \in \bar{A}_\gamma} g(x))\mu(\bar{A}_\gamma) = \int_X \psi d\mu = \int_{E_1} \psi|_{E_1} d\mu_{E_1} \leq \int_{E_1} g|_{E_1} d\mu_{E_1} < \epsilon/6$, where the equalities follows from Proposition 11.75; and the first inequality follows from Proposition 11.83. Then, $\left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| < \mu(E_1) + \epsilon/6 < \delta + \epsilon/6 \leq \epsilon/3$.

This completes the proof of the claim. $\square$

Claim 11.86.2 $\left\| \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right\| < \epsilon/6$.

Proof of claim: Since $\|\phi_n(x)\| \leq g_n(x)$, $\forall x \in X$, then $\|w_i\| \leq g_n(x)$, $\forall x \in A_i$, $i = 1, \dots, \bar{n} + 1$. Then, $\left\| \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right\| \leq \sum_{i=1}^{\bar{n}+1} \|w_i\| \mu(A_i \cap E_1) \leq \int_{E_1} g_n|_{E_1} d\mu_{E_1} < \epsilon/6$, where the second inequality follows from Definition 11.79. $\square$

Claim 11.86.3 $\forall \gamma \in \Gamma \setminus \bar{\Gamma}, \forall i \in \{1, \dots, \bar{n} + 1\}$, we have $\|\hat{y}_\gamma - w_i\| \mu(A_{\gamma, i}) \leq 3\epsilon \mu(A_{\gamma, i})$.

Proof of claim: $\forall \gamma \in \Gamma \setminus \bar{\Gamma}, \forall i \in \{1, \dots, \bar{n} + 1\}, i_\gamma \in \{1, \dots, \bar{n} + 1\}$ and $\hat{U}_\gamma \subseteq V_{i_\gamma}$. The result is trivial if $A_{\gamma, i} = \emptyset$. If $A_{\gamma, i} \neq \emptyset$, then $\exists x \in A_{\gamma, i} = A_i \cap$

$f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1)$. This implies that $\phi_n(x) = w_i$ and $f(x) \in \mathcal{B}_y(w_i, \bar{\epsilon}) \cap \hat{U}_\gamma \subseteq \mathcal{B}_y(w_i, \bar{\epsilon}) \cap V_{i_\gamma} \subseteq \mathcal{B}_y(w_i, \bar{\epsilon}) \cap \bar{V}_{i_\gamma} = \mathcal{B}_y(w_i, \bar{\epsilon}) \cap \mathcal{B}_y(w_{i_\gamma}, \bar{\epsilon})$. Note that $\hat{y}_\gamma \in \hat{U}_\gamma \subseteq \mathcal{B}_y(w_{i_\gamma}, \bar{\epsilon})$. Therefore, we have $\|\hat{y}_\gamma - w_i\| \leq \|\hat{y}_\gamma - w_{i_\gamma}\| + \|w_{i_\gamma} - f(x)\| + \|f(x) - w_i\| < 3\bar{\epsilon}$. Hence, the result holds. $\square$

Note that $F_R = \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma))$ and $\int_X \phi_n d\mu = \sum_{i=1}^{\bar{n}} w_i \mu(A_i)$, by Proposition 11.75. Then,

$$\begin{aligned}
\|F_R - \int_X \phi_n d\mu\| &= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma)) - \sum_{i=1}^{\bar{n}} w_i \mu(A_i) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma (\mu(f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1)) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1)) \right. \\
&\quad \left. - \sum_{i=1}^{\bar{n}+1} w_i (\mu(A_i \cap E_1) + \mu(A_i \cap (X \setminus E_1))) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \left(\sum_{i=1}^{\bar{n}+1} \mu(A_{\gamma,i}) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right) - \sum_{i=1}^{\bar{n}+1} w_i (\mu(A_i \cap E_1) \right. \\
&\quad \left. + \mu(A_{\gamma,i})) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} \hat{y}_\gamma \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) - \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right. \\
&\quad \left. - \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} w_i \mu(A_{\gamma,i}) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - w_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right. \\
&\quad \left. - \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right\| \\
&< \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - w_i) \mu(A_{\gamma,i}) \right\| + \frac{\epsilon}{3} + \frac{\epsilon}{6} \\
&= \frac{\epsilon}{2} + \left\| \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - w_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - w_i) \mu(A_{\gamma,i}) \right\| \\
&= \frac{\epsilon}{2} + \left\| \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - w_i) \mu(A_{\gamma,i}) \right\| \leq \frac{\epsilon}{2} + \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} \|\hat{y}_\gamma - w_i\| \mu(A_{\gamma,i}) \\
&\leq \frac{\epsilon}{2} + \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} 3\bar{\epsilon} \mu(A_{\gamma,i}) \leq \frac{\epsilon}{2} + \frac{3\mu(X)\epsilon}{6\mu(X) + 1} < \epsilon
\end{aligned}$$

where the first inequality follows from Claims 11.86.1 and 11.86.2; and the seventh equality follows from the fact that $A_{\gamma,i} = \emptyset$, $\forall \gamma \in \bar{\Gamma}$, $\forall i \in \{1, \dots, \bar{n} + 1\}$; and the third inequality follows from Claim 11.86.3.

In summary, we have shown that $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_1 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq n_1$, $\exists \bar{R} \in \mathcal{I}(\mathcal{Y})$ such that $\forall R \in \mathcal{I}(\mathcal{Y})$ with $\bar{R} \preceq R$, we have $\|F_R - \int_X \phi_n d\mu\| < \epsilon$. Then, the net $(F_R)_{R \in \mathcal{I}(\mathcal{Y})}$ is a Cauchy net. By Proposition 4.44, the net admits a limit $\int_X f d\mu \in \mathcal{Y}$. By Propositions 4.30, 3.66, and 3.67, we have $\|\int_X f d\mu - \int_X \phi_n d\mu\| = \lim_{R \in \mathcal{I}(\mathcal{Y})} \|F_R - \int_X \phi_n d\mu\| \leq \epsilon$, $\forall n \geq n_1$. Hence, we have $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n d\mu \in \mathcal{Y}$ and f is integrable over X . This completes the proof of the lemma. $\square$

Lemma 11.87 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $U \subseteq \mathcal{W}$ be a conic segment, $f : X \rightarrow U \subseteq \mathcal{W}$ be $\mathcal{B}$ -measurable, and $g : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. Assume that $\|f(x)\| \leq g(x)$, $\forall x \in X$, and g is integrable over X . Then, there exists a sequence of simple functions $(\varphi_k)_{k=1}^\infty$, $\varphi_k : X \rightarrow U$, $\forall k \in \mathbb{N}$, such that $\lim_{k \in \mathbb{N}} \varphi_k = f$ a.e. in $\mathcal{X}$, $\|\varphi_k(x)\| \leq \|f(x)\|$, $\forall x \in X$, $\forall k \in \mathbb{N}$, and f is integrable over X with $\int_X f d\mu = \lim_{k \in \mathbb{N}} \int_X \varphi_k d\mu \in \mathcal{Y}$.*

Furthermore, if $\mathcal{Y}$ admits a positive cone P and $U \subseteq P \subseteq \mathcal{Y}$, then $\int_X f d\mu \in P$.

Proof By Proposition 11.66, there exists a sequence of simple functions $(\varphi_k)_{k=1}^\infty$, $\varphi_k : X \rightarrow U$, $\forall k \in \mathbb{N}$, such that $\|\varphi_k(x)\| \leq \|f(x)\| \leq g(x)$, $\forall x \in X$, $\forall k \in \mathbb{N}$, and $\lim_{k \in \mathbb{N}} \varphi_k = f$ a.e. in $\mathcal{X}$. By Lemma 11.86, f is integrable over X and $\int_X f d\mu = \lim_{k \in \mathbb{N}} \int_X \varphi_k d\mu \in \mathcal{Y}$.

Furthermore, if $\mathcal{Y}$ admits a positive cone P and $U \subseteq P \subseteq \mathcal{Y}$, then, by Propositions 8.41 and 11.75, $\varphi_k : X \rightarrow U \subseteq P \subseteq \mathcal{Y}$, $\forall k \in \mathbb{N}$, and $\int_X \varphi_k d\mu \in P$. Then, by Proposition 4.13, $\int_X f d\mu \in \bar{P} = P$, since P is closed. This completes the proof of the lemma. $\square$

Theorem 11.88 (Lebesgue Dominated Convergence) *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $f_n : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $g_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $g : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. Assume that*

(i) $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$, $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$;

(ii) $\|f_n(x)\| \leq g_n(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$;

(iii) g_n 's and g are integrable over X and $\int_X g d\mu = \lim_{n \in \mathbb{N}} \int_X g_n d\mu \in \mathbb{R}$.

Then, f is integrable over X and $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu \in \mathcal{Y}$.

Proof $\forall n \in \mathbb{N}$, by Lemma 11.87, there exists a sequence of simple functions $(\bar{\varphi}_{n,k})_{k=1}^\infty$, $\bar{\varphi}_{n,k} : X \rightarrow \mathcal{W}$, $\forall k \in \mathbb{N}$, such that $\lim_{k \in \mathbb{N}} \bar{\varphi}_{n,k} =$

f_n a.e. in $\mathcal{X}$, $\|\bar{\varphi}_{n,k}(x)\| \leq \|f_n(x)\|$, $\forall x \in X$, $\forall k \in \mathbb{N}$, f_n is integrable over X , and $\int_X f_n d\mu = \lim_{k \in \mathbb{N}} \int_X \bar{\varphi}_{n,k} d\mu \in \mathcal{Y}$. By Proposition 11.58, $\lim_{k \in \mathbb{N}} \bar{\varphi}_{n,k} = f_n$ in measure in $\mathcal{X}$. Then, $\exists k_n \in \mathbb{N}$ such that $\mu(E_n) := \mu(\{x \in X \mid \|\bar{\varphi}_{n,k_n}(x) - f_n(x)\| \geq 2^{-n}\}) < 2^{-n}$ and $\|\int_X f_n d\mu - \int_X \bar{\varphi}_{n,k_n} d\mu\| < 2^{-n}$. Denote $\psi_n := \bar{\varphi}_{n,k_n}$.

By Proposition 11.58, $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have $2^{-n} < \epsilon/2$ and $\mu(\bar{E}_n) := \mu(\{x \in X \mid \|f_n(x) - f(x)\| \geq \epsilon/2\}) < \epsilon/2$. Note that, by Lemma 11.43, $\mathcal{B} \ni F_n := \{x \in X \mid \|\psi_n(x) - f(x)\| \geq \epsilon\} \subseteq \{x \in X \mid \|\psi_n(x) - f_n(x)\| + \|f_n(x) - f(x)\| \geq \epsilon\} \subseteq E_n \cup \bar{E}_n$. Then, we have $\mu(F_n) \leq \mu(E_n) + \mu(\bar{E}_n) < \epsilon$. This implies that $\lim_{n \in \mathbb{N}} \psi_n = f$ in measure in $\mathcal{X}$. By Proposition 11.57, there exists a subsequence $(\psi_{n_i})_{i=1}^\infty$ of $(\psi_n)_{n=1}^\infty$ that converges to f a.e. in $\mathcal{X}$. By Lemma 11.86, f is integrable over X and $\int_X f d\mu = \lim_{i \in \mathbb{N}} \int_X \psi_{n_i} d\mu \in \mathcal{Y}$. By the fact that $\|\int_X f_{n_i} d\mu - \int_X \psi_{n_i} d\mu\| < 2^{-n_i}$, $\forall i \in \mathbb{N}$, then we have $\lim_{i \in \mathbb{N}} \int_X f_{n_i} d\mu = \lim_{i \in \mathbb{N}} \int_X \psi_{n_i} d\mu = \int_X f d\mu \in \mathcal{Y}$. So far, we have shown that, for the sequence $(f_n)_{n=1}^\infty$, there exists a subsequence $(f_{n_i})_{i=1}^\infty$ such that $\lim_{i \in \mathbb{N}} \int_X f_{n_i} d\mu = \int_X f d\mu \in \mathcal{Y}$. This then holds for any subsequence of $(f_n)_{n=1}^\infty$. By Proposition 3.71, we have $\lim_{n \in \mathbb{N}} \int_X f_n d\mu = \int_X f d\mu$.

This completes the proof of the theorem. $\square$

Proposition 11.89 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $f_i : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $g_i : X \rightarrow [0, \infty) \subset \mathbb{R}$ be integrable over $\mathcal{X}$, $i = 1, 2$. Assume that $\|f_i(x)\| \leq g_i(x)$, $\forall x \in X$, $i = 1, 2$. Then, the following statements hold.*

- (i) f_1 , f_2 , and $f_1 + f_2$ are integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu \in \mathcal{Y}$.
- (ii) $\forall A \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, Af_1 is integrable over $\mathcal{X}$ and $\int_X (Af_1) d\mu = A \int_X f_1 d\mu \in \mathcal{Z}$.
- (iii) $\forall c \in \mathbb{K}$, cf_1 is integrable over $\mathcal{X}$ and $\int_X (cf_1) d\mu = c \int_X f_1 d\mu \in \mathcal{Y}$.
- (iv) $\forall E \in \mathcal{B}$, $f_1|_E$ is integrable over $\mathcal{E}$ and $\int_X (f_1 \chi_{E,X}) d\mu = \int_E f_1|_E d\mu_E \in \mathcal{Y}$, where $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$ is the finite measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. We will henceforth denote $\int_E f_1|_E d\mu_E$ by $\int_E f_1 d\mu$.
- (v) If $f_1 = f_2$ a.e. in $\mathcal{X}$ then $\int_X f_1 d\mu = \int_X f_2 d\mu \in \mathcal{Y}$.
- (vi) $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, $\sum_{i=1}^\infty \int_{E_i} f_1 d\mu = \int_{\bigcup_{i=1}^\infty E_i} f_1 d\mu \in \mathcal{Y}$.
- (vii) $0 \leq \|\int_X f_1 d\mu\| \leq \int_X \mathcal{P} \circ f_1 d\mu \leq \int_X g_1 d\mu < +\infty$, where $\mathcal{P} \circ f_1 : X \rightarrow [0, \infty) \subset \mathbb{R}$ is as defined on page 380.

(viii) If $\mathcal{Y}$ admits a positive cone P and $f_1 \leq f_2$ a.e. in $\mathcal{X}$, then $\int_X f_1 d\mu \leq \int_X f_2 d\mu$.

Proof By Lemma 11.87, there exist sequences of simple functions $(\phi_{i,n})_{n=1}^\infty, \phi_{i,n} : X \rightarrow \mathcal{W}, \forall n \in \mathbb{N}, i = 1, 2$ such that $i = 1, 2, \lim_{n \in \mathbb{N}} \phi_{i,n} = f_i$ a.e. in $\mathcal{X}$, $\|\phi_{i,n}(x)\| \leq \|f_i(x)\|, \forall x \in X, \forall n \in \mathbb{N}$, f_i is integrable over $\mathcal{X}$, and $\int_X f_i d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_{i,n} d\mu \in \mathcal{Y}$.

(i) By Propositions 7.23, 11.38, and 11.39, $f_1 + f_2, g_1 + g_2$, and $\phi_{1,n} + \phi_{2,n}$ are $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. By Proposition 11.83, $0 \leq \int_X (g_1 + g_2) d\mu = \int_X g_1 d\mu + \int_X g_2 d\mu < +\infty$. Note that, by Propositions 7.23, 11.52, and 11.53, $\lim_{n \in \mathbb{N}} (\phi_{1,n} + \phi_{2,n}) = f_1 + f_2$ a.e. in $\mathcal{X}$ and $\|\phi_{1,n}(x) + \phi_{2,n}(x)\| \leq \|\phi_{1,n}(x)\| + \|\phi_{2,n}(x)\| \leq g_1(x) + g_2(x), \forall x \in X, \forall n \in \mathbb{N}$. Clearly, $\phi_{1,n} + \phi_{2,n} : X \rightarrow \mathcal{W}$ and $f_1 + f_2 : X \rightarrow \mathcal{W}, \forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.88, $f_1 + f_2$ is integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) d\mu = \lim_{n \in \mathbb{N}} \int_X (\phi_{1,n} + \phi_{2,n}) d\mu = \lim_{n \in \mathbb{N}} (\int_X \phi_{1,n} d\mu + \int_X \phi_{2,n} d\mu) = \int_X f_1 d\mu + \int_X f_2 d\mu \in \mathcal{Y}$, where the second equality follows from Proposition 11.75; and the third equality follows from Propositions 3.66, 3.67, and 7.23.

(ii) $\forall A \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, by Proposition 11.83 and the fact that $\int_X g_1 d\mu \in \mathbb{R}$, we have $\int_X (\|A\| g_1) d\mu = \|A\| \int_X g_1 d\mu \in \mathbb{R}$. By Propositions 7.62 and 11.38, Af_1 and $A\phi_{1,n}$ are $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. Clearly, $A\phi_{1,n} : X \rightarrow \hat{\mathcal{W}}$ and $Af_1 : X \rightarrow \hat{\mathcal{W}}, \forall n \in \mathbb{N}$, where $\hat{\mathcal{W}} := \{z \in \mathcal{Z} \mid z = Aw, w \in \mathcal{W}\}$ is a separable subspace of $\mathcal{Z}$. By Propositions 11.52 and 7.62, $\lim_{n \in \mathbb{N}} A\phi_{1,n} = Af_1$ a.e. in $\mathcal{X}$. By Proposition 11.75, we have $\int_X (A\phi_{1,n}) d\mu = A \int_X \phi_{1,n} d\mu \in \mathcal{Z}$. Note that $\|A\phi_{1,n}(x)\| \leq \|A\| \cdot \|\phi_{1,n}(x)\| \leq \|A\| g_1(x), \forall x \in X, \forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.88, Af_1 is integrable over $\mathcal{X}$ and $\int_X (Af_1) d\mu = \lim_{n \in \mathbb{N}} \int_X (A\phi_{1,n}) d\mu = \lim_{n \in \mathbb{N}} A \int_X \phi_{1,n} d\mu = A \int_X f_1 d\mu \in \mathcal{Z}$, where the last equality follows from Propositions 3.66 and 7.62.

(iii) This follows immediately from (ii).

(iv) $(\phi_{1,n}|_E)_{n=1}^\infty$ is a sequence of simple functions that converges to $f_1|_E$ a.e. in $\mathcal{E}$. Note that $\|\phi_{1,n}|_E(x)\| \leq g_1|_E(x), \forall x \in E, \forall n \in \mathbb{N}$. By Proposition 11.83, $0 \leq \int_E g_1|_E d\mu_E \leq \int_X g_1 d\mu < +\infty$. By Lebesgue Dominated Convergence Theorem 11.88, we have $\int_E f_1|_E d\mu_E = \lim_{n \in \mathbb{N}} \int_E \phi_{1,n}|_E d\mu_E \in \mathcal{Y}$. Note that $(\phi_{1,n}\chi_{E,X})_{n=1}^\infty$ is a sequence of simple functions that converges to $f_1\chi_{E,X}$ a.e. in $\mathcal{X}$, and $\|\phi_{1,n}(x)\chi_{E,X}(x)\| \leq g_1(x), \forall x \in X, \forall n \in \mathbb{N}$. By Propositions 7.23, 11.38, and 11.39, $\phi_{1,n}\chi_{E,X}$ and $f_1\chi_{E,X}$ are $\mathcal{B}$ -measurable. Again by Lebesgue Dominated Convergence Theorem 11.88, we have $\int_X (f_1\chi_{E,X}) d\mu = \lim_{n \in \mathbb{N}} \int_X (\phi_{1,n}\chi_{E,X}) d\mu = \lim_{n \in \mathbb{N}} \int_E \phi_{1,n}|_E d\mu_E = \int_E f_1|_E d\mu_E \in \mathcal{Y}$, where the second equality follows from Proposition 11.75.

(v) If $f_1 = f_2$ a.e. in $\mathcal{X}$, then $\lim_{n \in \mathbb{N}} \phi_{1,n} = f_2$ a.e. in $\mathcal{X}$ by Proposition 11.54. By Lebesgue Dominated Convergence Theorem 11.88, we have $\int_X f_2 d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_{1,n} d\mu = \int_X f_1 d\mu$.

(vi) $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$ and $\bar{E}_n := \bigcup_{i=1}^n E_i \in \mathcal{B}$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} f_1(x) \chi_{\bar{E}_n, X}(x) = f_1(x) \chi_{E, X}(x)$, $\forall x \in X$. Then, we have $\sum_{i=1}^\infty \int_{E_i} f_1 d\mu = \lim_{n \in \mathbb{N}} \sum_{i=1}^n \int_X (f_1 \chi_{E_i, X}) d\mu = \lim_{n \in \mathbb{N}} \int_X (f_1 \chi_{\bar{E}_n, X}) d\mu = \int_X (f_1 \chi_{E, X}) d\mu = \int_E f_1 d\mu \in \mathcal{Y}$, where the first equality follows from (iv); the second equality follows from (i); the third equality follows from Lebesgue Dominated Convergence Theorem 11.88; the last equality follows from (iv); and the last step follows from (iv).

(vii) Note that, by Proposition 11.75, $\|\int_X \phi_{1,n} d\mu\| \leq \int_X \mathcal{P} \circ \phi_{1,n} d\mu \in \mathbb{R}$, $\forall n \in \mathbb{N}$. By Propositions 7.21 and 11.38, $\mathcal{P} \circ \phi_{1,n}$ and $\mathcal{P} \circ f_1$ are $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. By Propositions 7.21 and 11.52, we have $\lim_{n \in \mathbb{N}} \mathcal{P} \circ \phi_{1,n} = \mathcal{P} \circ f_1$ a.e. in $\mathcal{X}$. Note that $0 \leq \mathcal{P} \circ \phi_{1,n}(x) \leq g_1(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Then, $0 \leq \|\int_X f_1 d\mu\| = \|\lim_{n \in \mathbb{N}} \int_X \phi_{1,n} d\mu\| = \lim_{n \in \mathbb{N}} \|\int_X \phi_{1,n} d\mu\| \leq \lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \phi_{1,n} d\mu = \int_X \mathcal{P} \circ f_1 d\mu \leq \int_X g_1 d\mu < +\infty$, where the second equality follows from Propositions 7.21 and 3.66; the third equality follows from Lebesgue Dominated Convergence Theorem 11.88; and the third inequality follows from Proposition 11.83.

(viii) Let $f := f_2 - f_1$. Then, $f \geq \vartheta_{\mathcal{Y}}$ a.e. in $\mathcal{X}$, $f : X \rightarrow \mathcal{W}$, and f is $\mathcal{B}$ -measurable by Propositions 11.38, 11.39, and 7.23. Note that P is a closed set, $P \in \mathcal{B}_{\mathcal{B}}(\mathcal{Y})$, and $f_{\text{inv}}(P) \in \mathcal{B}$. Define $\bar{f} : X \rightarrow \mathcal{W} \cap P \subseteq \mathcal{Y}$ by $\bar{f}(x) = \begin{cases} f(x) & x \in f_{\text{inv}}(P) \\ \vartheta_{\mathcal{Y}} & x \in X \setminus f_{\text{inv}}(P) \end{cases}$. Then, $f = \bar{f}$ a.e. in $\mathcal{X}$, $\bar{f} : X \rightarrow \mathcal{W} \cap P$, and $\|\bar{f}(x)\| \leq \|f(x)\| \leq \|f_1(x)\| + \|f_2(x)\| \leq g_1(x) + g_2(x)$, $\forall x \in X$. By Proposition 11.41, $\bar{f}$ is $\mathcal{B}$ -measurable. By Proposition 11.83, $g_1 + g_2$ is integrable over $\mathcal{X}$. By Lemma 11.87, $\int_X \bar{f} d\mu \in P$. By (i) and (v), we have $\int_X f_2 d\mu - \int_X f_1 d\mu = \int_X f d\mu = \int_X \bar{f} d\mu \geq \vartheta_{\mathcal{Y}}$. Then, $\int_X f_1 d\mu \leq \int_X f_2 d\mu$.

This completes the proof of the proposition. $\square$

Definition 11.90 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a normed linear space, and $f : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. f is said to be absolutely integrable if $\mathcal{P} \circ f$ is integrable over $\mathcal{X}$, that is, $0 \leq \int_X \mathcal{P} \circ f d\mu < +\infty$.

Note that $\mathcal{P} \circ f$ is $\mathcal{B}$ -measurable by Propositions 7.21 and 11.38. We will show that the desired properties of integration are valid if the integrand is absolutely integrable.

Theorem 11.91 (Lebesgue Dominated Convergence) Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $f_n : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $g_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $g : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. Assume that

(i) $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$, $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$;

(ii) $\|f_n(x)\| \leq g_n(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$;

(iii) g_n 's and g are integrable over $\mathcal{X}$ and $\int_X g d\mu = \lim_{n \in \mathbb{N}} \int_X g_n d\mu \in \mathbb{R}$.

Then, f is integrable over $\mathcal{X}$ and $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathcal{Y}$. Furthermore, f is absolutely integrable and $0 \leq \int_X (\mathcal{P} \circ f) \, d\mu = \lim_{n \in \mathbb{N}} \int_X (\mathcal{P} \circ f_n) \, d\mu \leq \int_X g \, d\mu < \infty$.

Proof We will show that f is integrable over $\mathcal{X}$ and $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathcal{Y}$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\mu(X) < +\infty$; Case 2: $\mu(X) = +\infty$.

Case 1: $\mu(X) < +\infty$. The result follows immediately from Lebesgue Dominated Convergence Theorem 11.88.

Case 2: $\mu(X) = +\infty$. Let $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ and $(F_{n,A})_{A \in \mathcal{M}(\mathcal{X})}$ be the nets for the functions f and f_n as defined in Definition 11.71, respectively, $\forall n \in \mathbb{N}$. $\forall A \in \mathcal{M}(\mathcal{X})$, we have $A \in \mathcal{B}$ and $\mu(A) < \infty$. Let $\mathcal{A} := (A, \mathcal{B}_A, \mu_A)$ be the finite measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. By Proposition 11.83, $g|_A$ and $g_n|_A$'s are integrable over $\mathcal{A}$, $\lim_{n \in \mathbb{N}} g_n|_A = g|_A$ a.e. in $\mathcal{A}$, and $\int_A g|_A \, d\mu_A = \lim_{n \in \mathbb{N}} \int_A g_n|_A \, d\mu_A$. By Lebesgue Dominated Convergence Theorem 11.88, we have $f|_A$ and $f_n|_A$'s are integrable over $\mathcal{A}$ and $\int_A f|_A \, d\mu_A = \lim_{n \in \mathbb{N}} \int_A f_n|_A \, d\mu_A \in \mathcal{Y}$. Note that $F_A = \int_A f|_A \, d\mu_A \in \mathcal{Y}$ and $F_{n,A} = \int_A f_n|_A \, d\mu_A \in \mathcal{Y}$, $\forall n \in \mathbb{N}$. Then, the nets $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ and $(F_{n,A})_{A \in \mathcal{M}(\mathcal{X})}$'s are well defined.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Definition 11.79, $\exists$ a simple function $\phi : X \rightarrow [0, \infty) \subset \mathbb{R}$ such that $0 \leq \phi(x) \leq g(x)$, $\forall x \in X$, and $\int_X g \, d\mu - \epsilon/5 < \int_X \phi \, d\mu \leq \int_X g \, d\mu < +\infty$. Let $A_0 := \{x \in X \mid \phi(x) > 0\}$. Then, $A_0 \in \mathcal{B}$, $\mu(A_0) < +\infty$, and $A_0 \in \mathcal{M}(\mathcal{X})$. $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have $0 \leq \int_{A \setminus A_0} g|_{A \setminus A_0} \, d\mu_{A \setminus A_0} = \int_A g|_A \, d\mu_A - \int_{A_0} g|_{A_0} \, d\mu_{A_0} \leq \int_X g \, d\mu - \int_{A_0} \phi|_{A_0} \, d\mu_{A_0} = \int_X g \, d\mu - \int_X \phi \, d\mu < \epsilon/5$, where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Note that $\|F_A - F_{A_0}\| = \left\| \int_A f|_A \, d\mu_A - \int_{A_0} f|_{A_0} \, d\mu_{A_0} \right\| = \left\| \int_{A \setminus A_0} f|_{A \setminus A_0} \, d\mu_{A \setminus A_0} \right\| = \lim_{n \in \mathbb{N}} \left\| \int_{A \setminus A_0} f_n|_{A \setminus A_0} \, d\mu_{A \setminus A_0} \right\| \leq \lim_{n \in \mathbb{N}} \int_{A \setminus A_0} g_n|_{A \setminus A_0} \, d\mu_{A \setminus A_0} = \int_{A \setminus A_0} g|_{A \setminus A_0} \, d\mu_{A \setminus A_0} < \epsilon/5$, where the second equality follows from Proposition 11.89; the third equality follows from the preceding paragraph; the first inequality follows from Proposition 11.89; and the fourth equality follows from the preceding paragraph. Then, the net $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ is a Cauchy net, which admits a limit, by Proposition 4.44, $\int_X f \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_A \in \mathcal{Y}$. Hence, f is integrable over $\mathcal{X}$.

$\forall n \in \mathbb{N}$, by Definition 11.79, $\exists$ a simple function $\phi_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ such that $0 \leq \phi_n(x) \leq g_n(x)$, $\forall x \in X$, and $\int_X g_n \, d\mu - \epsilon/5 < \int_X \phi_n \, d\mu \leq \int_X g_n \, d\mu < +\infty$. Let $A_n := \{x \in X \mid \phi_n(x) > 0\}$. Then, $A_n \in \mathcal{B}$, $\mu(A_n) < +\infty$, and $A_n \in \mathcal{M}(\mathcal{X})$. $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_n \subseteq A$, we have $0 \leq \int_{A \setminus A_n} g_n|_{A \setminus A_n} \, d\mu_{A \setminus A_n} = \int_A g_n|_A \, d\mu_A - \int_{A_n} g_n|_{A_n} \, d\mu_{A_n} \leq \int_X g_n \, d\mu - \int_{A_n} \phi_n|_{A_n} \, d\mu_{A_n} = \int_X g_n \, d\mu - \int_X \phi_n \, d\mu < \epsilon/5$, where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Note that $\|F_{n,A} - F_{n,A_n}\| =$

$\left\| \int_A f_n|_A d\mu_A - \int_{A_n} f_n|_{A_n} d\mu_{A_n} \right\| = \left\| \int_{A \setminus A_n} f_n|_{A \setminus A_n} d\mu_{A \setminus A_n} \right\| \leq \int_{A \setminus A_n} g_n|_{A \setminus A_n} d\mu_{A \setminus A_n} < \epsilon/5$, where the second equality and the first inequality follow from Proposition 11.89. Then, the net $(F_{n,A})_{A \in \mathcal{M}(\mathcal{X})}$ is a Cauchy net, which admits a limit, by Proposition 4.44, $\int_X f_n d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_{n,A} \in \mathcal{Y}$.

Let $\bar{g}_n := g_n \wedge g$, $\forall n \in \mathbb{N}$. Then, $0 \leq \bar{g}_n(x) \leq g(x)$, $\forall x \in X$, and $\bar{g}_n$ is $\mathcal{B}$ -measurable by Proposition 11.40, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} \bar{g}_n = g$ a.e. in $\mathcal{X}$ by Propositions 11.52 and 11.53. By Fatou's Lemma 11.80, $\int_X g d\mu \leq \liminf_{n \in \mathbb{N}} \int_X \bar{g}_n d\mu$. By Proposition 11.83, $0 \leq \int_X \bar{g}_n d\mu \leq \int_X g d\mu < +\infty$, $\forall n \in \mathbb{N}$. Then, we have $\int_X g d\mu \leq \liminf_{n \in \mathbb{N}} \int_X \bar{g}_n d\mu \leq \limsup_{n \in \mathbb{N}} \int_X \bar{g}_n d\mu \leq \int_X g d\mu$. Therefore, by Proposition 3.83, $\int_X g d\mu = \lim_{n \in \mathbb{N}} \int_X \bar{g}_n d\mu \in \mathbb{R}$. This coupled with (iii) and the fact that $F_{A_0} = \lim_{n \in \mathbb{N}} F_{n,A_0}$, implies that $\exists n_0 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have $0 \leq \int_X g d\mu - \int_X \bar{g}_n d\mu < \epsilon/5$, $|\int_X g d\mu - \int_X g_n d\mu| < \epsilon/5$, and $\|F_{A_0} - F_{n,A_0}\| < \epsilon/5$. $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have $0 \leq \int_{A \setminus A_0} g_n|_{A \setminus A_0} d\mu_{A \setminus A_0} = \int_{A \setminus A_0} (g_n - \bar{g}_n)|_{A \setminus A_0} d\mu_{A \setminus A_0} + \int_{A \setminus A_0} \bar{g}_n|_{A \setminus A_0} d\mu_{A \setminus A_0} \leq \int_X (g_n - \bar{g}_n) d\mu + \int_{A \setminus A_0} g|_{A \setminus A_0} d\mu_{A \setminus A_0} < \int_X g_n d\mu - \int_X \bar{g}_n d\mu + \epsilon/5 \leq |\int_X g_n d\mu - \int_X g d\mu| + \int_X g d\mu - \int_X \bar{g}_n d\mu + \epsilon/5 < 3\epsilon/5$, where the first equality, the second inequality, and the third inequality follow from Proposition 11.83. Furthermore, $\|F_{n,A} - F_{n,A_0}\| = \left\| \int_A f_n|_A d\mu_A - \int_{A_0} f_n|_{A_0} d\mu_{A_0} \right\| = \left\| \int_{A \setminus A_0} f_n|_{A \setminus A_0} d\mu_{A \setminus A_0} \right\| \leq \int_{A \setminus A_0} g_n|_{A \setminus A_0} d\mu_{A \setminus A_0} < 3\epsilon/5$, where the second equality and the first inequality follows from Proposition 11.89. This implies that $\left\| \int_X f d\mu - \int_X f_n d\mu \right\| \leq \left\| \int_X f d\mu - F_{A_0} \right\| + \|F_{A_0} - F_{n,A_0}\| + \|F_{n,A_0} - \int_X f_n d\mu\| < \lim_{A \in \mathcal{M}(\mathcal{X})} \|F_A - F_{A_0}\| + \epsilon/5 + \lim_{A \in \mathcal{M}(\mathcal{X})} \|F_{n,A_0} - F_{n,A}\| \leq \epsilon/5 + \epsilon/5 + 3\epsilon/5 = \epsilon$, where the second inequality follows from Propositions 7.21 and 3.66. Hence, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu \in \mathcal{Y}$.

Then, in both cases, we have shown that f is integrable over $\mathcal{X}$ and $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu \in \mathcal{Y}$.

Note that $0 \leq \mathcal{P} \circ f_n(x) \leq g_n(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \mathcal{P} \circ f_n = \mathcal{P} \circ f$ a.e. in $\mathcal{X}$ by Propositions 7.21 and 11.52, and $\mathcal{P} \circ f$ and $\mathcal{P} \circ f_n$'s are $\mathcal{B}$ -measurable by Propositions 7.21 and 11.38. Then, by what we have proved so far, $\int_X \mathcal{P} \circ f d\mu = \lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ f_n d\mu \leq \lim_{n \in \mathbb{N}} \int_X g_n d\mu = \int_X g d\mu < \infty$. Hence, f is absolutely integrable.

This completes the proof of the theorem. $\square$

Proposition 11.92 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $f_i : X \rightarrow \mathcal{W}$ be absolutely integrable over $\mathcal{X}$, $i = 1, 2$. Then, the following statements hold.*

- (i) f_i is integrable over $\mathcal{X}$ and $\int_X f_i d\mu \in \mathcal{Y}$, $i = 1, 2$.
- (ii) $f_1 + f_2$ is absolutely integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu \in \mathcal{Y}$.

- (iii) $\forall A \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, Af_1 is absolutely integrable over $\mathcal{X}$ and $\int_X (Af_1) d\mu = A \int_X f_1 d\mu \in \mathcal{Z}$.
- (iv) $\forall c \in \mathbb{K}$, cf_1 is absolutely integrable over $\mathcal{X}$ and $\int_X (cf_1) d\mu = c \int_X f_1 d\mu \in \mathcal{Y}$.
- (v) $\forall H \in \mathcal{B}$, $f_1|_H$ is absolutely integrable over $\mathcal{H}$ and $\int_X (f_1 \chi_{H,X}) d\mu = \int_H f_1|_H d\mu_H \in \mathcal{Y}$, where $\mathcal{H} := (H, \mathcal{B}_H, \mu_H)$ is the measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. We will henceforth denote $\int_H f_1|_H d\mu_H$ by $\int_H f_1 d\mu$.
- (vi) If $f_1 = f_2$ a.e. in $\mathcal{X}$ then $\int_X f_1 d\mu = \int_X f_2 d\mu \in \mathcal{Y}$.
- (vii) $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, $\sum_{i=1}^\infty \int_{E_i} f_1 d\mu = \int_{\bigcup_{i=1}^\infty E_i} f_1 d\mu \in \mathcal{Y}$.
- (viii) $0 \leq \|\int_X f_1 d\mu\| \leq \int_X \mathcal{P} \circ f_1 d\mu < +\infty$.
- (ix) If $\mathcal{Y}$ admits a positive cone P and $f_1 \leq f_2$ a.e. in $\mathcal{X}$, then $\int_X f_1 d\mu \leq \int_X f_2 d\mu$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu(X) < +\infty$; Case 2: $\mu(X) = +\infty$. Case 1: $\mu(X) < +\infty$. The results follow immediately from Propositions 11.89 and 11.83 and their proof.

Case 2: $\mu(X) = +\infty$. $f_1, f_2, \mathcal{P} \circ f_1$, and $\mathcal{P} \circ f_2$ are $\mathcal{B}$ -measurable by Propositions 7.21 and 11.38.

(i) By Lebesgue Dominated Convergence Theorem 11.91, we have that f_i is integrable over $\mathcal{X}$ and $\int_X f_i d\mu \in \mathcal{Y}$, $i = 1, 2$. Let $(F_{i,E})_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X f_i d\mu$ as defined in Definition 11.71, $i = 1, 2$. $\forall E \in \mathcal{M}(\mathcal{X})$, $F_{i,E} = \int_E f_i|_E d\mu_E$, $i = 1, 2$. Then, by Definition 11.71, we have the net $(F_{i,E})_{E \in \mathcal{M}(\mathcal{X})}$ is well-defined, $i = 1, 2$, and $\lim_{E \in \mathcal{M}(\mathcal{X})} F_{i,E} = \int_X f_i d\mu \in \mathcal{Y}$, $i = 1, 2$.

(ii) By Propositions 11.38, 11.39, 7.21, and 7.23, $f_1 + f_2$ and $\mathcal{P} \circ (f_1 + f_2)$ are $\mathcal{B}$ -measurable. Note that $\mathcal{P} \circ (f_1 + f_2)(x) = \|f_1(x) + f_2(x)\| \leq \|f_1(x)\| + \|f_2(x)\| = (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x)$, $\forall x \in X$. By Proposition 11.83, $0 \leq \int_X \mathcal{P} \circ (f_1 + f_2) d\mu \leq \int_X (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2) d\mu = \int_X \mathcal{P} \circ f_1 d\mu + \int_X \mathcal{P} \circ f_2 d\mu < +\infty$. Hence, $f_1 + f_2$ is absolutely integrable over $\mathcal{X}$. Let $(\bar{F}_E)_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X (f_1 + f_2) d\mu$ as defined in Definition 11.71. $\forall E \in \mathcal{M}(\mathcal{X})$, $\bar{F}_E = \int_E (f_1 + f_2)|_E d\mu_E = \int_E f_1|_E d\mu_E + \int_E f_2|_E d\mu_E = F_{1,E} + F_{2,E}$, where the second equality follows from Proposition 11.89. Then, $\int_X (f_1 + f_2) d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} \bar{F}_E = \lim_{E \in \mathcal{M}(\mathcal{X})} (F_{1,E} + F_{2,E}) = \int_X f_1 d\mu + \int_X f_2 d\mu \in \mathcal{Y}$, where the third equality follows from Propositions 3.66, 3.67, and 7.23. Hence, $f_1 + f_2$ is integrable over $\mathcal{X}$.

(iii) By Propositions 11.38, 7.21, and 7.62, Af_1 and $\mathcal{P} \circ (Af_1)$ are $\mathcal{B}$ -measurable. Note that $\mathcal{P} \circ (Af_1)(x) = \|Af_1(x)\| \leq \|A\| \|f_1(x)\| = \|A\| \mathcal{P} \circ f_1(x)$, $\forall x \in X$. By Proposition 11.83 and the fact that $\int_X \mathcal{P} \circ f_1 d\mu \in \mathbb{R}$,

we have $0 \leq \int_X \mathcal{P} \circ (Af_1) d\mu \leq \int_X (\|A\| \mathcal{P} \circ f_1) d\mu = \|A\| \int_X \mathcal{P} \circ f_1 d\mu < +\infty$. Hence, Af_1 is absolutely integrable over $\mathcal{X}$. Let $(\bar{F}_E)_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X (Af_1) d\mu$ as defined in Definition 11.71. $\forall E \in \mathcal{M}(\mathcal{X})$, $\bar{F}_E = \int_E (Af_1)|_E d\mu_E = A \int_E f_1|_E d\mu_E = AF_{1,E}$, where the second equality follows from Proposition 11.89. Then, by Propositions 3.66 and 7.62, we have $\int_X (Af_1) d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} \bar{F}_E = \lim_{E \in \mathcal{M}(\mathcal{X})} AF_{1,E} = A \int_X f_1 d\mu \in \mathcal{Z}$. Hence, Af_1 is integrable over $\mathcal{X}$.

(iv) This follows immediately from (iii).

(v) Fix any $H \in \mathcal{B}$. By Proposition 11.41, $f_1|_H$ and $\mathcal{P} \circ (f_1|_H) = (\mathcal{P} \circ f_1)|_H$ are $\mathcal{B}_H$ -measurable. By Proposition 11.83, $f_1|_H$ is absolutely integrable over $\mathcal{H}$. Then, by (i), $f_1|_H$ is integrable over $\mathcal{H}$. Again, by (i), $f_1\chi_{H,X}$ is integrable over $\mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists A_0 \in \mathcal{M}(\mathcal{X})$ such that $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have $\|\int_A (f_1\chi_{H,X})|_A d\mu_A - \int_X (f_1\chi_{H,X}) d\mu\| < \epsilon/2$. We will distinguish two exhaustive and mutually exclusive subcases: Case 2a: $\mu(H) = +\infty$; Case 2b: $\mu(H) < +\infty$.

Case 2a: $\mu(H) = +\infty$. $\exists E_0 \in \mathcal{M}(\mathcal{H})$ such that $\forall E \in \mathcal{M}(\mathcal{H})$ with $E_0 \subseteq E$, we have $\|\int_E f_1|_E d\mu_E - \int_H f_1|_H d\mu_H\| < \epsilon/2$. Let $A_1 := A_0 \cup E_0 \in \mathcal{M}(\mathcal{X})$ and $E_1 := A_1 \cap H \in \mathcal{M}(\mathcal{H})$. By Proposition 11.89, $\int_{A_1} (f_1\chi_{H,X})|_{A_1} d\mu_{A_1} = \int_{H \cap A_1} f_1|_{H \cap A_1} d\mu_{H \cap A_1} = \int_{E_1} f_1|_{E_1} d\mu_{E_1}$. Then, we have $\|\int_X (f_1\chi_{H,X}) d\mu - \int_H f_1|_H d\mu_H\| \leq \left\| \int_X (f_1\chi_{H,X}) d\mu - \int_{A_1} (f_1\chi_{H,X})|_{A_1} d\mu_{A_1} \right\| + \left\| \int_{E_1} f_1|_{E_1} d\mu_{E_1} - \int_H f_1|_H d\mu_H \right\| < \epsilon/2 + \epsilon/2 = \epsilon$. By the arbitrariness of ϵ , we have $\int_X (f_1\chi_{H,X}) d\mu = \int_H f_1|_H d\mu_H \in \mathcal{Y}$.

Case 2b: $\mu(H) < +\infty$. Let $A_1 := A_0 \cup H \in \mathcal{M}(\mathcal{X})$. By Proposition 11.89, $\int_{A_1} (f_1\chi_{H,X})|_{A_1} d\mu_{A_1} = \int_H f_1|_H d\mu_H$. Then, we have $\|\int_X (f_1\chi_{H,X}) d\mu - \int_H f_1|_H d\mu_H\| = \left\| \int_X (f_1\chi_{H,X}) d\mu - \int_{A_1} (f_1\chi_{H,X})|_{A_1} d\mu_{A_1} \right\| < \epsilon/2$. By the arbitrariness of ϵ , we have $\int_X (f_1\chi_{H,X}) d\mu = \int_H f_1|_H d\mu_H \in \mathcal{Y}$.

Hence, in both subcases, $\int_X (f_1\chi_{H,X}) d\mu = \int_H f_1|_H d\mu_H \in \mathcal{Y}$.

(vi) This follows immediately from Lebesgue Dominated Convergence Theorem 11.91.

(vii) $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$ and $\bar{E}_n := \bigcup_{i=1}^n E_i \in \mathcal{B}$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} f_1(x)\chi_{\bar{E}_n,X}(x) = f_1(x)\chi_{E,X}(x)$, $\forall x \in X$. Then, we have $\sum_{i=1}^\infty \int_{E_i} f_1 d\mu = \lim_{n \in \mathbb{N}} \sum_{i=1}^n \int_{E_i} f_1 d\mu = \lim_{n \in \mathbb{N}} \int_X (f_1\chi_{\bar{E}_n,X}) d\mu = \lim_{n \in \mathbb{N}} \int_X (f_1\chi_{E,X}) d\mu = \int_E f_1 d\mu \in \mathcal{Y}$, where the first equality follows from (v); the second equality follows from (ii); the third equality follows from Lebesgue Dominated Convergence Theorem 11.91; the last equality follows from (v); and the last inclusion follows from (v).

(viii) By (i), $\mathcal{P} \circ f_1$ is integrable over $\mathcal{X}$. $\forall E \in \mathcal{M}(\mathcal{X})$, by Proposition 11.89, $\|F_{1,E}\| \leq \int_E \mathcal{P} \circ f_1 d\mu \leq \int_X \mathcal{P} \circ f_1 d\mu < +\infty$, where the second inequality follows from Proposition 11.83. Then, $0 \leq \|\int_X f_1 d\mu\| = \lim_{E \in \mathcal{M}(\mathcal{X})} \|F_{1,E}\| \leq \int_X \mathcal{P} \circ f_1 d\mu < +\infty$, where the equality follows from Propositions 3.66 and 7.21.

(ix) $\forall E \in \mathcal{M}(\mathcal{X})$, by Proposition 11.89 and (v), we have $\int_E (f_2 - f_1)|_E \cdot d\mu_E = \int_E f_2|_E d\mu_E - \int_E f_1|_E d\mu_E \in P$. Note that, by (i) and (iv), $\int_X f_2 d\mu - \int_X f_1 d\mu = \int_X (f_2 - f_1) d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} \int_E (f_2 - f_1)|_E d\mu_E \in \overline{P} = P$, where the second to last step follows from Proposition 3.68; and the last step follows from the fact that P is closed. Therefore, we have $\int_X f_1 d\mu \leq \int_X f_2 d\mu$.

This completes the proof of the proposition. $\square$

Proposition 11.93 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}_i$ be a Banach space over $\mathbb{K}$, $\mathcal{W}_i \subseteq \mathcal{Y}_i$ be a separable subspace, $f_i : X \rightarrow \mathcal{W}_i$, $i = 1, 2$, and $h : X \rightarrow \mathcal{W}_1 \times \mathcal{W}_2$ be defined by $h(x) = (f_1(x), f_2(x))$, $\forall x \in X$. Then, h is absolutely integrable over $\mathcal{X}$ if, and only if, f_1 and f_2 are absolutely integrable over $\mathcal{X}$. In this case, we have $\int_X h d\mu = (\int_X f_1 d\mu, \int_X f_2 d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2$.*

Proof Let f_1 and f_2 be absolutely integrable over $\mathcal{X}$. By Proposition 11.39, h is $\mathcal{B}$ -measurable. By Propositions 7.21 and 11.38, $\mathcal{P} \circ h$ is $\mathcal{B}$ -measurable. $\forall x \in X$, $\mathcal{P} \circ h(x) = \|(f_1(x), f_2(x))\| = (\|f_1(x)\|^2 + \|f_2(x)\|^2)^{1/2} \leq \mathcal{P} \circ f_1(x) + \mathcal{P} \circ f_2(x)$. By Proposition 11.83, $0 \leq \int_X (\mathcal{P} \circ h) d\mu \leq \int_X (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2) d\mu = \int_X (\mathcal{P} \circ f_1) d\mu + \int_X (\mathcal{P} \circ f_2) d\mu < \infty$. Hence, h is absolutely integrable over $\mathcal{X}$. By Propositions 4.31 and 7.22, $\mathcal{Y}_1 \times \mathcal{Y}_2$ is a Banach space over $\mathbb{K}$. By Propositions 7.22, 3.28, and 4.4, $\mathcal{W}_1 \times \mathcal{W}_2 \subseteq \mathcal{Y}_1 \times \mathcal{Y}_2$ is a separable subspace. By Proposition 11.92, h , f_1 , and f_2 are integrable over $\mathcal{X}$ and $\int_X h d\mu \in \mathcal{Y}_1 \times \mathcal{Y}_2$.

On the other hand, let h be absolutely integrable over $\mathcal{X}$. Then, by Proposition 11.39, f_1 and f_2 are $\mathcal{B}$ -measurable. Fix any $i \in \{1, 2\}$. By Propositions 7.21 and 11.38, $\mathcal{P} \circ f_i$ is $\mathcal{B}$ -measurable. $\forall x \in X$, $\mathcal{P} \circ f_i(x) \leq \mathcal{P} \circ h(x)$. Then, by Proposition 11.83, $0 \leq \int_X \mathcal{P} \circ f_i d\mu \leq \int_X \mathcal{P} \circ h d\mu < \infty$. Hence, f_i is absolutely integrable over $\mathcal{X}$. By Proposition 11.92, f_i is integrable over $\mathcal{X}$ and $\int_X f_i d\mu \in \mathcal{Y}_i$.

Hence, h is absolutely integrable over $\mathcal{X}$ if, and only if, f_1 and f_2 are absolutely integrable over $\mathcal{X}$.

Let h , f_1 , and f_2 be absolutely integrable over $\mathcal{X}$. We will show that $\int_X h d\mu = (\int_X f_1 d\mu, \int_X f_2 d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ by distinguishing two exhaustive and mutually exclusives cases: Case 1: $\mu(X) < \infty$; Case 2: $\mu(X) = \infty$.

Case 1: $\mu(X) < \infty$. By Lemma 11.87, $\forall i \in \{1, 2\}$, $\exists$ sequence of simple functions $(\varphi_{i,n})_{n=1}^\infty$, $\varphi_{i,n} : X \rightarrow \mathcal{W}_i$, $\forall n \in \mathbb{N}$, such that $\lim_{n \in \mathbb{N}} \varphi_{i,n} = f_i$ a.e. in $\mathcal{X}$, $\|\varphi_{i,n}(x)\| \leq \|f_i(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\int_X f_i d\mu = \lim_{n \in \mathbb{N}} \int_X \varphi_{i,n} d\mu \in \mathcal{Y}_i$. By Proposition 11.53, we have $\lim_{n \in \mathbb{N}} \psi_n = h$ a.e. in $\mathcal{X}$, where $\psi_n : X \rightarrow \mathcal{W}_1 \times \mathcal{W}_2$ is defined by $\psi_n(x) = (\varphi_{1,n}(x), \varphi_{2,n}(x))$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Clearly, ψ_n is a simple function, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91, $\int_X h d\mu = \lim_{n \in \mathbb{N}} \int_X \psi_n d\mu = \lim_{n \in \mathbb{N}} (\int_X \varphi_{1,n} d\mu, \int_X \varphi_{2,n} d\mu) = (\int_X f_1 d\mu, \int_X f_2 d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2$, where the second equality follows from Proposition 11.75; and the third equality follows from Proposition 3.67.

Case 2: $\mu(X) = \infty$. Let $(H_E)_{E \in \mathcal{M}(\mathcal{X})}$ and $(F_{i,E})_{E \in \mathcal{M}(\mathcal{X})}$ be the nets for $\int_X h \, d\mu$ and $\int_X f_i \, d\mu$ as defined in Definition 11.71, respectively, $i = 1, 2$. $\forall E \in \mathcal{M}(\mathcal{X})$, by Case 1,

$$H_E = \int_E h|_E \, d\mu_E = \left(\int_E f_1|_E \, d\mu_E, \int_E f_2|_E \, d\mu_E \right) = (F_{1,E}, F_{2,E})$$

where $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$ is the finite measure subspace of $\mathcal{X}$. Since h , f_1 , and f_2 are integrable over $\mathcal{X}$, we have $\int_X h \, d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} H_E = \lim_{E \in \mathcal{M}(\mathcal{X})} (F_{1,E}, F_{2,E}) = (\int_X f_1 \, d\mu, \int_X f_2 \, d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2$, where the last equality follows from Proposition 3.67.

Hence, in both cases, we have $\int_X h \, d\mu = (\int_X f_1 \, d\mu, \int_X f_2 \, d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2$. This completes the proof of the proposition. $\square$

Proposition 11.94 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $U \subseteq \mathcal{W}$ be a conic segment, and $f : X \rightarrow U$ be absolutely integrable over $\mathcal{X}$. Then, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : X \rightarrow U$, $\forall n \in \mathbb{N}$, such that $\|\varphi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$, $\lim_{n \in \mathbb{N}} \int_X \varphi_n \, d\mu = \int_X f \, d\mu \in \mathcal{Y}$, $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \varphi_n \, d\mu = \int_X \mathcal{P} \circ f \, d\mu \in [0, \infty) \subset \mathbb{R}$, and $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ (\varphi_n - f) \, d\mu = 0$.*

Proof By Proposition 11.66, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : X \rightarrow U$, $\forall n \in \mathbb{N}$, such that $\|\varphi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$. By Lebesgue Dominated Convergence Theorem 11.91, $\lim_{n \in \mathbb{N}} \int_X \varphi_n \, d\mu = \int_X f \, d\mu \in \mathcal{Y}$. Note that $\lim_{n \in \mathbb{N}} \mathcal{P} \circ \varphi_n = \mathcal{P} \circ f$ a.e. in $\mathcal{X}$, by Propositions 7.21 and 11.52. Again by Lebesgue Dominated Convergence Theorem 11.91, we have $\lim_{n \in \mathbb{N}} \int_X (\mathcal{P} \circ \varphi_n) \, d\mu = \int_X (\mathcal{P} \circ f) \, d\mu \in [0, \infty) \subset \mathbb{R}$. By Lemma 11.43, $\mathcal{P} \circ (\varphi_n - f)$ is $\mathcal{B}$ -measurable. By Propositions 7.21, 7.23, 11.53, and 11.52, we have $\lim_{n \in \mathbb{N}} \mathcal{P} \circ (\varphi_n - f) = 0$ a.e. in $\mathcal{X}$. Note that $0 \leq \mathcal{P} \circ (\varphi_n - f)(x) = \|\varphi_n(x) - f(x)\| \leq \|\varphi_n(x)\| + \|f(x)\| \leq 2\mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91 and Proposition 11.75, we have $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ (\varphi_n - f) \, d\mu = 0$. This completes the proof of the proposition. $\square$

Proposition 11.95 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $U \subseteq \mathcal{W}$ be a conic segment, and $f : X \rightarrow U$ be absolutely integrable over $\mathcal{X}$. Then, there exists a sequence of simple functions $(\phi_n)_{n=1}^\infty$, $\phi_n : X \rightarrow U$, $\forall n \in \mathbb{N}$, such that $\|\phi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu = \int_X f \, d\mu \in \mathcal{Y}$, $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \phi_n \, d\mu = \int_X \mathcal{P} \circ f \, d\mu \in [0, \infty) \subset \mathbb{R}$, and $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ (\phi_n - f) \, d\mu = 0$.*

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu(X) < +\infty$; Case 2: $\mu(X) = +\infty$.

Case 1: $\mu(X) < +\infty$. Then, $\mathcal{X}$ is a finite measure space. The result follows immediately from Proposition 11.94.

Case 2: $\mu(X) = +\infty$. Since $\int_X \mathcal{P} \circ f \, d\mu < +\infty$, then, by Definition 11.71, $\forall n \in \mathbb{N}$, $\exists E_n \in \mathcal{B}$ with $\mu(E_n) < +\infty$ such that $\int_X \mathcal{P} \circ f \, d\mu - 2^{-n-1} < \int_{E_n} \mathcal{P} \circ f \, d\mu \leq \int_X \mathcal{P} \circ f \, d\mu$. By Case 1, $\exists$ a simple function $\bar{\phi}_n : E_n \rightarrow U$ such that $\|\bar{\phi}_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in E_n$, $\left\| \int_{E_n} \bar{\phi}_n \, d\mu_{E_n} - \int_{E_n} f \, d\mu \right\| < 2^{-n-1}$, $\left| \int_{E_n} \mathcal{P} \circ \bar{\phi}_n \, d\mu_{E_n} - \int_{E_n} \mathcal{P} \circ f \, d\mu \right| < 2^{-n-1}$, and $0 \leq \int_{E_n} \mathcal{P} \circ (\bar{\phi}_n - f)|_{E_n} \, d\mu_{E_n} < 2^{-n-1}$, where $(E_n, \mathcal{B}_{E_n}, \mu_{E_n})$ is the finite measure subspace of $\mathcal{X}$. This simple function $\bar{\phi}_n$ may be extended to a simple function $\phi_n : X \rightarrow U$ such that $\phi_n|_{E_n} = \bar{\phi}_n$ and $\phi_n(x) = \vartheta_y$, $\forall x \in X \setminus E_n$. Then, we have $\|\phi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, and $\left\| \int_X \phi_n \, d\mu - \int_X f \, d\mu \right\| = \left\| \int_X (\phi_n \chi_{E_n, X}) \, d\mu - \int_{E_n} f \, d\mu - \int_{X \setminus E_n} f \, d\mu \right\| = \left\| \int_{E_n} \bar{\phi}_n \, d\mu_{E_n} - \int_{E_n} f \, d\mu - \int_{X \setminus E_n} f \, d\mu \right\| \leq \left\| \int_{E_n} \bar{\phi}_n \, d\mu_{E_n} - \int_{E_n} f \, d\mu \right\| + \left\| \int_{X \setminus E_n} f \, d\mu \right\| < 2^{-n-1} + \int_{X \setminus E_n} \mathcal{P} \circ f \, d\mu = 2^{-n-1} + \int_X \mathcal{P} \circ f \, d\mu - \int_{E_n} \mathcal{P} \circ f \, d\mu < 2^{-n}$, where the first and the second equalities follow from Proposition 11.92; the second inequality follows from Proposition 11.92; and the third equality follows from Proposition 11.83. Next, we have $\left| \int_X \mathcal{P} \circ \phi_n \, d\mu - \int_X \mathcal{P} \circ f \, d\mu \right| \leq \left| \int_X ((\mathcal{P} \circ \phi_n) \chi_{E_n, X}) \, d\mu - \int_{E_n} \mathcal{P} \circ f \, d\mu \right| + \left| \int_X \mathcal{P} \circ f \, d\mu - \int_{E_n} \mathcal{P} \circ f \, d\mu \right| < \left| \int_{E_n} \mathcal{P} \circ \bar{\phi}_n \, d\mu_{E_n} - \int_{E_n} \mathcal{P} \circ f \, d\mu \right| + 2^{-n-1} < 2^{-n}$, where the second inequality follows from Proposition 11.92. Finally, we have $0 \leq \int_X \mathcal{P} \circ (\phi_n - f) \, d\mu = \int_{E_n} \mathcal{P} \circ ((\phi_n - f)|_{E_n}) \, d\mu_{E_n} + \int_{X \setminus E_n} \mathcal{P} \circ ((\phi_n - f)|_{X \setminus E_n}) \, d\mu_{X \setminus E_n} = \int_{E_n} \mathcal{P} \circ (\bar{\phi}_n - f)|_{E_n} \, d\mu_{E_n} + \int_{X \setminus E_n} \mathcal{P} \circ (f|_{X \setminus E_n}) \, d\mu_{X \setminus E_n} < 2^{-n-1} + \int_{X \setminus E_n} \mathcal{P} \circ f \, d\mu < 2^{-n}$, where the first equality follows from Propositions 11.83. Hence, $(\phi_n)_{n=1}^\infty$ is the sequence we seek.

This completes the proof of the proposition. $\square$

Proposition 11.96 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, and $f_1 : X \rightarrow \mathbb{R}$ and $f_2 : X \rightarrow \mathbb{R}$ be $\mathcal{B}$ -measurable. Assume that $\int_E f_1 \, d\mu \leq \int_E f_2 \, d\mu$, $\forall E \in \mathcal{B}$ with $\int_E \mathcal{P} \circ f_i \, d\mu < \infty$, $i = 1, 2$. Then, $f_1 \leq f_2$ a.e. in $\mathcal{X}$.*

Proof By Propositions 7.21, 7.23, 11.38, and 11.39, $f_1 - f_2$, $\mathcal{P} \circ f_1$, and $\mathcal{P} \circ f_2$ are $\mathcal{B}$ -measurable. Let $E := \{x \in X \mid f_2(x) - f_1(x) < 0\} \in \mathcal{B}$. Suppose $\mu(E) > 0$. $E = \bigcup_{n=1}^\infty E_n := \bigcup_{n=1}^\infty \{x \in X \mid f_2(x) - f_1(x) \leq -1/n, |f_1(x)| \leq n, |f_2(x)| \leq n\}$. By Proposition 11.7, $\mu(E) = \lim_{n \in \mathbb{N}} \mu(E_n)$ and $\exists n \in \mathbb{N}$ such that $\mu(E_n) > 0$. Since $\mathcal{X}$ is σ -finite, then $\exists (X_m)_{m=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{m=1}^\infty X_m$ and $\mu(X_m) < +\infty$, $\forall m \in \mathbb{N}$. Without loss of generality, we may assume that $X_m \subseteq X_{m+1}$, $\forall m \in \mathbb{N}$. Then, by Proposition 11.7, $\mu(E_n) = \lim_{m \in \mathbb{N}} \mu(E_n \cap X_m)$. Then, $\exists m \in \mathbb{N}$ such that $\mu(\bar{E}) := \mu(E_n \cap X_m) \in (0, +\infty) \subset \mathbb{R}$. By Definition 11.79, $\int_{\bar{E}} \mathcal{P} \circ f_i \, d\mu \leq n\mu(\bar{E}) < \infty$, $i = 1, 2$. Then, by Propositions 11.75 and 11.92, $\mathbb{R} \ni \int_{\bar{E}} f_1 \, d\mu - \int_{\bar{E}} f_2 \, d\mu - \mu(\bar{E})/n = \int_{\bar{E}} (f_1 - f_2 - 1/n) \, d\mu \geq 0$. This contradicts with the assumption. Hence, $\mu(E) = 0$ and $f_1 \leq f_2$ a.e. in $\mathcal{X}$. $\square$

Proposition 11.97 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $f_i : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $i = 1, 2$. Assume that $f_1 \leq f_2$ a.e. in $\mathcal{X}$, $\mu(\{x \in X \mid f_1(x) < f_2(x)\}) > 0$, and $0 \leq \int_X f_1 d\mu < +\infty$. Then, $\int_X f_1 d\mu < \int_X f_2 d\mu$.*

Proof By Proposition 11.83, $\int_X f_1 d\mu \leq \int_X f_2 d\mu$. Suppose $\int_X f_2 d\mu = \int_X f_1 d\mu \in \mathbb{R}$. Then, f_1 and f_2 are absolutely integrable over $\mathcal{X}$. Let $g := f_2 - f_1$. Then, by Propositions 7.23, 11.38, and 11.39, we have $g : X \rightarrow \mathbb{R}$ is $\mathcal{B}$ -measurable and $g \geq 0$ a.e. in $\mathcal{X}$. By Proposition 11.92, g is absolutely integrable over $\mathcal{X}$ and $\int_X g d\mu = 0$. $\forall E \in \mathcal{B}$, by Proposition 11.92, we have $\int_E g d\mu + \int_{X \setminus E} g d\mu = \int_X g d\mu = 0$ and both of the summand on the left-hand-side are nonnegative. Then, $\int_E g d\mu = 0$. By Proposition 11.96, we have $g = 0$ a.e. in $\mathcal{X}$. This contradicts with the fact that $\mu(\{x \in X \mid g(x) > 0\}) = \mu(\{x \in X \mid f_1(x) < f_2(x)\}) > 0$. Therefore, we must have $\int_X f_1 d\mu < \int_X f_2 d\mu$. This completes the proof of the proposition. $\square$

Theorem 11.98 (Jensen's Inequality) *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space with $\mu(X) = 1$, $\mathcal{Y}$ be a real Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $\Omega \subseteq \mathcal{Y}$ be a nonempty closed convex set, $f : X \rightarrow \Omega \cap \mathcal{W}$ be absolutely integrable over $\mathcal{X}$, and $G : \Omega \rightarrow \mathbb{R}$ be a convex functional. Assume that $G \circ f$ is absolutely integrable over $\mathcal{X}$ and the epigraph $[G, \Omega]$ is closed. Then, $\int_X f d\mu \in \Omega$ and $G(\int_X f d\mu) \leq \int_X (G \circ f) d\mu \in \mathbb{R}$.*

Proof By Proposition 11.92, $y_0 := \int_X f d\mu \in \mathcal{Y}$. We will first show that $y_0 \in \Omega$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $y_0 \in \Omega$; Case 2: $y_0 \notin \Omega$.

Case 1: $y_0 \in \Omega$. Then, $\Omega \cap \mathcal{W}$ is a conic segment. By Lemma 11.87, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : \mathcal{X} \rightarrow \Omega \cap \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$, $\|\varphi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $y_0 = \int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu \in \mathcal{Y}$. Fix any $n \in \mathbb{N}$. Let φ_n admit the canonical representation $\varphi_n = \sum_{i=1}^{\bar{n}} y_i \chi_{A_i, X}$. Let $y_{\bar{n}+1} := y_0 \in \Omega \cap \mathcal{W}$ and $A_{\bar{n}+1} := X \setminus (\bigcup_{i=1}^{\bar{n}} A_i) \in \mathcal{B}$. Then, by Proposition 11.75, $\int_X \varphi_n d\mu = \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i)$. Note that $1 = \mu(X) = \sum_{i=1}^{\bar{n}+1} \mu(A_i)$ and the summands are nonnegative. Since Ω is convex, then $\int_X \varphi_n d\mu \in \Omega$. Since Ω is closed, then, by Proposition 4.13, $y_0 = \lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu \in \Omega$.

Case 2: $y_0 \notin \Omega$. Note that $\mu(X) = 1$ implies that $X \neq \emptyset$. Then, $\Omega \cap \mathcal{W} \neq \emptyset$. Let $\bar{y} \in \Omega \cap \mathcal{W}$ and $\bar{\Omega} := \Omega - \bar{y}$. Then, $y_0 \in \bar{\Omega}$ and, by Proposition 7.16, $\bar{\Omega}$ is a closed convex set. Let $\bar{f} := f - \bar{y}$. Then, we have $\bar{f} : X \rightarrow \bar{\Omega} \cap \mathcal{W}$. By Propositions 11.38, 11.39, and 7.23, $\bar{f}$ is $\mathcal{B}$ -measurable. Note that $\mathcal{P} \circ \bar{f}(x) = \|f(x) - \bar{y}\| \leq \mathcal{P} \circ f(x) + \|\bar{y}\|$, $\forall x \in X$. Then, by Propositions 11.83 and 11.75, $0 \leq \int_X (\mathcal{P} \circ \bar{f}) d\mu \leq \int_X (\mathcal{P} \circ f + \|\bar{y}\|) d\mu = \int_X (\mathcal{P} \circ f) d\mu + \int_X \|\bar{y}\| d\mu = \int_X (\mathcal{P} \circ f) d\mu + \|\bar{y}\| < +\infty$. Hence, $\bar{f}$ is absolutely integrable over $\mathcal{X}$. By Case 1, we have $\int_X \bar{f} d\mu \in \bar{\Omega}$. By Propositions 11.92 and 11.75, we have $\int_X f d\mu = \int_X \bar{f} d\mu + \bar{y} \in \Omega$.

Hence, in both cases, we have $y_0 \in \Omega$. By Proposition 8.32, we have $G(y_0) = \sup_{y_* \in \Omega_{\text{conj}}} (\langle y_*, y_0 \rangle - G_{\text{conj}}(y_*))$, where $G_{\text{conj}} : \Omega_{\text{conj}} \rightarrow \mathbb{R}$ is the conjugate functional to G . $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists y_{*\epsilon} \in \Omega_{\text{conj}}$, such that $G(y_0) - \epsilon \leq \langle y_{*\epsilon}, y_0 \rangle - G_{\text{conj}}(y_{*\epsilon}) \leq \langle y_{*\epsilon}, y_0 \rangle - \langle y_{*\epsilon}, y \rangle + G(y) = G(y) - \langle y_{*\epsilon}, y - y_0 \rangle$, $\forall y \in \Omega$. Then, $\forall x \in X$, we have $G(y_0) - \epsilon \leq G(f(x)) - \langle y_{*\epsilon}, f(x) - y_0 \rangle$. By Proposition 11.92, we have

$$\begin{aligned} G(y_0) - \epsilon &= \int_X (G(y_0) - \epsilon) d\mu \leq \int_X (G \circ f - \langle y_{*\epsilon}, f - y_0 \rangle) d\mu \\ &= \int_X (G \circ f) d\mu - \int_X \langle y_{*\epsilon}, f \rangle d\mu + \int_X \langle y_{*\epsilon}, y_0 \rangle d\mu \\ &= \int_X (G \circ f) d\mu - \langle y_{*\epsilon}, \int_X f d\mu \rangle + \langle y_{*\epsilon}, y_0 \rangle = \int_X (G \circ f) d\mu \in \mathbb{R} \end{aligned}$$

By the arbitrariness of ϵ , we have $\mathbb{R} \ni G(\int_X f d\mu) \leq \int_X (G \circ f) d\mu \in \mathbb{R}$. This completes the proof of the theorem. $\square$

11.6 Banach Space Valued Measures

Definition 11.99 Let $(X, \mathcal{B})$ be a measurable space and $\mathcal{Y}$ be a normed linear space. A $\mathcal{Y}$ -valued pre-measure μ on $(X, \mathcal{B})$ is a function $\mu : \mathcal{B} \rightarrow \mathcal{Y}$ such that

$$(i) \quad \mu(\emptyset) = \vartheta_{\mathcal{Y}};$$

$$(ii) \quad \forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{B}, \text{ which is pairwise disjoint, we have } \sum_{i=1}^{\infty} \|\mu(E_i)\| < +\infty \text{ and } \mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i).$$

Then, the triple $(X, \mathcal{B}, \mu)$ is said to be a $\mathcal{Y}$ -valued pre-measure space.

Define $\mathcal{P} \circ \mu : \mathcal{B} \rightarrow [0, \infty] \subset \mathbb{R}_e$ by, $\forall E \in \mathcal{B}$, $\mathcal{P} \circ \mu(E) := \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(E_i)\|$. $\mathcal{P} \circ \mu$ is said to be the total variation of μ .

Proposition 11.100 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued pre-measure space, where $\mathcal{Y}$ is a normed linear space. Then, $\mathcal{P} \circ \mu$ defines a measure on the measurable space $(X, \mathcal{B})$.

Proof Clearly, $\mathcal{P} \circ \mu(\emptyset) = 0$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^{\infty} E_i$. $\forall i \in \mathbb{N}$ with $\mathcal{P} \circ \mu(E_i) > 0$, $\forall 0 \leq t_i < \mathcal{P} \circ \mu(E_i)$, $\exists n_i \in \mathbb{Z}_+$ and $\exists$ pairwise disjoint $(E_{i,j})_{j=1}^{n_i} \subseteq \mathcal{B}$ with $E_i = \bigcup_{j=1}^{n_i} E_{i,j}$ such that $t_i < \sum_{j=1}^{n_i} \|\mu(E_{i,j})\| \leq \mathcal{P} \circ \mu(E_i)$. $\forall i \in \mathbb{N}$ with $\mathcal{P} \circ \mu(E_i) = 0$, let $t_i = 0$, $n_i = 1$, $E_{i,1} = E_i$. Then, $0 = t_i \leq \|\mu(E_{i,1})\| \leq \mathcal{P} \circ \mu(E_i) = 0$.

Claim 11.100.1 $0 \leq t := \sum_{i=1}^{\infty} t_i \leq \mathcal{P} \circ \mu(E)$.

Proof of claim: Clearly, $t \geq 0$ since $t_i \geq 0, \forall i \in \mathbb{N}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $t < +\infty$; Case 2: $t = +\infty$.

Case 1: $t < +\infty$. Then, $t \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \|\mu(E_{i,j})\| \geq 0$. Therefore, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}$ such that $t - \epsilon < \sum_{i=1}^N \sum_{j=1}^{n_i} \|\mu(E_{i,j})\| \leq \sum_{i=1}^N \sum_{j=1}^{n_i} \|\mu(E_{i,j})\| + \left\| \mu(E \setminus \left(\bigcup_{i=1}^N \bigcup_{j=1}^{n_i} E_{i,j} \right)) \right\| \leq \mathcal{P} \circ \mu(E)$. By the arbitrariness of ϵ , we have $t \leq \mathcal{P} \circ \mu(E)$.

Case 2: $t = +\infty$. Then $t = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \|\mu(E_{i,j})\| = \infty, \forall M \in [0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}$ such that

$$\begin{aligned} M &\leq \sum_{i=1}^N \sum_{j=1}^{n_i} \|\mu(E_{i,j})\| \\ &\leq \sum_{i=1}^N \sum_{j=1}^{n_i} \|\mu(E_{i,j})\| + \left\| \mu(E \setminus \left(\bigcup_{i=1}^N \bigcup_{j=1}^{n_i} E_{i,j} \right)) \right\| \leq \mathcal{P} \circ \mu(E) \end{aligned}$$

This implies that $\mathcal{P} \circ \mu(E) = +\infty = t$. This completes the proof of the claim. $\square$

Hence, $\sum_{i=1}^{\infty} t_i \leq \mathcal{P} \circ \mu(E)$. By the arbitrariness of t_i 's, we have $\sum_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i) \leq \mathcal{P} \circ \mu(E)$.

On the other hand, $\forall n \in \mathbb{Z}_+, \forall$ pairwise disjoint $(A_j)_{j=1}^n \subseteq \mathcal{B}$ with $E = \bigcup_{j=1}^n A_j$, we have

$$\begin{aligned} \sum_{j=1}^n \|\mu(A_j)\| &= \sum_{j=1}^n \left\| \mu \left(\bigcup_{i=1}^{\infty} (E_i \cap A_j) \right) \right\| = \sum_{j=1}^n \left\| \sum_{i=1}^{\infty} \mu(E_i \cap A_j) \right\| \\ &\leq \sum_{j=1}^n \sum_{i=1}^{\infty} \|\mu(E_i \cap A_j)\| = \sum_{i=1}^{\infty} \sum_{j=1}^n \|\mu(E_i \cap A_j)\| \leq \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i) \end{aligned}$$

By the arbitrariness of n and $(A_j)_{j=1}^n$, we have $\mathcal{P} \circ \mu(E) \leq \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i)$.

Hence, $\mathcal{P} \circ \mu(E) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i)$. This implies that $\mathcal{P} \circ \mu$ is a measure on $(X, \mathcal{B})$. This completes the proof of the proposition. $\square$

Definition 11.101 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued pre-measure space, where $\mathcal{Y}$ is a normed linear space. We will say $\mathcal{X}$ is finite if $\mathcal{P} \circ \mu(X) < +\infty$. We will say a property holds almost everywhere in $\mathcal{X}$ if it holds almost everywhere in $\bar{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$. We will say that $\mathcal{X}$ is complete if the measure space $\bar{\mathcal{X}}$ is complete.

Fact 11.102 Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued pre-measure space, where $\mathcal{Y}$ is a normed linear space. Then, μ is finitely additive, i. e., $\mu(\bigcup_{i=1}^n E_i) = \sum_{i=1}^n \mu(E_i)$, where $n \in \mathbb{Z}_+$ and $(E_i)_{i=1}^n \subseteq \mathcal{B}$ is pairwise disjoint.

Proof Let $E_{n+i} = \emptyset, \forall i \in \mathbb{N}$. Clearly, $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$ is pairwise disjoint. Therefore, $\mu(\bigcup_{i=1}^n E_i) = \mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^n \mu(E_i)$. This completes the proof of the fact. $\square$

Proposition 11.103 *Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued pre-measure space, where $\mathcal{Y}$ is a normed linear space, and $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$ be such that $E_n \supseteq E_{n+1}$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} \mu(E_n) = \mu(\bigcap_{n=1}^\infty E_n) \in \mathcal{Y}$.*

Proof Let $E := \bigcap_{n=1}^\infty E_n \in \mathcal{B}$. Note that $E_1 = E \cup (\bigcup_{n=1}^\infty (E_n \setminus E_{n+1}))$, and the sets in the union are pairwise disjoint and measurable. Therefore, $\mu(E_1) = \mu(E) + \sum_{n=1}^\infty \mu(E_n \setminus E_{n+1})$ and $\|\mu(E)\| + \sum_{n=1}^\infty \|\mu(E_n \setminus E_{n+1})\| < +\infty$. Note that $E_n = E_{n+1} \cup (E_n \setminus E_{n+1})$, $n = 1, 2, \dots$, which implies that $\mu(E_n) = \mu(E_{n+1}) + \mu(E_n \setminus E_{n+1})$, by Fact 11.102. This leads to $\mu(E_1) = \mu(E) + \sum_{n=1}^\infty (\mu(E_n) - \mu(E_{n+1})) = \mu(E) + \mu(E_1) - \lim_{n \in \mathbb{N}} \mu(E_n) \in \mathcal{Y}$. Hence, we have $\lim_{n \in \mathbb{N}} \mu(E_n) = \mu(E) \in \mathcal{Y}$. This completes the proof of the proposition. $\square$

Proposition 11.104 *Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued pre-measure space, where $\mathcal{Y}$ is a normed linear space, and $(A_i)_{i=1}^\infty \subseteq \mathcal{B}$ with $A_i \subseteq A_{i+1}$, $\forall i \in \mathbb{N}$. Then, $\mu(\bigcup_{i=1}^\infty A_i) = \lim_{i \in \mathbb{N}} \mu(A_i) \in \mathcal{Y}$.*

Proof Let $A := \bigcup_{i=1}^\infty A_i \in \mathcal{B}$. Note that $A = A_1 \cup (\bigcup_{i=1}^\infty (A_{i+1} \setminus A_i))$, where the sets in the union are pairwise disjoint and measurable. Then, by countable additivity, $\mu(A) = \mu(A_1) + \sum_{i=1}^\infty \mu(A_{i+1} \setminus A_i) = \lim_{n \in \mathbb{N}} (\mu(A_1) + \sum_{i=1}^n \mu(A_{i+1} \setminus A_i)) = \lim_{n \in \mathbb{N}} \mu(A_{n+1}) = \lim_{i \in \mathbb{N}} \mu(A_i) \in \mathcal{Y}$, where the third equality follows from Fact 11.102. This completes the proof of the proposition. $\square$

Proposition 11.105 *Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued pre-measure space, where $\mathcal{Y}$ is a normed linear space. Then, $\forall E_1, E_2 \in \mathcal{B}$ with $\mathcal{P} \circ \mu(E_1 \triangle E_2) = 0$, we have $\mu(E_1) = \mu(E_2) \in \mathcal{Y}$.*

Proof By Proposition 11.100, $0 = \mathcal{P} \circ \mu(E_1 \triangle E_2) = \mathcal{P} \circ \mu(E_1 \setminus E_2) + \mathcal{P} \circ \mu(E_2 \setminus E_1)$. Then, $\mathcal{P} \circ \mu(E_1 \setminus E_2) = \mathcal{P} \circ \mu(E_2 \setminus E_1) = 0$ and $\mu(E_1 \setminus E_2) = \vartheta_{\mathcal{Y}} = \mu(E_2 \setminus E_1)$. Hence, we have $\mu(E_1) = \mu(E_1 \cap E_2) + \mu(E_1 \setminus E_2) = \mu(E_1 \cap E_2) + \mu(E_2 \setminus E_1) = \mu(E_2)$, where the first and last equality follows from Fact 11.102. This completes the proof of the proposition. $\square$

Proposition 11.106 *Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued pre-measure space, where $\mathcal{Y}$ is a normed linear space. Then, there is a unique complete $\mathcal{Y}$ -valued pre-measure space $(X, \mathcal{B}_0, \mu_0)$ such that*

(i) $\mathcal{B} \subseteq \mathcal{B}_0$;

(ii) $\forall E \in \mathcal{B}$, $\mu(E) = \mu_0(E)$ and $\mathcal{P} \circ \mu(E) = \mathcal{P} \circ \mu_0(E)$;

(iii) $\forall E \subseteq X$, $E \in \mathcal{B}_0$ if, and only if, $E = A \cup B$, where $A, B \subseteq X$ with $B \in \mathcal{B}$, and there exists a $C \in \mathcal{B}$, such that $A \subseteq C$ and $\mathcal{P} \circ \mu(C) = 0$.

Then, $(X, \mathcal{B}_0, \mu_0)$ is said to be the completion of $(X, \mathcal{B}, \mu)$. Then, $(X, \mathcal{B}_0, \mathcal{P} \circ \mu_0)$ is the completion of $(X, \mathcal{B}, \mathcal{P} \circ \mu)$.

Furthermore, let $(X, \mathcal{B}_1, \mu_1)$ be another complete $\mathcal{Y}$ -valued pre-measure space that satisfies (i) and (ii), then, $\mathcal{B}_0 \subseteq \mathcal{B}_1$, $\mu_0(E) = \mu_1(E)$, and $\mathcal{P} \circ \mu_0(E) = \mathcal{P} \circ \mu_1(E)$, $\forall E \in \mathcal{B}_0$.

Proof By Proposition 11.100, $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ is a measure space. By Proposition 11.12, it admits the completion $(X, \mathcal{B}_0, \nu_0)$, where $\mathcal{B}_0$ satisfies (i) and (iii) and $\nu_0(E) = \mathcal{P} \circ \mu(E)$, $\forall E \in \mathcal{B}$. Define $\mu_0 : \mathcal{B}_0 \rightarrow \mathcal{Y}$ as following. $\forall E \in \mathcal{B}_0$, there exists $A, B, C \subseteq X$ such that $E = A \cup B$, $B, C \in \mathcal{B}$, $A \subseteq C$, and $\mathcal{P} \circ \mu(C) = 0$. $\mu_0(E) := \mu(B) \in \mathcal{Y}$.

The mapping μ_0 is well defined because of the following. $\forall E \in \mathcal{B}_0$, let $A_1, A_2, B_1, B_2, C_1, C_2 \subseteq X$ be such that

$$\begin{aligned} E &= A_1 \cup B_1 = A_2 \cup B_2 \\ B_1, C_1, B_2, C_2 &\in \mathcal{B} \\ \mathcal{P} \circ \mu(C_1) &= \mathcal{P} \circ \mu(C_2) = 0; \quad A_1 \subseteq C_1; \quad A_2 \subseteq C_2 \end{aligned}$$

Then, $\mathcal{B} \ni B_1 \setminus B_2 \subseteq E \setminus B_2 = A_2 \setminus B_2 \subseteq C_2$ and $\mathcal{B} \ni B_2 \setminus B_1 \subseteq E \setminus B_1 = A_1 \setminus B_1 \subseteq C_1$. This implies, by Proposition 11.4, $0 \leq \mathcal{P} \circ \mu(B_1 \triangle B_2) = \mathcal{P} \circ \mu(B_1 \setminus B_2) + \mathcal{P} \circ \mu(B_2 \setminus B_1) \leq \mathcal{P} \circ \mu(C_2) + \mathcal{P} \circ \mu(C_1) = 0$. By Proposition 11.105, we have $\mu(B_1) = \mu(B_2)$. This shows that μ_0 is well defined.

$\forall E \in \mathcal{B} \subseteq \mathcal{B}_0$, $E = E \cup \emptyset$, then, $\mu_0(E) = \mu(E)$. This proves that μ_0 agrees with μ on $\mathcal{B}$.

Next, we will show that μ_0 is a $\mathcal{Y}$ -valued pre-measure on the measurable space $(X, \mathcal{B}_0)$. Since $\emptyset \in \mathcal{B}$, then $\mu_0(\emptyset) = \mu(\emptyset) = \vartheta_{\mathcal{Y}}$. Consider any sequence of pairwise disjoint subsets $(E_n)_{n=1}^{\infty} \subseteq \mathcal{B}_0$. $\forall n \in \mathbb{N}$, there exists $A_n, B_n, C_n \subseteq X$ such that $E_n = A_n \cup B_n$, $B_n, C_n \in \mathcal{B}$, $A_n \subseteq C_n$, and $\mathcal{P} \circ \mu(C_n) = 0$. Since $B_n \subseteq E_n$, $\forall n \in \mathbb{N}$, then, $(B_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ is pairwise disjoint. Hence, $\mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n)$ and $\sum_{n=1}^{\infty} \|\mu(B_n)\| < +\infty$. Since $\bigcup_{n=1}^{\infty} E_n = (\bigcup_{n=1}^{\infty} A_n) \cup (\bigcup_{n=1}^{\infty} B_n)$, $\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} C_n$, $0 \leq \mathcal{P} \circ \mu(\bigcup_{n=1}^{\infty} C_n) \leq \sum_{n=1}^{\infty} \mathcal{P} \circ \mu(C_n) = 0$, and $\bigcup_{n=1}^{\infty} B_n, \bigcup_{n=1}^{\infty} C_n \in \mathcal{B}$ then, $\mu_0(\bigcup_{n=1}^{\infty} E_n) = \mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} \mu_0(E_n)$. Also, $\sum_{n=1}^{\infty} \|\mu_0(E_n)\| = \sum_{n=1}^{\infty} \|\mu(B_n)\| < +\infty$. This proves that μ_0 is countably additive. Hence, $(X, \mathcal{B}_0, \mu_0)$ is a $\mathcal{Y}$ -valued pre-measure space.

Next, we will show that $\mathcal{P} \circ \mu_0 = \nu_0$. $\forall E \in \mathcal{B}_0$, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_i)_{i=1}^n \subseteq \mathcal{B}_0$ with $E = \bigcup_{i=1}^n E_i$, $\forall i \in \{1, \dots, n\}$, $E_i = A_i \cup B_i$ such that $A_i \subseteq C_i$, $B_i, C_i \in \mathcal{B}$, and $\mathcal{P} \circ \mu(C_i) = 0$. Then, $(B_i)_{i=1}^n \subseteq \mathcal{B}$ is pairwise disjoint, $E = (\bigcup_{i=1}^n A_i) \cup (\bigcup_{i=1}^n B_i) =: A \cup B$, $A \subseteq \bigcup_{i=1}^n C_i =: C \in \mathcal{B}$, $B \in \mathcal{B}$, and $0 \leq \mathcal{P} \circ \mu(C) \leq \sum_{i=1}^n \mathcal{P} \circ \mu(C_i) = 0$, where the last inequality follows from Proposition 11.6. By Proposition 11.12 and its proof, $\nu_0(E) = \mathcal{P} \circ \mu(B) = \sum_{i=1}^n \mathcal{P} \circ \mu(B_i) \geq \sum_{i=1}^n \|\mu(B_i)\| = \sum_{i=1}^n \|\mu_0(E_i)\|$. Hence, $\nu_0(E) \geq \mathcal{P} \circ \mu_0(E)$.

On the other hand, $\nu_0(E) = \mathcal{P} \circ \mu(B) \leq \mathcal{P} \circ \mu_0(B) \leq \mathcal{P} \circ \mu_0(E)$, where the first inequality follows from Definition 11.99 and the fact that $\mu = \mu_0|_{\mathcal{B}}$; and the second inequality follows from Proposition 11.4. Therefore, $\mathcal{P} \circ \mu_0 = \nu_0$. $\forall E \in \mathcal{B}$, $\mathcal{P} \circ \mu_0(E) = \nu_0(E) = \mathcal{P} \circ \mu(E)$. This shows that (ii) is satisfied. This proves that $(X, \mathcal{B}_0, \mu_0)$ is a complete $\mathcal{Y}$ -valued pre-measure space satisfying (i), (ii), and (iii) and $(X, \mathcal{B}_0, \mathcal{P} \circ \mu_0)$ is the completion of $(X, \mathcal{B}, \mathcal{P} \circ \mu)$.

Let $(X, \mathcal{B}_1, \mu_1)$ be any complete $\mathcal{Y}$ -valued pre-measure space that satisfies the (i) and (ii). By Proposition 11.100, $(X, \mathcal{B}_1, \mathcal{P} \circ \mu_1)$ is a complete measure space that extends $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. By Proposition 11.12 and the fact that $(X, \mathcal{B}_0, \mathcal{P} \circ \mu_0)$ is the completion of $(X, \mathcal{B}, \mathcal{P} \circ \mu)$, then $\mathcal{B}_0 \subseteq \mathcal{B}_1$ and $(\mathcal{P} \circ \mu_1)|_{\mathcal{B}_0} = \mathcal{P} \circ \mu_0$. $\forall E \in \mathcal{B}_0$, there exists $A, B, C \subseteq X$ such that $E = A \cup B$, $A \subseteq C$, $B, C \in \mathcal{B}$, and $\mathcal{P} \circ \mu(C) = 0$. Then, $\mathcal{P} \circ \mu_1(C) = 0$, $E = B \cup (E \setminus B)$, $\mathcal{B}_0 \ni E \setminus B = A \setminus B \subseteq C$, and $0 \leq \mathcal{P} \circ \mu_1(E \setminus B) \leq \mathcal{P} \circ \mu_1(C) = 0$. This leads to $\mu_1(E \setminus B) = \vartheta_{\mathcal{Y}}$ and $\mu_1(E) = \mu_1(B) + \mu_1(E \setminus B) = \mu(B) = \mu_0(E)$. Therefore, $\mu_0 = \mu_1|_{\mathcal{B}_0}$.

Therefore, $(X, \mathcal{B}_0, \mu_0)$ is a complete $\mathcal{Y}$ -valued pre-measure space that satisfies (i), (ii), and (iii). It is the unique such space since $\mathcal{B}_0$ is unique by (iii), and μ_0 is also unique since any complete $\mathcal{Y}$ -valued pre-measure space $(X, \mathcal{B}_0, \mu_1)$ satisfying (i) and (ii) will be such that $\mu_1 = \mu_0$ and $\mathcal{P} \circ \mu_0 = \mathcal{P} \circ \mu_1$.

This completes the proof of the proposition. $\square$

Proposition 11.107 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued pre-measure space, where $\mathcal{Y}$ is a normed linear space, $A \in \mathcal{B}$, $\hat{\mathcal{B}} \subseteq \mathcal{B}$ be a σ -algebra on X , and $\mathcal{B}_A := \{C \subseteq A \mid C \in \mathcal{B}\}$. Then, the following statements hold.*

(i) $(X, \hat{\mathcal{B}}, \mu|_{\hat{\mathcal{B}}})$ and $\mathcal{A} := (A, \mathcal{B}_A, \mu_A := \mu|_{\mathcal{B}_A})$ are $\mathcal{Y}$ -valued pre-measure spaces, and $(\mathcal{P} \circ \mu)|_{\mathcal{B}_A} = \mathcal{P} \circ (\mu|_{\mathcal{B}_A}) = \mathcal{P} \circ \mu_A$. The $\mathcal{Y}$ -valued pre-measure space $\mathcal{A}$ is said to be the subspace of the $\mathcal{Y}$ -valued pre-measure space $\mathcal{X}$.

(ii) If, in addition, $\mathcal{X}$ is complete, then $\mathcal{A}$ is also complete.

(iii) Let $\bar{\mathcal{X}} := (X, \bar{\mathcal{B}}, \bar{\mu})$ be the completion of $\mathcal{X}$ as defined in Proposition 11.106, and $\bar{\mathcal{A}} := (A, \bar{\mathcal{B}}_A, \bar{\mu}_A)$ be the $\mathcal{Y}$ -valued pre-measure subspace of $\bar{\mathcal{X}}$. Then, $\bar{\mathcal{A}}$ is the completion of $\mathcal{A}$.

Proof (i) Since $\hat{\mathcal{B}}$ is a σ -algebra on X , then $(X, \hat{\mathcal{B}})$ is a measurable space. $\forall B \in \hat{\mathcal{B}} \subseteq \mathcal{B}$, $\mu|_{\hat{\mathcal{B}}}(B) = \mu(B) \in \mathcal{Y}$. Then, $\mu|_{\hat{\mathcal{B}}} : \hat{\mathcal{B}} \rightarrow \mathcal{Y}$. Clearly, $\mu|_{\hat{\mathcal{B}}}(\emptyset) = \mu(\emptyset) = \vartheta_{\mathcal{Y}}$. $\forall$ pairwise disjoint sequence $(E_i)_{i=1}^{\infty} \subseteq \hat{\mathcal{B}} \subseteq \mathcal{B}$, we have $E := \bigcup_{i=1}^{\infty} E_i \in \hat{\mathcal{B}}$, $\sum_{i=1}^{\infty} \|\mu|_{\hat{\mathcal{B}}}(E_i)\| = \sum_{i=1}^{\infty} \|\mu(E_i)\| < +\infty$, and $\mu|_{\hat{\mathcal{B}}}(E) = \mu(E) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \mu|_{\hat{\mathcal{B}}}(E_i)$. Hence, $(X, \hat{\mathcal{B}}, \mu|_{\hat{\mathcal{B}}})$ is a $\mathcal{Y}$ -valued pre-measure space.

Clearly, $\mathcal{B}_A$ is a σ -algebra on A as we have shown in Proposition 11.13. $\emptyset \in \mathcal{B}_A$ and $\mu|_{\mathcal{B}_A}(\emptyset) = \mu(\emptyset) = \vartheta_{\mathcal{Y}}$. $\forall$ pairwise disjoint sequence $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}_A \subseteq \mathcal{B}$, we have $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}_A$, $\sum_{i=1}^{\infty} \|\mu|_{\mathcal{B}_A}(E_i)\| = \sum_{i=1}^{\infty} \|\mu(E_i)\| < +\infty$, and $\mu|_{\mathcal{B}_A}(E) = \mu(E) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \mu|_{\mathcal{B}_A}(E_i)$. Hence $\mathcal{A} = (A, \mathcal{B}_A, \mu|_{\mathcal{B}_A})$ is a $\mathcal{Y}$ -valued pre-measure space. By Definition 11.99, $(\mathcal{P} \circ \mu)|_{\mathcal{B}_A} = \mathcal{P} \circ (\mu|_{\mathcal{B}_A}) = \mathcal{P} \circ \mu_A$.

(ii) If, in addition, $\mathcal{X}$ is complete, then $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ is a complete measure space. $\forall E_A \subseteq A$ such that there exists $B \in \mathcal{B}_A$ with $E_A \subseteq B$ and $\mathcal{P} \circ (\mu|_{\mathcal{B}_A})(B) = 0$, we have $B \in \mathcal{B}$ and $\mathcal{P} \circ \mu(B) = (\mathcal{P} \circ \mu)|_{\mathcal{B}_A}(B) =$

$\mathcal{P} \circ (\mu|_{\mathcal{B}_A})(B) = \mathcal{P} \circ \mu_A(B) = 0$. Then, $E_A \in \mathcal{B}$ by the completeness of $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. Thus, $E_A \in \mathcal{B}_A$. By the arbitrariness of E_A , $(A, \mathcal{B}_A, \mathcal{P} \circ \mu_A)$ is complete. Hence, $\mathcal{A}$ is complete.

(iii) Note that $\bar{\mathcal{B}}_A = \{C \subseteq A \mid C \in \bar{\mathcal{B}}\}$, $\bar{\mu}_A = \bar{\mu}|_{\bar{\mathcal{B}}_A}$, and $\mathcal{P} \circ \bar{\mu}_A = (\mathcal{P} \circ \bar{\mu})|_{\bar{\mathcal{B}}_A}$. By (ii), $\bar{\mathcal{A}}$ is a complete $\mathcal{Y}$ -valued pre-measure space. Since $\mathcal{B} \subseteq \bar{\mathcal{B}}$, then $\mathcal{B}_A \subseteq \bar{\mathcal{B}}_A$. $\forall E \in \mathcal{B}_A$, $\mu_A(E) = \mu|_{\mathcal{B}_A}(E) = \mu(E) = \bar{\mu}(E) = \bar{\mu}_A(E)$ and, by (i), $\mathcal{P} \circ \mu_A(E) = (\mathcal{P} \circ \mu)|_{\mathcal{B}_A}(E) = \mathcal{P} \circ \mu(E) = \mathcal{P} \circ \bar{\mu}(E) = (\mathcal{P} \circ \bar{\mu})|_{\bar{\mathcal{B}}_A}(E) = \mathcal{P} \circ \bar{\mu}_A(E)$, where the third equality follows from Proposition 11.106. $\forall E \subseteq A$, assume that $E = E_A \cup E_B$ with $E_B, E_C \in \mathcal{B}_A$, $E_A \subseteq E_C$, and $\mathcal{P} \circ \mu_A(E_C) = 0$. Then, $\mathcal{P} \circ \bar{\mu}_A(E_C) = 0$ and $E_A \in \bar{\mathcal{B}}_A$ since $\bar{\mathcal{A}}$ is complete. Then, $E \in \bar{\mathcal{B}}_A$. $\forall E \subseteq A$, assume that $E \in \bar{\mathcal{B}}_A$. Then, $E \subseteq A$ and $E \in \bar{\mathcal{B}}$. By Proposition 11.106, $E = E_A \cup E_B$ with $E_B, E_C \in \mathcal{B}$, $E_A \subseteq E_C$, and $\mathcal{P} \circ \mu(E_C) = 0$. Then, $E_B, E_C \cap A \in \mathcal{B}_A$, $E_A \subseteq E_C \cap A$, and $0 \leq \mathcal{P} \circ \mu_A(E_C \cap A) = \mathcal{P} \circ \mu(E_C \cap A) \leq \mathcal{P} \circ \mu(E_C) = 0$, where the last inequality follows from Proposition 11.4. Hence, (i) — (iii) of Proposition 11.106 are satisfied and $\bar{\mathcal{A}}$ is the completion of $\mathcal{A}$.

This completes the proof of the proposition. $\square$

Definition 11.108 Let $(X, \mathcal{B}, \nu)$ be a measure space and $\mathcal{Y}$ be a normed linear space. A $\mathcal{Y}$ -valued measure μ on $(X, \mathcal{B})$ is a function from $\mathcal{B}$ to $\mathcal{Y}$ that satisfies

- (i) $\mu(\emptyset) = \vartheta_{\mathcal{Y}}$;
- (ii) $\forall E \in \mathcal{B}$ with $\nu(E) = +\infty$, $\mu(E)$ is undefined;
- (iii) $\forall E \in \mathcal{B}$ with $\nu(E) < +\infty$, $\mu(E) \in \mathcal{Y}$ and $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^{\infty} E_i$, we have $\sum_{i=1}^{\infty} \|\mu(E_i)\| < +\infty$ and $\mu(E) = \sum_{i=1}^{\infty} \mu(E_i)$; (countable additivity)
- (iv) $\forall E \in \mathcal{B}$ with $\nu(E) < +\infty$, we have

$$\nu(E) = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(E_i)\|$$

Then, $(X, \mathcal{B}, \mu)$ is said to be a $\mathcal{Y}$ -valued measure space; and ν is said to be the total variation of μ , denoted by $\mathcal{P} \circ \mu$.

Definition 11.109 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space. We will say $\mathcal{X}$ is finite if $\mathcal{P} \circ \mu(X) < +\infty$. We will say $\mathcal{X}$ is σ -finite if $\bar{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$ is σ -finite. We will say a property holds almost everywhere in $\mathcal{X}$ if it holds almost everywhere in $\bar{\mathcal{X}}$. We will say that $\mathcal{X}$ is complete if the measure space $\bar{\mathcal{X}}$ is complete. We will say that a sequence of measurable functions on $\mathcal{X}$ converges in measure in $\mathcal{X}$ if it converges in measure in $\bar{\mathcal{X}}$.

Clearly, $(X, \mathcal{B}, \mu)$ is a finite $\mathcal{Y}$ -valued measure space if, and only if, it is a finite $\mathcal{Y}$ -valued pre-measure space. In this case, the definitions of total variations of μ coincide considering μ as a $\mathcal{Y}$ -valued measure or as $\mathcal{Y}$ -valued pre-measure. Hence, a $\mathcal{Y}$ -valued pre-measure space becomes a finite $\mathcal{Y}$ -valued measure space when we show that its total variation is finite.

Fact 11.110 *Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space. Then, μ is finitely additive, i. e., $\forall n \in \mathbb{Z}_+$ and $\forall$ pairwise disjoint $(E_i)_{i=1}^n \subseteq \mathcal{B}$, $\mu(\bigcup_{i=1}^n E_i) = \sum_{i=1}^n \mu(E_i)$, whenever the right-hand-side or the left-hand-side makes sense.*

Proof Let $E := \bigcup_{i=1}^n E_i \in \mathcal{B}$. Clearly, $\mathcal{P} \circ \mu(E) = \sum_{i=1}^n \mathcal{P} \circ \mu(E_i)$. Then, $E \in \text{dom}(\mu)$ whenever $(E_i)_{i=1}^n \subseteq \text{dom}(\mu)$. Let $E_{n+i} = \emptyset$, $\forall i \in \mathbb{N}$. Clearly, $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$ is pairwise disjoint. Therefore, $\mu(E) = \mu(\bigcup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \mu(E_i) = \sum_{i=1}^n \mu(E_i) \in \mathcal{Y}$, when $E \in \text{dom}(\mu)$. This completes the proof of the fact. $\square$

Proposition 11.111 *Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space, and $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$ be such that $\mathcal{P} \circ \mu(E_1) < +\infty$ and $E_n \supseteq E_{n+1}$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} \mu(E_n) = \mu(\bigcap_{n=1}^\infty E_n) \in \mathcal{Y}$.*

Proof Clearly, $E_1 \in \text{dom}(\mu)$ and then $E_n \in \text{dom}(\mu)$, $\forall n \in \mathbb{N}$. Let $E := \bigcap_{n=1}^\infty E_n \in \mathcal{B}$. Note that $E_1 = E \cup (\bigcup_{n=1}^\infty (E_n \setminus E_{n+1}))$, and the sets in the union are pairwise disjoint and measurable. Therefore, $\mu(E_1) = \mu(E) + \sum_{n=1}^\infty \mu(E_n \setminus E_{n+1}) \in \mathcal{Y}$ and $\|\mu(E)\| + \sum_{n=1}^\infty \|\mu(E_n \setminus E_{n+1})\| < \infty$. Note that $E_n = E_{n+1} \cup (E_n \setminus E_{n+1})$, $n = 1, 2, \dots$, which implies that $\mu(E_n) = \mu(E_{n+1}) + \mu(E_n \setminus E_{n+1}) \in \mathcal{Y}$, by Fact 11.110. This leads to $\mu(E_1) = \mu(E) + \sum_{n=1}^\infty (\mu(E_n) - \mu(E_{n+1})) = \mu(E) + \mu(E_1) - \lim_{n \in \mathbb{N}} \mu(E_n) \in \mathcal{Y}$. Hence, we have $\lim_{n \in \mathbb{N}} \mu(E_n) = \mu(E) \in \mathcal{Y}$. This completes the proof of the proposition. $\square$

Proposition 11.112 *Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space, and $(A_i)_{i=1}^\infty \subseteq \mathcal{B}$ be such that $A_i \subseteq A_{i+1}$, $\forall i \in \mathbb{N}$, and $A := \bigcup_{i=1}^\infty A_i \in \text{dom}(\mu)$. Then, $\mu(A) = \lim_{i \in \mathbb{N}} \mu(A_i) \in \mathcal{Y}$.*

Proof $A \in \text{dom}(\mu)$ implies that $\mathcal{P} \circ \mu(A) < +\infty$. Then, $\mathcal{P} \circ \mu(A_i) \leq \mathcal{P} \circ \mu(A) < +\infty$ and $A_i \in \text{dom}(\mu)$, $\forall i \in \mathbb{N}$. Note that $A = A_1 \cup (\bigcup_{i=1}^\infty (A_{i+1} \setminus A_i))$, where the sets in the union are pairwise disjoint and measurable. Then, by countable additivity, $\mu(A) = \mu(A_1) + \sum_{i=1}^\infty \mu(A_{i+1} \setminus A_i) = \lim_{n \in \mathbb{N}} (\mu(A_1) + \sum_{i=1}^n \mu(A_{i+1} \setminus A_i)) = \lim_{n \in \mathbb{N}} \mu(A_{n+1}) = \lim_{i \in \mathbb{N}} \mu(A_i) \in \mathcal{Y}$, where the third equality follows from Fact 11.110. This completes the proof of the proposition. $\square$

Proposition 11.113 *Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space. Then, $\forall E_1, E_2 \in \mathcal{B}$ with $\mathcal{P} \circ \mu(E_1 \triangle E_2) = 0$, we have $\mu(E_1) = \mu(E_2)$ whenever one of them makes sense.*

Proof By Fact 11.3, $0 = \mathcal{P} \circ \mu(E_1 \triangle E_2) = \mathcal{P} \circ \mu(E_1 \setminus E_2) + \mathcal{P} \circ \mu(E_2 \setminus E_1)$. Then, $\mathcal{P} \circ \mu(E_1 \setminus E_2) = \mathcal{P} \circ \mu(E_2 \setminus E_1) = 0$ and $\mu(E_1 \setminus E_2) = \vartheta_{\mathcal{Y}} = \mu(E_2 \setminus E_1)$. Without loss of generality, assume $E_1 \in \text{dom}(\mu)$. Then, $\mathcal{P} \circ \mu(E_1) < +\infty$. This implies that $\mathcal{P} \circ \mu(E_2) = \mathcal{P} \circ \mu(E_1 \cap E_2) + \mathcal{P} \circ \mu(E_2 \setminus E_1) = \mathcal{P} \circ \mu(E_1 \cap E_2) + \mathcal{P} \circ \mu(E_1 \setminus E_2) = \mathcal{P} \circ \mu(E_1) < +\infty$ and $E_2 \in \text{dom}(\mu)$. By Fact 11.110, we have $\mu(E_1) = \mu(E_1 \cap E_2) + \mu(E_1 \setminus E_2) = \mu(E_1 \cap E_2) + \mu(E_2 \setminus E_1) = \mu(E_2)$. This completes the proof of the proposition. $\square$

Proposition 11.114 *Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space. Then, there is a unique complete $\mathcal{Y}$ -valued measure space $(X, \mathcal{B}_0, \mu_0)$ such that*

$$(i) \quad \mathcal{B} \subseteq \mathcal{B}_0;$$

$$(ii) \quad (\mathcal{P} \circ \mu_0)|_{\mathcal{B}} = \mathcal{P} \circ \mu \text{ and } \mu_0|_{\mathcal{B}} = \mu;$$

$$(iii) \quad \forall E \subseteq X, E \in \mathcal{B}_0 \text{ if, and only if, } E = A \cup B, \text{ where } A, B \subseteq X \text{ with } B \in \mathcal{B}, \text{ and there exists a } C \in \mathcal{B}, \text{ such that } A \subseteq C \text{ and } \mathcal{P} \circ \mu(C) = 0.$$

Then, $(X, \mathcal{B}_0, \mu_0)$ is said to be the completion of $(X, \mathcal{B}, \mu)$. Then, $(X, \mathcal{B}_0, \mathcal{P} \circ \mu_0)$ is the completion of $(X, \mathcal{B}, \mathcal{P} \circ \mu)$.

Furthermore, let $(X, \mathcal{B}_1, \mu_1)$ be another complete $\mathcal{Y}$ -valued measure space that satisfies (i) and (ii), then, $\mathcal{B}_0 \subseteq \mathcal{B}_1$, $\mathcal{P} \circ \mu_0 = (\mathcal{P} \circ \mu_1)|_{\mathcal{B}_0}$, and $\mu_0 = \mu_1|_{\mathcal{B}_0}$.

Proof By Definition 11.108, $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ is a measure space. By Proposition 11.12, it admits the completion $(X, \mathcal{B}_0, \nu_0)$, where $\mathcal{B}_0$ satisfies (i) and (iii) and $\nu_0(E) = \mathcal{P} \circ \mu(E)$, $\forall E \in \mathcal{B}$. Define a function μ_0 from $\mathcal{B}_0$ to $\mathcal{Y}$ as following. $\forall E \in \mathcal{B}_0$ with $\nu_0(E) = +\infty$, $\mu_0(E)$ is undefined. $\forall E \in \mathcal{B}_0$ with $\nu_0(E) < +\infty$, there exists $A, B, C \subseteq X$ such that $E = A \cup B$, $B, C \in \mathcal{B}$, $A \subseteq C$, and $\mathcal{P} \circ \mu(C) = 0$. By Proposition 11.12 and its proof, $\nu_0(E) = \mathcal{P} \circ \mu(B) < +\infty$ and $B \in \text{dom}(\mu)$. We will set $\mu_0(E) := \mu(B) \in \mathcal{Y}$.

The mapping μ_0 is well defined because of the following. $\forall E \in \mathcal{B}_0$ with $\nu_0(E) < +\infty$, let $A_1, A_2, B_1, B_2, C_1, C_2 \subseteq X$ be such that

$$\begin{aligned} E &= A_1 \cup B_1 = A_2 \cup B_2 \\ B_1, C_1, B_2, C_2 &\in \mathcal{B} \\ \mathcal{P} \circ \mu(C_1) &= \mathcal{P} \circ \mu(C_2) = 0; \quad A_1 \subseteq C_1; \quad A_2 \subseteq C_2 \end{aligned}$$

Then, $\mathcal{B} \ni B_1 \setminus B_2 \subseteq E \setminus B_2 = A_2 \setminus B_2 \subseteq C_2$ and $\mathcal{B} \ni B_2 \setminus B_1 \subseteq E \setminus B_1 = A_1 \setminus B_1 \subseteq C_1$. This implies, by Proposition 11.4, $0 \leq \mathcal{P} \circ \mu(B_1 \triangle B_2) = \mathcal{P} \circ \mu(B_1 \setminus B_2) + \mathcal{P} \circ \mu(B_2 \setminus B_1) \leq \mathcal{P} \circ \mu(C_2) + \mathcal{P} \circ \mu(C_1) = 0$. By Proposition 11.113, we have $\mu(B_1) = \mu(B_2) \in \mathcal{Y}$. This shows that μ_0 is well defined.

$\forall E \in \mathcal{B} \subseteq \mathcal{B}_0$, $E = E \cup \emptyset$, then, $\nu_0(E) = \mathcal{P} \circ \mu(E)$. If $\mathcal{P} \circ \mu(E) = +\infty$, then $\mu(E)$ is undefined and $\mu_0(E)$ is also undefined. If $\mathcal{P} \circ \mu(E) < +\infty$, then $\mu_0(E) = \mu(E) \in \mathcal{Y}$. This proves that $\mu_0|_{\mathcal{B}} = \mu$.

Next, we will show that $(X, \mathcal{B}_0, \mu_0)$ is a $\mathcal{Y}$ -valued measure space with $\mathcal{P} \circ \mu_0 = \nu_0$. Since $\emptyset \in \mathcal{B}$, then $\mu_0(\emptyset) = \mu(\emptyset) = \vartheta_{\mathcal{Y}}$. Hence, (i) of Definition 11.108 is satisfied. (ii) of Definition 11.108 is satisfied by the definition of μ_0 . Consider any $E \in \mathcal{B}_0$ with $\nu_0(E) < +\infty$, then $\mu_0(E) \in \mathcal{Y}$. $\forall$ pairwise disjoint $(E_n)_{n=1}^\infty \subseteq \mathcal{B}_0$ with $E = \bigcup_{n=1}^\infty E_n$, $\forall n \in \mathbb{N}$, there exists $A_n, B_n, C_n \subseteq X$ such that $E_n = A_n \cup B_n$, $B_n, C_n \in \mathcal{B}$, $A_n \subseteq C_n$, and $\mathcal{P} \circ \mu(C_n) = 0$. Since $B_n \subseteq E_n$, $\forall n \in \mathbb{N}$, then, $(B_n)_{n=1}^\infty \subseteq \mathcal{B}$ is pairwise disjoint. Let $A := \bigcup_{n=1}^\infty A_n$, $B := \bigcup_{n=1}^\infty B_n \in \mathcal{B}$, and $C := \bigcup_{n=1}^\infty C_n \in \mathcal{B}$. Then, $E = A \cup B$, $A \subseteq C$, and $\mathcal{P} \circ \mu(C) = 0$. This implies that $\mathcal{P} \circ \mu(B) = \nu_0(E) < +\infty$, $\mu(B) = \sum_{n=1}^\infty \mu(B_n) \in \mathcal{Y}$, and $\sum_{n=1}^\infty \|\mu(B_n)\| < +\infty$. Hence, $\mu_0(E) = \mu(B) = \sum_{n=1}^\infty \mu(B_n) = \sum_{n=1}^\infty \mu_0(E_n) \in \mathcal{Y}$. Also, $\sum_{n=1}^\infty \|\mu_0(E_n)\| = \sum_{n=1}^\infty \|\mu(B_n)\| < +\infty$. This proves that (iii) of Definition 11.108 is satisfied and μ_0 is countably additive.

$\forall E \in \mathcal{B}_0$ with $\nu_0(E) < +\infty$, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_i)_{i=1}^n \subseteq \mathcal{B}_0$ with $E = \bigcup_{i=1}^n E_i$, $\forall i \in \{1, \dots, n\}$, $\exists A_i, B_i, C_i \subseteq X$ such that $E_i = A_i \cup B_i$, $B_i, C_i \in \mathcal{B}$, $A_i \subseteq C_i$, and $\mathcal{P} \circ \mu(C_i) = 0$. Then, $(B_i)_{i=1}^n \subseteq \mathcal{B}$ is pairwise disjoint. Let $A := \bigcup_{i=1}^n A_i$, $B := \bigcup_{i=1}^n B_i \in \mathcal{B}$, and $C := \bigcup_{i=1}^n C_i \in \mathcal{B}$, we have $E = A \cup B$, $A \subseteq C$, and $\mathcal{P} \circ \mu(C) = 0$. By Proposition 11.12 and its proof, $+\infty > \nu_0(E) = \mathcal{P} \circ \mu(B) = \sum_{i=1}^n \mathcal{P} \circ \mu(B_i) \geq \sum_{i=1}^n \|\mu(B_i)\| = \sum_{i=1}^n \|\mu_0(E_i)\|$. Hence,

$$\nu_0(E) \geq \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}_0, \bigcup_{i=1}^n E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu_0(E_i)\| =: s_E$$

On the other hand,

$$\begin{aligned} +\infty &> \nu_0(E) = \mathcal{P} \circ \mu(B) \\ &= \sup_{n \in \mathbb{Z}_+, (B_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n B_i = B, B_i \cap B_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(B_i)\| \\ &\leq \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}_0, \bigcup_{i=1}^n E_i = B, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu_0(E_i)\| \leq s_E \end{aligned}$$

Hence, we have $\nu_0(E) = s_E$ and (iv) of Definition 11.108 is satisfied. Hence, $(X, \mathcal{B}_0, \mu_0)$ is a $\mathcal{Y}$ -valued measure space with total variation $\mathcal{P} \circ \mu_0 = \nu_0$. Since ν_0 is complete, then $(X, \mathcal{B}_0, \mu_0)$ is a complete $\mathcal{Y}$ -valued measure space. Since $\nu_0|_{\mathcal{B}} = \mathcal{P} \circ \mu$, then $(\mathcal{P} \circ \mu_0)|_{\mathcal{B}} = \mathcal{P} \circ \mu$. Hence, (i) – (iii) are satisfied by $(X, \mathcal{B}_0, \mu_0)$. Clearly, $(X, \mathcal{B}_0, \mathcal{P} \circ \mu_0)$ is the completion of $(X, \mathcal{B}, \mathcal{P} \circ \mu)$.

Let $(X, \mathcal{B}_1, \mu_1)$ be any complete $\mathcal{Y}$ -valued measure space that satisfies (i) and (ii). $(X, \mathcal{B}_1, \mathcal{P} \circ \mu_1)$ is a complete measure space that extends $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. By Proposition 11.12 and the fact that $(X, \mathcal{B}_0, \mathcal{P} \circ \mu_0)$ is the completion of $(X, \mathcal{B}, \mathcal{P} \circ \mu)$, then $\mathcal{B}_0 \subseteq \mathcal{B}_1$ and $(\mathcal{P} \circ \mu_1)|_{\mathcal{B}_0} = \mathcal{P} \circ \mu_0$. $\forall E \in \mathcal{B}_0$ with $\mathcal{P} \circ \mu_0(E) = +\infty$, we have $\mathcal{P} \circ \mu_1(E) = +\infty$ and $\mu_0(E)$ and $\mu_1(E)$ are both undefined. $\forall E \in \mathcal{B}_0$ with $\mathcal{P} \circ \mu_0(E) < +\infty$, by (iii), there exists $A, B, C \subseteq X$ such that $E = A \cup B$, $A \subseteq C$, $B, C \in \mathcal{B}$, and $\mathcal{P} \circ \mu(C) = 0$.

Then, $\mu_0(E) = \mu(B) = \mu_1(B) \in \mathcal{Y}$ and $\mathcal{P} \circ \mu_1(E) = \mathcal{P} \circ \mu_0(E) < +\infty$. Note that $\mathcal{P} \circ \mu_1(C) = \mathcal{P} \circ \mu(C) = 0$, $E = B \cup (E \setminus B)$, $\mathcal{B}_0 \ni E \setminus B = A \setminus B \subseteq C$, and $0 \leq \mathcal{P} \circ \mu_1(E \setminus B) \leq \mathcal{P} \circ \mu_1(C) = 0$. This leads to $\mu_1(E \setminus B) = \vartheta_{\mathcal{Y}}$ and, by Fact 11.110, $\mu_1(E) = \mu_1(B) + \mu_1(E \setminus B) = \mu(B) = \mu_0(E) \in \mathcal{Y}$. Therefore, $\mu_0 = \mu_1|_{\mathcal{B}_0}$.

Therefore, $(X, \mathcal{B}_0, \mu_0)$ is a complete $\mathcal{Y}$ -valued measure space that satisfies (i), (ii), and (iii). It is the unique such space since $\mathcal{B}_0$ is unique by (iii), and μ_0 is also unique since any complete $\mathcal{Y}$ -valued measure space $(X, \mathcal{B}_0, \mu_1)$ satisfying (i) and (ii) will be such that $\mu_1 = \mu_0$ and $\mathcal{P} \circ \mu_0 = \mathcal{P} \circ \mu_1$.

This completes the proof of the proposition. $\square$

Proposition 11.115 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space, $A \in \mathcal{B}$, and $\mathcal{B}_A := \{C \subseteq A \mid C \in \mathcal{B}\}$. Then, the following statements hold.*

(i) $\mathcal{A} := (A, \mathcal{B}_A, \mu_A := \mu|_{\mathcal{B}_A})$ is a $\mathcal{Y}$ -valued measure space, which is said to be the subspace of $\mathcal{X}$, and $(\mathcal{P} \circ \mu)|_{\mathcal{B}_A} = \mathcal{P} \circ \mu_A$.

(ii) If, in addition, $\mathcal{X}$ is complete, then $\mathcal{A}$ is also complete.

(iii) Let $\bar{\mathcal{X}} := (X, \bar{\mathcal{B}}, \bar{\mu})$ be the completion of $\mathcal{X}$ as defined in Proposition 11.114, and $\bar{\mathcal{A}} := (A, \bar{\mathcal{B}}_A, \bar{\mu}_A)$ be the $\mathcal{Y}$ -valued measure subspace of $\bar{\mathcal{X}}$. Then, $\bar{\mathcal{A}}$ is the completion of $\mathcal{A}$.

Proof (i) Clearly, $\mathcal{B}_A$ is a σ -algebra on A as we have shown in Proposition 11.13. By Proposition 11.13, $\nu_A := (\mathcal{P} \circ \mu)|_{\mathcal{B}_A} : \mathcal{B}_A \rightarrow [0, +\infty] \subset \mathbb{R}_e$ is a measure on $(A, \mathcal{B}_A)$. Clearly, $\emptyset \in \mathcal{B}_A$ and $\mu_A(\emptyset) = \mu|_{\mathcal{B}_A}(\emptyset) = \mu(\emptyset) = \vartheta_{\mathcal{Y}}$. $\forall E \in \mathcal{B}_A$ with $\nu_A(E) = +\infty$, we have $\mathcal{P} \circ \mu(E) = +\infty$ and $\mu(E)$ is undefined. Then, $\mu_A(E)$ is also undefined. $\forall E \in \mathcal{B}_A$ with $\nu_A(E) < +\infty$, we have $\mathcal{P} \circ \mu(E) = \nu_A(E) < +\infty$. Then, $\mu_A(E) = \mu(E) \in \mathcal{Y}$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}_A \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^{\infty} E_i$, we have $\sum_{i=1}^{\infty} \|\mu_A(E_i)\| = \sum_{i=1}^{\infty} \|\mu(E_i)\| < +\infty$, and $\mu_A(E) = \mu(E) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \mu_A(E_i) \in \mathcal{Y}$. $\forall E \in \mathcal{B}_A$ with $\nu_A(E) < +\infty$, we have $\mathcal{P} \circ \mu(E) = \nu_A(E) < +\infty$. Then, $\nu_A(E) = \mathcal{P} \circ \mu(E) = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(E_i)\| = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}_A, \bigcup_{i=1}^n E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu_A(E_i)\|$. Hence, $\mathcal{A} = (A, \mathcal{B}_A, \mu_A)$ is a $\mathcal{Y}$ -valued measure space with $\mathcal{P} \circ \mu_A = \nu_A = (\mathcal{P} \circ \mu)|_{\mathcal{B}_A}$.

(ii) If, in addition, $\mathcal{X}$ is complete, then $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ is a complete measure space. Since $(A, \mathcal{B}_A, \mathcal{P} \circ \mu_A)$ is a subspace of $(X, \mathcal{B}, \mathcal{P} \circ \mu)$, then it is complete by Proposition 11.13. Hence, $\mathcal{A}$ is complete.

(iii) Note that $\bar{\mathcal{B}}_A = \{C \subseteq A \mid C \in \bar{\mathcal{B}}\}$, $\bar{\mu}_A = \bar{\mu}|_{\bar{\mathcal{B}}_A}$, and $\mathcal{P} \circ \bar{\mu}_A = (\mathcal{P} \circ \bar{\mu})|_{\bar{\mathcal{B}}_A}$. By (ii), $\bar{\mathcal{A}}$ is a complete $\mathcal{Y}$ -valued measure space. Since $\mathcal{B} \subseteq \bar{\mathcal{B}}$, then $\mathcal{B}_A \subseteq \bar{\mathcal{B}}_A$. Note that $\mathcal{P} \circ \mu_A = (\mathcal{P} \circ \mu)|_{\mathcal{B}_A} = ((\mathcal{P} \circ \bar{\mu})|_{\mathcal{B}})|_{\mathcal{B}_A} = (\mathcal{P} \circ \bar{\mu})|_{\mathcal{B}_A} = ((\mathcal{P} \circ \bar{\mu})|_{\bar{\mathcal{B}}_A})|_{\mathcal{B}_A} = (\mathcal{P} \circ \bar{\mu}_A)|_{\mathcal{B}_A}$, where the

first equality follows from (i); the second equality follows from Proposition 11.114; the third and fourth equalities follow from the fact that $\mathcal{B}_A \subseteq \mathcal{B}$ and $\mathcal{B}_A \subseteq \bar{\mathcal{B}}_A$; and the last equality follows from (i). Note also that $\mu_A = \mu|_{\mathcal{B}_A} = (\bar{\mu}|_{\mathcal{B}})|_{\mathcal{B}_A} = (\bar{\mu}|_{\bar{\mathcal{B}}_A})|_{\mathcal{B}_A} = \bar{\mu}_A|_{\mathcal{B}_A}$, where the first equality follows from (i), the second equality follows from Proposition 11.114; the third equality follows from the fact that $\mathcal{B}_A \subseteq \mathcal{B}$ and $\mathcal{B}_A \subseteq \bar{\mathcal{B}}_A$; and the last equality follows from (i). $\forall E \subseteq A$, assume that $E = E_A \cup E_B$ with $E_B, E_C \in \mathcal{B}_A$, $E_A \subseteq E_C$, and $\mathcal{P} \circ \mu_A(E_C) = 0$. Then, $\mathcal{P} \circ \bar{\mu}_A(E_C) = 0$ and $E_A \in \bar{\mathcal{B}}_A$ since $\bar{\mathcal{A}}$ is complete. Then, $E \in \bar{\mathcal{B}}_A$. On the other hand, $\forall E \subseteq A$, assume that $E \in \bar{\mathcal{B}}_A$. Then, $E \subseteq A$ and $E \in \bar{\mathcal{B}}$. By Proposition 11.114, $E = E_A \cup E_B$ with $E_B, E_C \in \mathcal{B}$, $E_A \subseteq E_C$, and $\mathcal{P} \circ \mu(E_C) = 0$. Then, $E_B, E_C \cap A \in \mathcal{B}_A$, $E_A \subseteq E_C \cap A$, and $0 \leq \mathcal{P} \circ \mu_A(E_C \cap A) = \mathcal{P} \circ \mu(E_C \cap A) \leq \mathcal{P} \circ \mu(E_C) = 0$, where the last inequality follows from Proposition 11.4. Hence, (i) — (iii) of Proposition 11.114 are satisfied and $\bar{\mathcal{A}}$ is the completion of $\mathcal{A}$.

This completes the proof of the proposition. $\square$

Proposition 11.116 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, and $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable. Define $\bar{\nu} : \mathcal{B} \rightarrow [0, +\infty] \subset \mathbb{R}_e$ by $\bar{\nu}(E) = \int_E \mathcal{P} \circ f \, d\mu$, $\forall E \in \mathcal{B}$. Then, $\bar{\nu}$ is a measure on $(X, \mathcal{B})$. Define a function ν from $\mathcal{B}$ to $\mathcal{Y}$ by $\nu(E)$ is undefined, $\forall E \in \mathcal{B}$ with $\bar{\nu}(E) = +\infty$, and $\nu(E) = \int_E f \, d\mu$, $\forall E \in \mathcal{B}$ with $\bar{\nu}(E) < +\infty$. Then, $(X, \mathcal{B}, \nu)$ is a $\mathcal{Y}$ -valued measure space with $\mathcal{P} \circ \nu = \bar{\nu}$. We will call ν the $\mathcal{Y}$ -valued measure with kernel f over $\mathcal{X}$. Furthermore, the following statements hold.*

- (i) *If $\mathcal{X}$ is σ -finite, then ν is also σ -finite.*
- (ii) *f is absolutely integrable over $\mathcal{X}$ if, and only if, ν is finite.*
- (iii) *If $\mathcal{P} \circ f > 0$ a.e. in $\mathcal{X}$ and $\mathcal{X}$ is σ -finite and complete, then $(X, \mathcal{B}, \nu)$ is a σ -finite complete $\mathcal{Y}$ -valued measure space.*

Proof Since f is $\mathcal{B}$ -measurable, by Propositions 7.21 and 11.38, $\mathcal{P} \circ f$ is $\mathcal{B}$ -measurable. Then, $\bar{\nu} : \mathcal{B} \rightarrow [0, +\infty] \subset \mathbb{R}_e$ is well defined by Definition 11.79. By Proposition 11.75, $\bar{\nu}(\emptyset) = \int_{\emptyset} (\mathcal{P} \circ f) \, d\mu = 0$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, by Proposition 11.83, we have $\bar{\nu}(\bigcup_{i=1}^{\infty} E_i) = \int_{\bigcup_{i=1}^{\infty} E_i} (\mathcal{P} \circ f) \, d\mu = \sum_{i=1}^{\infty} \int_{E_i} (\mathcal{P} \circ f) \, d\mu = \sum_{i=1}^{\infty} \bar{\nu}(E_i)$. Hence, $\bar{\nu}$ is a measure on $(X, \mathcal{B})$. By Proposition 11.75, $\nu(\emptyset) = \int_{\emptyset} f \, d\mu = \vartheta_{\mathcal{Y}}$. $\forall E \in \mathcal{B}$ with $\bar{\nu}(E) = +\infty$, $\nu(E)$ is undefined. $\forall E \in \mathcal{B}$ with $\bar{\nu}(E) < +\infty$, by Proposition 11.92, $\nu(E) = \int_E f \, d\mu = \int_E f|_E \, d\mu_E \in \mathcal{Y}$, where $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$ is the measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$ with $\bigcup_{i=1}^{\infty} E_i = E$, we have $\sum_{i=1}^{\infty} \|\nu(E_i)\| = \sum_{i=1}^{\infty} \left\| \int_{E_i} f \, d\mu \right\| \leq \sum_{i=1}^{\infty} \int_{E_i} (\mathcal{P} \circ f) \, d\mu = \int_E (\mathcal{P} \circ f) \, d\mu = \bar{\nu}(E) < +\infty$, where the first equality follows from the fact that $E_i \subseteq E$ and $0 \leq \bar{\nu}(E_i) \leq \bar{\nu}(E) < +\infty$; the first inequality follows from Proposition 11.92; and the second equality follows from Proposition 11.83. Then, $\sum_{i=1}^{\infty} \nu(E_i) =$

$\sum_{i=1}^{\infty} \int_{E_i} f \, d\mu = \sum_{i=1}^{\infty} \int_{E_i} f|_E \, d\mu_E = \int_E f|_E \, d\mu_E = \int_E f \, d\mu = \nu(E)$, where the second, the third, and the fourth equalities follow from Proposition 11.92.

$\forall E \in \mathcal{B}$ with $\bar{\nu}(E) < +\infty$, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_i)_{i=1}^n \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^n E_i$, $f|_E$ is absolutely integrable over $\mathcal{E}$. $\sum_{i=1}^n \|\nu(E_i)\| = \sum_{i=1}^n \left\| \int_{E_i} f \, d\mu \right\| \leq \sum_{i=1}^n \int_{E_i} \mathcal{P} \circ f \, d\mu = \int_E \mathcal{P} \circ f \, d\mu = \bar{\nu}(E) < +\infty$, where the first equality follows from the fact that $E_i \subseteq E$ and $0 \leq \bar{\nu}(E_i) \leq \bar{\nu}(E) < +\infty$; the first inequality follows from Proposition 11.92; and the second equality follows from Proposition 11.83. Hence, $s_E := \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\nu(E_i)\| \leq \bar{\nu}(E)$.

On the other hand, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Proposition 11.95, $\exists$ a simple function $\phi : \mathcal{E} \rightarrow \mathcal{W}$ such that $0 \leq \left| \int_E (\mathcal{P} \circ f)|_E \, d\mu_E - \int_E (\mathcal{P} \circ \phi) \, d\mu_E \right| < \epsilon/2$ and $0 \leq \int_E \mathcal{P} \circ (\phi - f|_E) \, d\mu_E < \epsilon/2$. Let ϕ admit canonical representation $\phi = \sum_{i=1}^n w_i \chi_{E_i, E}$, where $n \in \mathbb{Z}_+$, $w_1, \dots, w_n \in \mathcal{W}$ are distinct and none equals to $\emptyset_{\mathcal{W}}$, $E_1, \dots, E_n \in \mathcal{B}_E$ are nonempty, pairwise disjoint, and $\mu_E(E_i) < +\infty$, $\forall i \in \{1, \dots, n\}$. Then, we have

$$\begin{aligned}
 \bar{\nu}(E) &= \int_E (\mathcal{P} \circ f) \, d\mu = \int_E (\mathcal{P} \circ f)|_E \, d\mu_E < \int_E (\mathcal{P} \circ \phi) \, d\mu_E + \epsilon/2 \\
 &= \sum_{i=1}^n \|\nu(E_i)\| + \sum_{i=1}^n \|w_i\| \mu_E(E_i) + \epsilon/2 - \sum_{i=1}^n \|\nu(E_i)\| \\
 &= \sum_{i=1}^n \|\nu(E_i)\| + \epsilon/2 + \sum_{i=1}^n \left\| \int_{E_i} \phi \, d\mu_E \right\| - \sum_{i=1}^n \left\| \int_{E_i} f|_E \, d\mu_E \right\| \\
 &\leq s_E + \epsilon/2 + \sum_{i=1}^n \left\| \int_{E_i} \phi \, d\mu_E - \int_{E_i} f|_E \, d\mu_E \right\| \\
 &= s_E + \epsilon/2 + \sum_{i=1}^n \left\| \int_{E_i} (\phi - f|_E) \, d\mu_E \right\| \\
 &\leq s_E + \epsilon/2 + \sum_{i=1}^n \int_{E_i} \mathcal{P} \circ (\phi - f|_E) \, d\mu_E \\
 &= s_E + \epsilon/2 + \int_{\bigcup_{i=1}^n E_i} \mathcal{P} \circ (\phi - f|_E) \, d\mu_E < s_E + \epsilon
 \end{aligned}$$

where the second equality follows from Proposition 11.92; the third equality follows from Proposition 11.75; the fourth equality follows from Proposition 11.75 and the fact that $0 \leq \bar{\nu}(E_i) \leq \bar{\nu}(E) < +\infty$; the fifth equality and the third inequality follow from Proposition 11.92; , and the last equality and the last inequality follows from Proposition 11.83. Hence, by the arbitrariness of ϵ , we have $\bar{\nu}(E) \leq s_E$. Therefore, $\bar{\nu}(E) = s_E$.

The above shows that $(X, \mathcal{B}, \nu)$ is a $\mathcal{Y}$ -valued measure space with $\mathcal{P} \circ \nu = \bar{\nu}$.

(i) Let $\mathcal{X}$ be σ -finite. Then, $\exists (X_i)_{i=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{i=1}^\infty X_i$ and $\mu(X_i) < +\infty$, $\forall i \in \mathbb{N}$. $\forall i \in \mathbb{N}$, $\forall m \in \mathbb{N}$, let $X_{i,m} := X_i \cap \{x \in X \mid \|f(x)\| \leq m\} \in \mathcal{B}$. Clearly, $X = \bigcup_{i=1}^\infty X_i = \bigcup_{i=1}^\infty \bigcup_{m=1}^\infty X_{i,m}$. By Bounded Convergence Theorem 11.77, we have $\bar{\nu}(X_{i,m}) = \int_{X_{i,m}} (\mathcal{P} \circ f) d\mu \leq m\mu(X_{i,m}) \leq m\mu(X_i) < +\infty$. Then, $\bar{\nu}$ is σ -finite. Hence, $(X, \mathcal{B}, \nu)$ is a σ -finite $\mathcal{Y}$ -valued measure space.

(ii) Let f be absolutely integrable over $\mathcal{X}$. Then, $\bar{\nu}(X) = \int_X (\mathcal{P} \circ f) d\mu < +\infty$. Hence, $(X, \mathcal{B}, \nu)$ is a finite $\mathcal{Y}$ -valued (pre-)measure space. On the other hand, let ν be finite. Then, $\bar{\nu}(X) = \int_X (\mathcal{P} \circ f) d\mu < +\infty$. Hence, f is absolutely integrable over $\mathcal{X}$.

(iii) By (i), $(X, \mathcal{B}, \nu)$ is σ -finite. $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \nu(E) = 0$, then $\int_E (\mathcal{P} \circ f) d\mu = 0$. By Proposition 11.96, $\mathcal{P} \circ f = 0$ a.e. in $\mathcal{E}$. Since $\mathcal{P} \circ f > 0$ a.e. in $\mathcal{X}$. Then, $\mu(E) = 0$. $\forall A \subseteq E$, by the completeness of $\mathcal{X}$, $A \in \mathcal{B}$. Hence, $(X, \mathcal{B}, \nu)$ is a σ -finite complete $\mathcal{Y}$ -valued measure space.

This completes the proof of the proposition. $\square$

Proposition 11.117 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space, $(X_n)_{n=1}^\infty \subseteq \mathcal{B}$ be such that $X = \bigcup_{n=1}^\infty X_n$ and $\mathcal{P} \circ \mu(X_n) < +\infty$, $\forall n \in \mathbb{N}$, and $\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$. Then, any $\mathcal{Y}$ -valued measure space $\bar{\mathcal{X}} := (X, \bar{\mathcal{B}}, \bar{\mu})$ with $\mathcal{X}_n$ as its finite $\mathcal{Y}$ -valued measure subspace, $\forall n \in \mathbb{N}$, must be identical to $\mathcal{X}$.*

Proof $\forall E \in \mathcal{B}$, $E \cap X_n \in \mathcal{B}_n \subseteq \bar{\mathcal{B}}$, $\forall n \in \mathbb{N}$. Then, $E = \bigcup_{n=1}^\infty (E \cap X_n) \in \bar{\mathcal{B}}$. Then, $\mathcal{B} \subseteq \bar{\mathcal{B}}$. On the other hand, $\forall E \in \bar{\mathcal{B}}$, $E \cap X_n \in \mathcal{B}_n \subseteq \mathcal{B}$, $\forall n \in \mathbb{N}$. Then, $E = \bigcup_{n=1}^\infty (E \cap X_n) \in \mathcal{B}$. Then, $\bar{\mathcal{B}} \subseteq \mathcal{B}$. Hence, $\bar{\mathcal{B}} = \mathcal{B}$.

$\forall E \in \mathcal{B}$ with $E \in \text{dom}(\mu)$, we have $\mathcal{P} \circ \mu(E) < +\infty$ and $\mu(E) \in \mathcal{Y}$. Let $E_n := E \cap (X_n \setminus \bigcup_{i=1}^{n-1} X_i) \in \mathcal{B}_n$, $\forall n \in \mathbb{N}$. Then, $\mu(E_n) = \mu_n(E_n) = \bar{\mu}(E_n)$, $\forall n \in \mathbb{N}$. Note that $\mathcal{P} \circ \mu(E) = \sum_{n=1}^\infty \mathcal{P} \circ \mu(E_n) = \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(E_n) = \sum_{n=1}^\infty \mathcal{P} \circ \bar{\mu}(E_n) = \mathcal{P} \circ \bar{\mu}(E) < +\infty$, where the second and third equalities follow from Proposition 11.115. Then, $E \in \text{dom}(\bar{\mu})$ and $\mu(E) = \sum_{n=1}^\infty \mu(E_n) = \sum_{n=1}^\infty \bar{\mu}(E_n) = \bar{\mu}(E) \in \mathcal{Y}$.

On the other hand, $\forall E \in \bar{\mathcal{B}} = \mathcal{B}$ with $E \in \text{dom}(\bar{\mu})$, we have $\mathcal{P} \circ \bar{\mu}(E) < +\infty$ and $\bar{\mu}(E) \in \mathcal{Y}$. Let $E_n := E \cap (X_n \setminus \bigcup_{i=1}^{n-1} X_i) \in \mathcal{B}_n$, $\forall n \in \mathbb{N}$. Then, $\mu(E_n) = \mu_n(E_n) = \bar{\mu}(E_n)$, $\forall n \in \mathbb{N}$. Note that $\mathcal{P} \circ \mu(E) = \sum_{n=1}^\infty \mathcal{P} \circ \mu(E_n) = \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(E_n) = \sum_{n=1}^\infty \mathcal{P} \circ \bar{\mu}(E_n) = \mathcal{P} \circ \bar{\mu}(E) < +\infty$, where the second and third equalities follow from Proposition 11.115. Then, $E \in \text{dom}(\mu)$ and $\mu(E) = \sum_{n=1}^\infty \mu(E_n) = \sum_{n=1}^\infty \bar{\mu}(E_n) = \bar{\mu}(E) \in \mathcal{Y}$.

Therefore, $\text{dom}(\mu) = \text{dom}(\bar{\mu})$ and $\mu = \bar{\mu}$. Hence, $\mathcal{X} = \bar{\mathcal{X}}$. This completes the proof of the proposition. $\square$

Proposition 11.118 *Let $\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)$ be a finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a Banach space, $\forall n \in \mathbb{N}$. $\forall n, m \in \mathbb{N}$, let $(X_n \cap X_m, \mathcal{B}_{n,m}, \mu_{n,m})$ be the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}_n$. Assume that $(\mathcal{X}_n)_{n=1}^\infty$ is consistent, that is $\mathcal{B}_{n,m} = \mathcal{B}_{m,n}$ and $\mu_{n,m} = \mu_{m,n}$,*

$\forall n, m \in \mathbb{N}$. Let $X := \bigcup_{n=1}^{\infty} X_n$, $\mathcal{B} := \{B \subseteq X \mid B \cap X_n \in \mathcal{B}_n, \forall n \in \mathbb{N}\}$, $\nu : \mathcal{B} \rightarrow [0, \infty] \subset \mathbb{R}_e$ be defined by $\nu(E) := \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(E_n)$, where $E_1 := E \cap X_1$ and $E_{n+1} := E \cap (X_{n+1} \setminus \bigcup_{i=1}^n X_i)$, $\forall n \in \mathbb{N}$, and μ be a function from $\mathcal{B}$ to $\mathcal{Y}$ defined by $\mu(E)$ is undefined, $\forall E \in \mathcal{B}$ with $\nu(E) = +\infty$, and $\mu(E) := \sum_{n=1}^{\infty} \mu_n(E_n) \in \mathcal{Y}$, $\forall E \in \mathcal{B}$ with $\nu(E) < +\infty$, where $(E_n)_{n=1}^{\infty}$ is defined as above. Then, $\mathcal{X} := (X, \mathcal{B}, \mu)$ is a σ -finite $\mathcal{Y}$ -valued measure space with $\mathcal{P} \circ \mu = \nu$ and $\mathcal{X}_n$ is the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$, $\forall n \in \mathbb{N}$, and $\mathcal{X}$ is unique among any $\mathcal{Y}$ -valued measure spaces on X . The above process is said to be the generation process on $(\mathcal{X}_n)_{n=1}^{\infty}$ for σ -finite $\mathcal{Y}$ -valued measure space $\mathcal{X}$.

Proof We will first show that $\mathcal{B}$ is a σ -algebra on X . Clearly, $\emptyset \in \mathcal{B}$ and $X \in \mathcal{B}$. $\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^{\infty} E_i$. $\forall n \in \mathbb{N}$, $E \cap X_n = \bigcup_{i=1}^{\infty} (E_i \cap X_n)$. Note that $\forall i \in \mathbb{N}$, $E_i \in \mathcal{B}$ implies $E_i \cap X_n \in \mathcal{B}_n$. Since $\mathcal{B}_n$ is a σ -algebra, then $E \cap X_n \in \mathcal{B}_n$. By the arbitrariness of n , we have $E \in \mathcal{B}$. $\forall E \in \mathcal{B}$, $E \cap X_n \in \mathcal{B}_n$, $\forall n \in \mathbb{N}$. Then, $(X \setminus E) \cap X_n = X_n \setminus E = X_n \setminus (X_n \cap E) \in \mathcal{B}_n$. By the arbitrariness of n , we have $X \setminus E \in \mathcal{B}$. Hence, $\mathcal{B}$ is a σ -algebra on X . $\forall n \in \mathbb{N}$, $\forall m \in \mathbb{N}$, $X_n \cap X_m \in \mathcal{B}_{n,m} = \mathcal{B}_{m,n} \subseteq \mathcal{B}_m$. By the arbitrariness of m , we have $X_n \in \mathcal{B}$.

Next, we will show that ν is a σ -finite measure on $(X, \mathcal{B})$. Clearly, $\nu(\emptyset) = 0$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$. $\forall i \in \mathbb{N}$, let $E_{i,1} := E_i \cap X_1 \in \mathcal{B}_1$ and $E_{i,n+1} := E_i \cap (X_{n+1} \setminus \bigcup_{j=1}^n X_j) \in \mathcal{B}_{n+1}$, $\forall n \in \mathbb{N}$. Let $B_1 := E \cap X_1 = \bigcup_{i=1}^{\infty} E_{i,1} \in \mathcal{B}_1$ and $B_{n+1} := E \cap (X_{n+1} \setminus \bigcup_{j=1}^n X_j) = \bigcup_{i=1}^{\infty} E_{i,n+1} \in \mathcal{B}_{n+1}$, $\forall n \in \mathbb{N}$. Then, $\nu(E) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(B_n) = \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_n(E_{i,n}) = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(E_{i,n}) = \sum_{i=1}^{\infty} \nu(E_i)$, where the first equality follows from the definition of ν ; the second equality follows from the fact that $\mathcal{P} \circ \mu_n$ is a measure on $(X_n, \mathcal{B}_n)$; the third equality follows from the fact that all summands are nonnegative; and the fourth equality follows from the definition of ν . Hence, ν is a measure on $(X, \mathcal{B})$. $\forall n \in \mathbb{N}$, let $X_{n,1} := X_n \cap X_1$ and $X_{n,i+1} := X_n \cap (X_{i+1} \setminus \bigcup_{j=1}^i X_j)$, $\forall i \in \mathbb{N}$. Then, $\nu(X_n) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_i(X_{n,i}) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_{i+1}((X_n \cap X_{i+1}) \setminus \bigcup_{j=1}^i X_j) + \mathcal{P} \circ \mu_1(X_n \cap X_1) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_{i+1,n}((X_n \cap X_{i+1}) \setminus \bigcup_{j=1}^i X_j) + \mathcal{P} \circ \mu_{1,n}(X_n \cap X_1) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_{n,i+1}((X_n \cap X_{i+1}) \setminus \bigcup_{j=1}^i X_j) + \mathcal{P} \circ \mu_{n,1}(X_n \cap X_1) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_n((X_n \cap X_{i+1}) \setminus \bigcup_{j=1}^i X_j) + \mathcal{P} \circ \mu_n(X_n \cap X_1) = \mathcal{P} \circ \mu_n(X_n) < +\infty$, where the first equality follows from the definition of ν ; the third equality follows from Proposition 11.115; the fourth equality follows from the consistency assumption; the fifth equality follows from Proposition 11.115; and the last equality follows from the fact that $\mathcal{P} \circ \mu_n$ is a measure on $(X_n, \mathcal{B}_n)$. Hence, ν is σ -finite.

Next, we will show that $(X, \mathcal{B}, \mu)$ is a σ -finite $\mathcal{Y}$ -valued measure space with $\mathcal{P} \circ \mu = \nu$. Clearly, $\mu(\emptyset) = \vartheta_{\mathcal{Y}}$. $\forall E \in \mathcal{B}$ with $\nu(E) = +\infty$, $\mu(E)$ is undefined. $\forall E \in \mathcal{B}$ with $\nu(E) < +\infty$, let $(B_n)_{n=1}^{\infty}$ be as defined in the previous paragraph. Then, $+\infty > \nu(E) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(B_n) \geq \sum_{n=1}^{\infty} \|\mu_n(B_n)\|$. This implies that $\mu(E) = \sum_{n=1}^{\infty} \mu_n(B_n) \in \mathcal{Y}$ by Propo-

sition 7.27. $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^\infty E_i$, $\forall i \in \mathbb{N}$, let $(E_{i,n})_{n=1}^\infty$ be as defined in the previous paragraph. Then, we have $\sum_{i=1}^\infty \|\mu(E_i)\| = \sum_{i=1}^\infty \|\sum_{n=1}^\infty \mu_n(E_{i,n})\| \leq \sum_{i=1}^\infty \sum_{n=1}^\infty \|\mu_n(E_{i,n})\| \leq \sum_{i=1}^\infty \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(E_{i,n}) = \sum_{i=1}^\infty \nu(E_i) = \nu(E) < +\infty$, where the first equality follows from the definition of μ ; the second inequality follows from the fact that $\mathcal{X}_n$ is a finite $\mathcal{Y}$ -valued measure space; the second equality follows from the definition of ν ; and the third equality follows from the fact that ν is a measure. We also have $\sum_{i=1}^\infty \mu(E_i) = \sum_{i=1}^\infty \sum_{n=1}^\infty \mu_n(E_{i,n}) = \sum_{n=1}^\infty \sum_{i=1}^\infty \mu_n(E_{i,n}) = \sum_{n=1}^\infty \mu_n(\bigcup_{i=1}^\infty E_{i,n}) = \sum_{n=1}^\infty \mu_n(B_n) = \mu(E)$, where the first equality follows from the definition of μ ; the second equality follows from the fact that the double summation is absolutely summable; the third equality follows from the fact that μ_n is a finite $\mathcal{Y}$ -valued measure; and the last equality follows from the definition of μ . Hence, (i) – (iii) of Definition 11.108 are satisfied.

$\forall E \in \mathcal{B}$ with $\nu(E) < +\infty$, let $(B_n)_{n=1}^\infty$ be as defined in the previous paragraph. $\forall m \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_i)_{i=1}^m \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^m E_i$, $\forall i \in \{1, \dots, m\}$, let $(E_{i,n})_{n=1}^\infty$ be as defined in the previous paragraph. Then, $\sum_{i=1}^m \|\mu(E_i)\| = \sum_{i=1}^m \|\sum_{n=1}^\infty \mu_n(E_{i,n})\| \leq \sum_{i=1}^m \sum_{n=1}^\infty \|\mu_n(E_{i,n})\| \leq \sum_{i=1}^m \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(E_{i,n}) = \sum_{i=1}^m \nu(E_i) = \nu(E) < \infty$. Hence,

$$s_E := \sup_{m \in \mathbb{Z}_+, (E_i)_{i=1}^m \subseteq \mathcal{B}, E = \bigcup_{i=1}^m E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq m} \sum_{i=1}^m \|\mu(E_i)\| \leq \nu(E)$$

On the other hand, $\nu(E) = \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(B_n) < +\infty$, where $B_1, B_2, \dots$ are pairwise disjoint. Since $(X_n, \mathcal{B}_n, \mu_n)$ is a finite $\mathcal{Y}$ -valued measure space, then, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\exists m_n \in \mathbb{Z}_+$, $\exists$ pairwise disjoint $(B_{n,i})_{i=1}^{m_n} \subseteq \mathcal{B}_n$ with $B_n = \bigcup_{i=1}^{m_n} B_{n,i}$ such that $\mathcal{P} \circ \mu_n(B_n) - 2^{-n}\epsilon < \sum_{i=1}^{m_n} \|\mu_n(B_{n,i})\| \leq \mathcal{P} \circ \mu_n(B_n)$. It is easy to see that $\mu_n(B_{n,i}) = \mu(B_{n,i})$, $\forall i \in \{1, \dots, m_n\}$. Then, $\nu(E) - \epsilon = \sum_{n=1}^\infty (\mathcal{P} \circ \mu_n(B_n) - 2^{-n}\epsilon) < \sum_{n=1}^\infty \sum_{i=1}^{m_n} \|\mu_n(B_{n,i})\| = \sum_{n=1}^\infty \sum_{i=1}^{m_n} \|\mu(B_{n,i})\| \leq \nu(E) < +\infty$. Then, $\exists N \in \mathbb{N}$ such that $\nu(E) \leq \sum_{n=1}^N \sum_{i=1}^{m_n} \|\mu(B_{n,i})\| + \epsilon \leq s_E + \epsilon$. By the arbitrariness of ϵ , we have $\nu(E) \leq s_E$. Therefore, $\nu(E) = s_E$ and (iv) of Definition 11.108 is satisfied. This shows that $(X, \mathcal{B}, \mu)$ is a σ -finite $\mathcal{Y}$ -valued measure space with $\mathcal{P} \circ \mu = \nu$.

Finally, we will show that $\mathcal{X}_n$ is the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$, $\forall n \in \mathbb{N}$. Fix any $n \in \mathbb{N}$. $\forall E \in \mathcal{B}$ with $E \subseteq X_n$, we have $E = E \cap X_n \in \mathcal{B}_n$. On the other hand, $\forall E \in \mathcal{B}_n$, we have $E \subseteq X_n$ and $\forall i \in \mathbb{N}$, $E \cap X_i = E \cap X_i \cap X_n \in \mathcal{B}_{n,i} = \mathcal{B}_{i,n} \subseteq \mathcal{B}_i$. Then, $E \in \mathcal{B}$. Hence, $\mathcal{B}_n = \{E \in \mathcal{B} \mid E \subseteq X_n\}$. $\forall E \in \mathcal{B}_n$, let $B_1 := E \cap X_1 \in \mathcal{B}_1$ and $B_{i+1} := E \cap (X_{i+1} \setminus \bigcup_{j=1}^i X_j) \in \mathcal{B}_{i+1}$, $\forall i \in \mathbb{N}$. Then, $\nu(E) = \sum_{i=1}^\infty \mathcal{P} \circ \mu_i(B_i)$. Note that $B_i \in \mathcal{B}_i$ and $B_i \subseteq X_n$, $\forall i \in \mathbb{N}$, this implies that $B_i \in \mathcal{B}_{i,n} = \mathcal{B}_{n,i} \subseteq \mathcal{B}_n$. Then, we have $\nu(E) = \sum_{i=1}^\infty \mathcal{P} \circ \mu_{i,n}(B_i) = \sum_{i=1}^\infty \mathcal{P} \circ \mu_{n,i}(B_i) = \sum_{i=1}^\infty \mathcal{P} \circ \mu_n(B_i) = \mathcal{P} \circ \mu_n(E) \leq \mathcal{P} \circ \mu_n(X_n) < +\infty$, where the first equality follows from the fact that $\mu_{i,n} = \mu_i|_{\mathcal{B}_{i,n}}$; the second equality follows from the consistency

assumption; the third equality follows from the fact that $\mu_{n,i} = \mu_n|_{\mathcal{B}_{n,i}}$; the fourth equality follows from countable additivity of $\mathcal{P} \circ \mu_n$; and the first inequality follows from Proposition 11.4. Then, $\mathcal{Y} \ni \mu(E) = \sum_{i=1}^{\infty} \mu_i(B_i) = \sum_{i=1}^{\infty} \mu_{i,n}(B_i) = \sum_{i=1}^{\infty} \mu_{n,i}(B_i) = \sum_{i=1}^{\infty} \mu_n(B_i) = \mu_n(E)$, where the first equality follows from the definition of μ ; the second equality follows from the fact that $\mu_{i,n} = \mu_i|_{\mathcal{B}_{i,n}}$; the third equality follows from the consistency assumption; the fourth equality follows from the fact that $\mu_{n,i} = \mu_n|_{\mathcal{B}_{n,i}}$; and the last equality follows from the fact that μ_n is a finite $\mathcal{Y}$ -valued measure. Hence, $\mu_n = \mu|_{\mathcal{B}_n}$. This shows that $\mathcal{X}_n$ is the subspace of $\mathcal{X}$. By Proposition 11.117, $\mathcal{X}$ is the unique $\mathcal{Y}$ -valued measure space on the set X such that $\mathcal{X}_n$ is the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$, $\forall n \in \mathbb{N}$.

This completes the proof of the proposition. $\square$

In the application of Proposition 11.118, if the sequence $(X_n)_{n=1}^{\infty}$ is pairwise disjoint. Then, the consistency requirement is automatically satisfied as long as $\mathcal{X}_n$ is a $\mathcal{Y}$ -valued measure space, $\forall n \in \mathbb{N}$.

Definition 11.119 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $\mathcal{W} \subseteq \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be a subspace, $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, and $\mathcal{I}(\mathcal{W}) := (\mathcal{R}(\mathcal{W}), \preceq)$ be the integration system on $\mathcal{W}$. $\forall R = \{(w_\alpha, U_\alpha) \mid \alpha \in \Lambda\} \in \mathcal{R}(\mathcal{W})$, define $F_R := \sum_{\alpha \in \Lambda} w_\alpha \mu(f_{\text{inv}}(U_\alpha)) \in \mathcal{Z}$. This defines a net $(F_R)_{R \in \mathcal{I}(\mathcal{W})}$. f is said to be integrable over $\mathcal{X}$ if the net admits a limit in $\mathcal{Z}$. In this case, $\lim_{R \in \mathcal{I}(\mathcal{W})} F_R \in \mathcal{Z}$ is said to be the integral of f over $\mathcal{X}$ and denoted by $\int_X f d\mu$ or $\int_X f(x) d\mu(x)$. When $\mathcal{Z} = \mathbb{R}$, we will denote $\int_X f d\mu := \int_X f(x) d\mu(x) := \lim_{R \in \mathcal{I}(\mathcal{W})} F_R \in \mathbb{R}_e$, whenever the limit exists.

Definition 11.120 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space with $\mathcal{P} \circ \mu(X) = +\infty$, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $\bar{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$ be the total variation, and $f : X \rightarrow \mathcal{W} := \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be $\mathcal{B}$ -measurable. Define $\mathcal{M}(\mathcal{X})$ to be the directed system, $\mathcal{M}(\bar{\mathcal{X}})$, of subsets of X as defined in Definition 11.71. $\forall A \in \mathcal{M}(\mathcal{X})$, let $(A, \mathcal{B}_A, \mu_A)$ be the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$ and define $F_A := \int_A f|_A d\mu_A \in \mathcal{Z}$, if the integral exists. This defines a net $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ when all of the integrals involved exist. f is said to be integrable over $\mathcal{X}$ if the net is well-defined (i. e., all integrals involved exist) and admits a limit in $\mathcal{Z}$. In this case, $\lim_{A \in \mathcal{M}(\mathcal{X})} F_A \in \mathcal{Z}$ is said to be the integral of f over $\mathcal{X}$ and denoted by $\int_X f d\mu$ or $\int_X f(x) d\mu(x)$. When $\mathcal{Z} = \mathbb{R}$, we will allow the net $(F_A)_{A \in \mathcal{M}(\mathcal{X})} \subseteq \mathbb{R}_e$ and denote $\int_X f d\mu := \int_X f(x) d\mu(x) := \lim_{A \in \mathcal{M}(\mathcal{X})} F_A \in \mathbb{R}_e$, whenever the limit exists.

Proposition 11.121 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $f : X \rightarrow \mathcal{W} := \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be $\mathcal{B}$ -measurable, and $\bar{\mathcal{X}} := (X, \bar{\mathcal{B}}, \bar{\mu})$ be a finite $\mathcal{Y}$ -valued measure space that satisfies $\mathcal{B} \subseteq \bar{\mathcal{B}}$ and $\mu = \bar{\mu}|_{\mathcal{B}}$. Then, $\int_X f d\mu = \int_X f d\bar{\mu}$, whenever one of the integrals exists.

Proof Let $\mathcal{J}(\mathcal{W})$ be the integration system on $\mathcal{W}$ as defined in Proposition 11.69, $(F_R)_{R \in \mathcal{J}(\mathcal{W})}$ be the net for $\int_X f d\mu$, and $(\bar{F}_R)_{R \in \mathcal{J}(\mathcal{W})}$ be the net for $\int_X f d\bar{\mu}$, as defined in Definition 11.119. $\forall R := \{(w_\alpha, U_\alpha) \mid \alpha \in \Lambda\} \in \mathcal{J}(\mathcal{W})$, $\forall \alpha \in \Lambda$, we have $f_{\text{inv}}(U_\alpha) \in \mathcal{B} \subseteq \bar{\mathcal{B}}$ since f is $\mathcal{B}$ -measurable, and $\mu(f_{\text{inv}}(U_\alpha)) = \bar{\mu}(f_{\text{inv}}(U_\alpha)) \in \mathcal{Y}$ since $\mu = \bar{\mu}|_{\mathcal{B}}$. Then, $F_R = \sum_{\alpha \in \Lambda} w_\alpha \mu(f_{\text{inv}}(U_\alpha)) = \sum_{\alpha \in \Lambda} w_\alpha \bar{\mu}(f_{\text{inv}}(U_\alpha)) = \bar{F}_R \in \mathcal{Z}$. Hence, $\int_X f d\mu = \lim_{R \in \mathcal{J}(\mathcal{W})} F_R = \lim_{R \in \mathcal{J}(\mathcal{W})} \bar{F}_R = \int_X f d\bar{\mu}$, whenever one of the integrals exists. This completes the proof of the proposition. $\square$

Lemma 11.122 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $f : X \rightarrow \mathcal{W} := \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be $\mathcal{B}$ -measurable, $\hat{X} \in \mathcal{B}$, and $\hat{\mathcal{X}} := (\hat{X}, \hat{\mathcal{B}}, \hat{\mu})$ be the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$ as defined in Proposition 11.115. Assume that $\mathcal{P} \circ \mu(X \setminus \hat{X}) = 0$. Then, $\int_X f d\mu = \int_{\hat{X}} f|_{\hat{X}} d\hat{\mu}$, whenever one of the integrals exists.

Proof Let $\mathcal{J}(\mathcal{W})$ be the integration system on $\mathcal{W}$ as defined in Proposition 11.69, $(F_R)_{R \in \mathcal{J}(\mathcal{W})}$ be the net for $\int_X f d\mu$, and $(\hat{F}_R)_{R \in \mathcal{J}(\mathcal{W})}$ be the net for $\int_{\hat{X}} f|_{\hat{X}} d\hat{\mu}$, as defined in Definition 11.119. $\forall R := \{(w_\alpha, U_\alpha) \mid \alpha \in \Lambda\} \in \mathcal{J}(\mathcal{W})$, $\forall \alpha \in \Lambda$, we have $f_{\text{inv}}(U_\alpha) \cap \hat{X} = f|_{\hat{X} \text{ inv}}(U_\alpha) \in \hat{\mathcal{B}} \subseteq \mathcal{B}$ since f is $\mathcal{B}$ -measurable, and $0 \leq \mathcal{P} \circ \mu(f_{\text{inv}}(U_\alpha) \cap (X \setminus \hat{X})) \leq \mathcal{P} \circ \mu(X \setminus \hat{X}) = 0$. This implies that $\mu(f_{\text{inv}}(U_\alpha)) = \mu(f_{\text{inv}}(U_\alpha) \cap \hat{X}) + \mu(f_{\text{inv}}(U_\alpha) \cap (X \setminus \hat{X})) = \mu(f|_{\hat{X} \text{ inv}}(U_\alpha)) = \hat{\mu}(f|_{\hat{X} \text{ inv}}(U_\alpha))$ by Fact 11.110. Then, $F_R = \sum_{\alpha \in \Lambda} w_\alpha \mu(f_{\text{inv}}(U_\alpha)) = \sum_{\alpha \in \Lambda} w_\alpha \hat{\mu}(f|_{\hat{X} \text{ inv}}(U_\alpha)) = \hat{F}_R \in \mathcal{Z}$. Hence, $\int_X f d\mu = \lim_{R \in \mathcal{J}(\mathcal{W})} F_R = \lim_{R \in \mathcal{J}(\mathcal{W})} \hat{F}_R = \int_{\hat{X}} f|_{\hat{X}} d\hat{\mu}$, whenever one of the integrals exists. This completes the proof of the proposition. $\square$

Proposition 11.123 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space with $\mathcal{P} \circ \mu(X) = +\infty$, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $f : X \rightarrow \mathcal{W} := \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be $\mathcal{B}$ -measurable, and $\bar{\mathcal{X}} := (X, \bar{\mathcal{B}}, \bar{\mu})$ be the completion of $\mathcal{X}$. Then, $\int_X f d\mu = \int_X f d\bar{\mu}$, whenever one of the integrals exists.

Proof Let $\hat{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu =: \nu)$ and $\hat{\bar{\mathcal{X}}} := (X, \bar{\mathcal{B}}, \mathcal{P} \circ \bar{\mu} =: \bar{\nu})$. Then, $\hat{\mathcal{X}}$ and $\hat{\bar{\mathcal{X}}}$ are measure spaces and $\hat{\bar{\mathcal{X}}}$ is the completion of $\hat{\mathcal{X}}$ by Proposition 11.114. Let $(F_A)_{A \in \mathcal{M}(\hat{\mathcal{X}})}$ be the net for $\int_X f d\mu$ and $(\bar{F}_{\bar{A}})_{\bar{A} \in \mathcal{M}(\hat{\bar{\mathcal{X}}})}$ be the net for $\int_X f d\bar{\mu}$ as defined in Definition 11.120. Clearly, $\mathcal{M}(\hat{\mathcal{X}}) \subseteq \mathcal{M}(\hat{\bar{\mathcal{X}}})$, since $\hat{\bar{\mathcal{X}}}$ is the completion of $\hat{\mathcal{X}}$. We will distinguish two exhaustive cases: Case 1: $\int_X f d\bar{\mu}$ exists; Case 2: $\int_X f d\mu$ exists.

Case 1: $\int_X f d\bar{\mu}$ exists. Then, the net $(\bar{F}_{\bar{A}})_{\bar{A} \in \mathcal{M}(\hat{\bar{\mathcal{X}}})}$ is well defined. $\forall A \in \mathcal{M}(\hat{\mathcal{X}})$, we have $A \in \mathcal{B}$ and $\nu(A) < +\infty$. Then, $A \in \mathcal{M}(\hat{\bar{\mathcal{X}}})$. By Propositions 11.121 and 11.114, $F_A = \int_A f|_A d\mu_A = \int_A f|_A d\bar{\mu}_A = \bar{F}_A$, where $(A, \mathcal{B}_A, \mu_A)$ is the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$ and

$(A, \bar{\mathcal{B}}_A, \bar{\mu}_A)$ is the finite complete $\mathcal{Y}$ -valued measure subspace of $\bar{\mathcal{X}}$. Then, $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ is well defined. $\forall \bar{A} \in \mathcal{M}(\bar{\mathcal{X}})$, we have $\bar{A} \in \bar{\mathcal{B}}$ with $\bar{\nu}(\bar{A}) < +\infty$. By Proposition 11.114 and its proof, $\exists \tilde{A}, \tilde{B}, \tilde{C} \subseteq X$ with $\bar{A} = \tilde{A} \cup \tilde{B}$, $\tilde{B}, \tilde{C} \in \mathcal{B}$, $\tilde{A} \subseteq \tilde{C}$, $\nu(\tilde{C}) = 0$, and $\bar{\nu}(\bar{A}) = \nu(\tilde{B})$. Then, $\bar{A} \subseteq \tilde{B} \cup \tilde{C} =: A \in \mathcal{B}$ and $\nu(A) \leq \nu(\tilde{B}) + \nu(\tilde{C}) = \nu(\tilde{B}) = \bar{\nu}(\bar{A}) < +\infty$. This shows that $A \in \mathcal{M}(\mathcal{X})$ and $\bar{A} \subseteq A$. Hence, $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ is a subnet of $(\bar{F}_{\bar{A}})_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})}$. By Proposition 3.70, $\int_X f d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_A = \lim_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})} \bar{F}_{\bar{A}} = \int_X f d\bar{\mu}$.

Case 2: $\int_X f d\mu$ exists. Then, the net $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ is well defined. $\forall \bar{A} \in \mathcal{M}(\bar{\mathcal{X}})$, we have $\bar{A} \in \bar{\mathcal{B}}$ with $\bar{\nu}(\bar{A}) < +\infty$. By Proposition 11.114 and its proof, $\exists \tilde{A}, \tilde{B}, \tilde{C} \subseteq X$ with $\bar{A} = \tilde{A} \cup \tilde{B}$, $\tilde{B}, \tilde{C} \in \mathcal{B}$, $\tilde{A} \subseteq \tilde{C}$, $\nu(\tilde{C}) = 0$, and $\bar{\nu}(\bar{A}) = \nu(\tilde{B})$. Let $\tilde{B} =: A \in \mathcal{B}$. Then, $0 \leq \bar{\nu}(\bar{A} \setminus A) = \bar{\nu}(\tilde{A} \setminus \tilde{B}) \leq \bar{\nu}(\tilde{C}) = \nu(\tilde{C}) = 0$ and $\nu(A) = \bar{\nu}(\bar{A}) < +\infty$. This shows that $A \in \mathcal{M}(\mathcal{X})$ and $A \subseteq \bar{A}$. Let $\mathcal{A} := (A, \mathcal{B}_A, \mu_A)$ be the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$, $\bar{\mathcal{A}} := (\bar{A}, \bar{\mathcal{B}}_{\bar{A}}, \bar{\mu}_{\bar{A}})$ be the finite complete $\mathcal{Y}$ -valued measure subspace of $\bar{\mathcal{X}}$, and $\hat{\mathcal{A}} := (A, \bar{\mathcal{B}}_{\bar{A}}, \bar{\mu}_{\bar{A}})$ be the finite complete $\mathcal{Y}$ -valued measure subspace of $\hat{\mathcal{A}}$. Since $\bar{\mathcal{X}}$ is the completion of $\mathcal{X}$, then, by Proposition 11.115, $\hat{\mathcal{A}}$ is the completion of $\mathcal{A}$. By Proposition 11.121, we have $\int_A f|_A d\bar{\mu}_A = \int_A f|_A d\mu_A = F_A$. By Lemma 11.122, we have $\bar{F}_{\bar{A}} = \int_{\bar{A}} f|_{\bar{A}} d\bar{\mu}_{\bar{A}} = \int_A f|_A d\bar{\mu}_{\bar{A}} = F_A$. Hence, $\bar{F}_{\bar{A}} = F_A$. Then, the net $(\bar{F}_{\bar{A}})_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})}$ is well defined.

Fix any open set U ($U \subseteq \mathbb{R}_e$ if $\mathcal{Z} = \mathbb{R}$ or $U \subseteq \mathcal{Z}$ if $\mathcal{Z} \neq \mathbb{R}$) with $\int_X f d\mu \in U$. Since $\lim_{A \in \mathcal{M}(\mathcal{X})} F_A = \int_X f d\mu \in U$, then $\exists A_0 \in \mathcal{M}(\mathcal{X})$ such that $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have $F_A \in U$. Note that $A_0 \in \mathcal{B} \subseteq \bar{\mathcal{B}}$ and $\nu(A_0) = \bar{\nu}(A_0) < +\infty$. Then, $A_0 \in \mathcal{M}(\bar{\mathcal{X}})$. $\forall \bar{A} \in \mathcal{M}(\bar{\mathcal{X}})$ with $A_0 \subseteq \bar{A}$, then $\bar{A} \in \bar{\mathcal{B}}$ and $\bar{\nu}(\bar{A}) < +\infty$. By Proposition 11.114 and its proof, $\exists \tilde{A}, \tilde{B}, \tilde{C} \subseteq X$ with $\bar{A} = \tilde{A} \cup \tilde{B}$, $\tilde{B}, \tilde{C} \in \mathcal{B}$, $\tilde{A} \subseteq \tilde{C}$, $\nu(\tilde{C}) = 0$, and $\bar{\nu}(\bar{A}) = \nu(\tilde{B})$. Let $A := A_0 \cup \tilde{B} \in \mathcal{B}$. Then, $A_0 \subseteq A \subseteq \bar{A}$, $0 \leq \bar{\nu}(\bar{A} \setminus A) = \bar{\nu}((\tilde{A} \cup \tilde{B}) \setminus (A_0 \cup \tilde{B})) \leq \bar{\nu}((\tilde{A} \cup \tilde{B}) \setminus \tilde{B}) \leq \bar{\nu}(\tilde{A} \setminus \tilde{B}) \leq \bar{\nu}(\tilde{C}) = \nu(\tilde{C}) = 0$, and $\nu(A) \leq \nu(\tilde{B}) + \nu(A_0) = \bar{\nu}(\bar{A}) + \nu(A_0) < +\infty$. This shows that $A \in \mathcal{M}(\mathcal{X})$. Let $\mathcal{A} := (A, \mathcal{B}_A, \mu_A)$ be the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$, $\bar{\mathcal{A}} := (\bar{A}, \bar{\mathcal{B}}_{\bar{A}}, \bar{\mu}_{\bar{A}})$ be the finite complete $\mathcal{Y}$ -valued measure subspace of $\bar{\mathcal{X}}$, and $\hat{\mathcal{A}} := (A, \bar{\mathcal{B}}_{\bar{A}}, \bar{\mu}_{\bar{A}})$ be the finite complete $\mathcal{Y}$ -valued measure subspace of $\hat{\mathcal{A}}$. Since $\bar{\mathcal{X}}$ is the completion of $\mathcal{X}$, then, by Proposition 11.115, $\hat{\mathcal{A}}$ is the completion of $\mathcal{A}$. By Proposition 11.121, we have $\int_A f|_A d\bar{\mu}_A = \int_A f|_A d\mu_A = F_A$. By Lemma 11.122, we have $\bar{F}_{\bar{A}} = \int_{\bar{A}} f|_{\bar{A}} d\bar{\mu}_{\bar{A}} = \int_A f|_A d\bar{\mu}_{\bar{A}} = F_A \in U$. Therefore, we have $\int_X f d\bar{\mu} = \lim_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})} \bar{F}_{\bar{A}} = \int_X f d\mu$.

Hence, in both cases, we have $\int_X f d\mu = \int_X f d\bar{\mu}$. This completes the proof of the proposition. $\square$

Definition 11.124 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, and $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$. $\phi : X \rightarrow \mathcal{Z}$ is said to be a simple function if $\exists n \in \mathbb{Z}_+$, $\exists z_1, \dots, z_n \in$

$\mathcal{Z}$, and $\exists A_1, \dots, A_n \in \mathcal{B}$ with $\mathcal{P} \circ \mu(A_i) < +\infty$, $i = 1, \dots, n$, such that $\phi(x) = \sum_{i=1}^n z_i \chi_{A_i, X}(x)$, $\forall x \in X$. We will say that a simple function ϕ is in canonical representation if $z_1, \dots, z_n$ are distinct and none equals to $\vartheta_{\mathcal{Z}}$, and $A_1, \dots, A_n$ are nonempty and pairwise disjoint.

Clearly, every simple function admits a unique canonical representation.

Proposition 11.125 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $\phi : X \rightarrow \mathcal{B}(\mathcal{Y}, \mathcal{Z}) =: \mathcal{W}$ be a simple function in canonical representation, i. e., $\exists n \in \mathbb{Z}_+$, $\exists w_1, \dots, w_n \in \mathcal{W}$, which are distinct and none equals to $\vartheta_{\mathcal{W}}$, $\exists A_1, \dots, A_n \in \mathcal{B}$, which are nonempty, pairwise disjoint, and $\mathcal{P} \circ \mu(A_i) < +\infty$, $i = 1, \dots, n$, such that $\phi(x) = \sum_{i=1}^n w_i \chi_{A_i, X}(x)$, $\forall x \in X$. Then, $\int_X \phi d\mu = \sum_{i=1}^n w_i \mu(A_i) =: I \in \mathcal{Z}$ and $0 \leq \left\| \int_X \phi d\mu \right\| \leq \int_X \mathcal{P} \circ \phi d\mathcal{P} \circ \mu < +\infty$.*

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{P} \circ \mu(X) < +\infty$; Case 2: $\mathcal{P} \circ \mu(X) = +\infty$.

Case 1: $\mathcal{P} \circ \mu(X) < +\infty$. Let $w_{n+1} := \vartheta_{\mathcal{W}}$ and $A_{n+1} := X \setminus (\bigcup_{i=1}^n A_i) \in \mathcal{B}$. Let $\epsilon_0 := \min\{1, \min_{1 \leq i < j \leq n+1} \|w_i - w_j\|\} > 0$, by the assumption. Let $(\Phi_R)_{R \in \mathcal{J}(\mathcal{W})}$ be the net as defined in Definition 11.119 for $\int_X \phi d\mu$.

$\forall \epsilon \in (0, \epsilon_0/2) \subset \mathbb{R}$, let $U_i := \mathcal{B}_{\mathcal{W}}(w_i, \epsilon) \in \mathcal{B}_{\mathcal{B}}(\mathcal{W})$, $i = 1, \dots, n+1$, which are clearly pairwise disjoint and nonempty. We will distinguish two exhaustive and mutually exclusive subcases: Case 1a: $\mathcal{W} = \bigcup_{i=1}^{n+1} U_i$; Case 1b: $\mathcal{W} \supset \bigcup_{i=1}^{n+1} U_i$.

Case 1a: $\mathcal{W} = \bigcup_{i=1}^{n+1} U_i$. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{W} = \mathcal{B}_{\mathcal{W}}(\vartheta_{\mathcal{W}}, \delta)$. Then, $\mathcal{W}$ must be the trivial normed linear space, i. e., $\mathcal{W}$ is a singleton set containing $\vartheta_{\mathcal{W}}$. There is a single representation of $\mathcal{W}$, which is $R_0 := \{(\vartheta_{\mathcal{W}}, \mathcal{W})\}$. Then, $\Phi_{R_0} = \vartheta_{\mathcal{Z}}$. Clearly, we must have $n = 0$, since $w_1, \dots, w_n \in \mathcal{Y}$ are distinct and none equals to $\vartheta_{\mathcal{W}}$. Then, $I = \vartheta_{\mathcal{Z}}$. Hence, $\forall R \in \mathcal{J}(\mathcal{W})$ with $R_0 \preceq R$, $\|\Phi_R - I\| = \|\Phi_{R_0} - I\| = 0$.

Case 1b: $\mathcal{W} \supset \bigcup_{i=1}^{n+1} U_i$. Let $U_{n+2} = \mathcal{W} \setminus (\bigcup_{i=1}^{n+1} U_i) \in \mathcal{B}_{\mathcal{B}}(\mathcal{W})$, which is nonempty. Let $w_{n+2} \in U_{n+2}$ be such that $\|w_{n+2}\| < \inf_{w \in U_{n+2}} \|w\| + 1$. Let $R_0 := \{(w_i, U_i) \mid i = 1, \dots, n+2\}$. It is easy to check that $R_0 \in \mathcal{J}(\mathcal{W})$. $\forall \tilde{R} \in \mathcal{J}(\mathcal{W})$ with $R_0 \preceq \tilde{R}$, then $\tilde{R} = \left\{ (\tilde{w}_\alpha, \tilde{U}_\alpha) \mid \alpha \in \Lambda \right\}$, where Λ is a finite index set, $(\tilde{U}_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{B}_{\mathcal{B}}(\mathcal{W})$ are nonempty and pairwise disjoint, $\bigcup_{\alpha \in \Lambda} \tilde{U}_\alpha = \mathcal{W}$, $\tilde{w}_\alpha \in \tilde{U}_\alpha$, and $\|\tilde{w}_\alpha\| < \inf_{w \in \tilde{U}_\alpha} \|w\| + 1$, $\forall \alpha \in \Lambda$. $\forall i \in \{1, \dots, n+2\}$, by $R_0 \preceq \tilde{R}$, $\exists! \alpha_i \in \Lambda$ such that $w_i = \tilde{w}_{\alpha_i}$ and, by $\tilde{R} \in \mathcal{J}(\mathcal{W})$, $\exists! \bar{\alpha}_i \in \Lambda$ such that $w_i \in \tilde{U}_{\bar{\alpha}_i}$. Since $(\tilde{U}_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{B}_{\mathcal{B}}(\mathcal{W})$ are nonempty and pairwise disjoint and $\tilde{w}_\alpha \in \tilde{U}_\alpha$, $\forall \alpha \in \Lambda$, then $\alpha_i = \bar{\alpha}_i$. Again by $R_0 \preceq \tilde{R}$, we have $\tilde{U}_{\alpha_i} \subseteq U_i$. Hence, $\alpha_1, \dots, \alpha_{n+2}$ are distinct. Let $\bar{\Lambda} := \{\alpha_1, \dots, \alpha_{n+2}\}$. $\forall \alpha \in \Lambda \setminus \bar{\Lambda}$, $\phi_{\text{inv}}(\tilde{U}_\alpha) = \emptyset$. Then, $\Phi_{\tilde{R}} = \sum_{\alpha \in \Lambda} \tilde{w}_\alpha \mu(\phi_{\text{inv}}(\tilde{U}_\alpha)) = \sum_{i=1}^{n+2} \tilde{w}_{\alpha_i} \mu(\phi_{\text{inv}}(\tilde{U}_{\alpha_i})) = \sum_{i=1}^{n+2} w_i \mu(\phi_{\text{inv}}(U_{\alpha_i}))$.

Note that $\forall i \in \{1, \dots, n+1\}$, $\phi_{\text{inv}}(\tilde{U}_{\alpha_i}) = A_i$ and $\mu(A_i) \in \mathcal{Y}$. Furthermore, $\phi_{\text{inv}}(\tilde{U}_{\alpha_{n+2}}) = \emptyset$. Hence, $\Phi_{\tilde{R}} = \sum_{i=1}^n w_i \mu(A_i) = I \in \mathcal{Z}$. Then, we have $\|\Phi_{\tilde{R}} - I\| = 0$.

In both subcases, we have $\exists R_0 \in \mathcal{I}(\mathcal{W})$, $\forall R \in \mathcal{I}(\mathcal{W})$ with $R_0 \preceq R$, we have $\|\Phi_R - I\| = 0 < \epsilon$. Hence, $\int_X \phi \, d\mu = \lim_{R \in \mathcal{I}(\mathcal{W})} \Phi_R = I$. Note that $0 \leq \|\int_X \phi \, d\mu\| \leq \sum_{i=1}^n \|w_i\| \|\mu(A_i)\| \leq \sum_{i=1}^n \|w_i\| \mathcal{P} \circ \mu(A_i) = \int_X \mathcal{P} \circ \phi \, d\mathcal{P} \circ \mu < +\infty$, where the second inequality follows from Proposition 7.64; the third inequality follows from Definition 11.108; the equality follows from what we have proved in this proposition; and the last inequality follows from Definition 11.124.

Case 2: $\mathcal{P} \circ \mu(X) = +\infty$. $\forall A \in \mathcal{M}(\mathcal{X})$, $\mathcal{P} \circ \mu(A) < +\infty$. $\phi|_A : A \rightarrow \mathcal{W}$ is a simple function. By Case 1, $\Phi_A = \int_A \phi|_A \, d\mu_A \in \mathcal{Z}$ is well-defined, where $(A, \mathcal{B}_A, \mu_A)$ is the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$. Hence, the net $(\Phi_A)_{A \in \mathcal{M}(\mathcal{X})}$ is well-defined. Take $A_0 := \bigcup_{i=1}^n A_i \in \mathcal{B}$. Then, $\mathcal{P} \circ \mu(A_0) = \sum_{i=1}^n \mathcal{P} \circ \mu(A_i) < +\infty$. Therefore, $A_0 \in \mathcal{M}(\mathcal{X})$. $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, by Case 1, $\Phi_A = \sum_{i=1}^n w_i \mu(A_i) = I \in \mathcal{Z}$. Then, $\int_X \phi \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} \Phi_A = I \in \mathcal{Z}$. Note that $0 \leq \|\int_X \phi \, d\mu\| \leq \sum_{i=1}^n \|w_i\| \|\mu(A_i)\| \leq \sum_{i=1}^n \|w_i\| \mathcal{P} \circ \mu(A_i) = \int_X \mathcal{P} \circ \phi \, d\mathcal{P} \circ \mu < +\infty$, where the second inequality follows from Proposition 7.64; the third inequality follows from Definition 11.108; the equality follows from what we have proved in this proposition; and the last inequality follows from Definition 11.124.

This completes the proof of the proposition. $\square$

Clearly, the above result holds for simple functions given in any form, not necessarily in the canonical representation. It also shows that the integrals of simple functions are linear.

Lemma 11.126 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $\phi_n : X \rightarrow \mathcal{W}$ be a simple function, $\forall n \in \mathbb{N}$, $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $g_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $g : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $\bar{\mathcal{W}} := \text{span}(\{z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \exists \cdot z = wy\}) \subseteq \mathcal{Z}$ be the Banach subspace of $\mathcal{Z}$. Assume that*

(i) $\lim_{n \in \mathbb{N}} \phi_n = f$ a.e. in $\mathcal{X}$ and $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$;

(ii) $\|\phi_n(x)\| \leq g_n(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$;

(iii) $\int_X g_n \, d\mathcal{P} \circ \mu \in \mathbb{R}$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \int_X g_n \, d\mathcal{P} \circ \mu = \int_X g \, d\mathcal{P} \circ \mu \in \mathbb{R}$.

Then, f is integrable over $\mathcal{X}$, $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu \in \bar{\mathcal{W}} \subseteq \mathcal{Z}$, and $0 \leq \|\int_X f \, d\mu\| \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty$.

Proof Let $\bar{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$ be the finite measure space as defined in Definition 11.108. Let $E := \{x \in X \mid (\phi_n(x))_{n=1}^\infty \text{ does not converge to } f(x) \text{ or } (g_n(x))_{n=1}^\infty \text{ does not converge to } g(x)\}$. Then, by (i), $E \in \mathcal{B}$ and

$\mathcal{P} \circ \mu(E) = 0$. Let $(F_R)_{R \in \mathcal{J}(\mathcal{W})}$ be the net for $\int_X f d\mu$ as defined in Definition 11.119.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by (i), (iii), and Lemma 11.85, $\exists \delta \in (0, \epsilon/6] \subset \mathbb{R}$ and $\exists n_0 \in \mathbb{N}$ such that $\forall A \in \mathcal{B}$ with $\mathcal{P} \circ \mu(A) < \delta$, $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have $0 \leq \int_A g d\mathcal{P} \circ \mu < \epsilon/6$ and $0 \leq \int_A g_n d\mathcal{P} \circ \mu < \epsilon/6$. By Egoroff's Theorem 11.55, $\exists \bar{E}_1 \in \mathcal{B}$ with $\mathcal{P} \circ \mu(\bar{E}_1) < \delta$ such that $\left(\phi_n|_{X \setminus \bar{E}_1}\right)_{n=1}^\infty$ converges uniformly to $f|_{X \setminus \bar{E}_1}$. Let $E_1 := E \cup \bar{E}_1$. Then, $E_1 \in \mathcal{B}$ with $\mathcal{P} \circ \mu(E_1) = \mathcal{P} \circ \mu(\bar{E}_1) < \delta$ such that $\left(\phi_n|_{X \setminus E_1}\right)_{n=1}^\infty$ converges uniformly to $f|_{X \setminus E_1}$. Then, $\exists n_1 \in \mathbb{N}$ with $n_1 \geq n_0$, $\forall n \in \mathbb{N}$ with $n \geq n_1$, $\forall x \in X \setminus E_1$, $\|\phi_n(x) - f(x)\| < \frac{\epsilon}{6\mathcal{P} \circ \mu(X) + 1} =: \bar{\epsilon} \in (0, \infty) \subset \mathbb{R}$. Fix any $n \geq n_1$. Let ϕ_n admit the canonical representation $\phi_n = \sum_{i=1}^{\bar{n}} w_i \chi_{A_i, X}$, where $\bar{n} \in \mathbb{Z}_+$, $w_1, \dots, w_{\bar{n}} \in \mathcal{W}$ are distinct and none equals to $\vartheta_{\mathcal{W}}$, and $A_1, \dots, A_{\bar{n}} \in \mathcal{B}$ are nonempty, pairwise disjoint, and $\mathcal{P} \circ \mu(A_i) < +\infty$, $i = 1, \dots, \bar{n}$. Let $w_{\bar{n}+1} := \vartheta_{\mathcal{W}}$ and $A_{\bar{n}+1} := X \setminus (\bigcup_{i=1}^{\bar{n}} A_i) \in \mathcal{B}$. Then, $X = \bigcup_{i=1}^{\bar{n}+1} A_i$ and the sets in the union are pairwise disjoint and of finite total variation. Define $\bar{V}_i := \mathcal{B}_{\mathcal{W}}(w_i, \bar{\epsilon})$, $i = 1, \dots, \bar{n} + 1$. Define $V_1 := \bar{V}_1 \in \mathcal{B}_{\mathcal{B}}(\mathcal{W})$, $V_{i+1} := \bar{V}_{i+1} \setminus (\bigcup_{j=1}^i \bar{V}_j) \in \mathcal{B}_{\mathcal{B}}(\mathcal{W})$, $i = 1, \dots, \bar{n}$. Let $V_{\bar{n}+2} := \mathcal{W} \setminus (\bigcup_{j=1}^{\bar{n}+1} \bar{V}_j) \in \mathcal{B}_{\mathcal{B}}(\mathcal{W})$. Clearly, we may form a representation $\bar{R} := \{(\bar{w}_i, V_i) \mid i = 1, \dots, \bar{n} + 2, V_i \neq \emptyset\} \in \mathcal{J}(\mathcal{W})$, where $\bar{w}_i \in V_i$ are any vectors that satisfy the assumption of Proposition 11.69. $\forall R \in \mathcal{J}(\mathcal{W})$ with $\bar{R} \preceq R$. Let $R := \{(\hat{w}_\gamma, \hat{U}_\gamma) \mid \gamma \in \Gamma\}$. $\forall \gamma \in \Gamma$, by $\bar{R} \preceq R$, $\exists! i_\gamma \in \{1, \dots, \bar{n} + 2\}$ such that $\hat{U}_\gamma \subseteq V_{i_\gamma}$. Define $\bar{\Gamma} := \{\gamma \in \Gamma \mid i_\gamma = \bar{n} + 2\}$. Let $A_{\gamma, i} := A_i \cap f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1) \in \mathcal{B}$, $i = 1, \dots, \bar{n} + 1$. Let $\bar{A}_\gamma := f_{\text{inv}}(\hat{U}_\gamma) \cap E_1 \setminus E$.

Claim 11.126.1 $\forall \gamma \in \bar{\Gamma}$, $f_{\text{inv}}(\hat{U}_\gamma) \subseteq E_1$; $\forall \gamma \in \Gamma_s := \{\gamma \in \Gamma \mid \bar{A}_\gamma \neq \emptyset\}$, $\|\hat{w}_\gamma \mu(\bar{A}_\gamma)\| \leq (1 + \inf_{x \in \bar{A}_\gamma} g(x)) \mathcal{P} \circ \mu(\bar{A}_\gamma)$; and $\left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| < \epsilon/3$.

Proof of claim: Fix any $\gamma \in \bar{\Gamma}$. $\hat{U}_\gamma \subseteq V_{\bar{n}+2} = \mathcal{W} \setminus (\bigcup_{j=1}^{\bar{n}+1} \bar{V}_j)$. Then, $\mathcal{B} \ni f_{\text{inv}}(\hat{U}_\gamma) \subseteq E_1$.

$\forall \gamma \in \Gamma_s$, $\bar{A}_\gamma = f_{\text{inv}}(\hat{U}_\gamma) \cap (E_1 \setminus E) \neq \emptyset$. $\forall x \in \bar{A}_\gamma$, we have $f(x) \in \hat{U}_\gamma$, $\lim_{m \in \mathbb{N}} \phi_m(x) = f(x)$, $\|\phi_m(x)\| \leq g_m(x)$, $\forall m \in \mathbb{N}$, and $\lim_{m \in \mathbb{N}} g_m(x) = g(x)$. Then, by Propositions 7.21 and 3.66, we have $\|f(x)\| = \lim_{m \in \mathbb{N}} \|\phi_m(x)\| \leq \lim_{m \in \mathbb{N}} g_m(x) = g(x)$. Then, by $R \in \mathcal{J}(\mathcal{W})$, $\|\hat{w}_\gamma\| < \inf_{w \in \hat{U}_\gamma} \|w\| + 1 \leq \|f(x)\| + 1 \leq g(x) + 1$. By the arbitrariness of x , we have $\|\hat{w}_\gamma\| \leq \inf_{x \in \bar{A}_\gamma} g(x) + 1$. Therefore, $\|\hat{w}_\gamma \mu(\bar{A}_\gamma)\| \leq (1 + \inf_{x \in \bar{A}_\gamma} g(x)) \|\mu(\bar{A}_\gamma)\| \leq (1 + \inf_{x \in \bar{A}_\gamma} g(x)) \mathcal{P} \circ \mu(\bar{A}_\gamma)$.

Note that

$$\left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| = \left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma (\mu(\bar{A}_\gamma) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1 \cap E)) \right\|$$

$$\begin{aligned}
&\leq \left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(\bar{A}_\gamma) \right\| + \sum_{\gamma \in \Gamma} \|\hat{w}_\gamma\| \left\| \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1 \cap E) \right\| \\
&\leq \left\| \sum_{\gamma \in \Gamma_s} \hat{w}_\gamma \mu(\bar{A}_\gamma) \right\| + \sum_{\gamma \in \Gamma} \|\hat{w}_\gamma\| \mathcal{P} \circ \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1 \cap E) \\
&\leq \sum_{\gamma \in \Gamma_s} (1 + \inf_{x \in \bar{A}_\gamma} g(x)) \mathcal{P} \circ \mu(\bar{A}_\gamma) \\
&= \sum_{\gamma \in \Gamma_s} \mathcal{P} \circ \mu(\bar{A}_\gamma) + \sum_{\gamma \in \Gamma_s} (\inf_{x \in \bar{A}_\gamma} g(x)) \mathcal{P} \circ \mu(\bar{A}_\gamma)
\end{aligned}$$

The simple function $\psi : X \rightarrow [0, \infty) \subset \mathbb{R}$, defined by $\psi(x) = \sum_{\gamma \in \Gamma_s} (\inf_{\bar{x} \in \bar{A}_\gamma} g(\bar{x})) \chi_{\bar{A}_\gamma, X}(x)$, $\forall x \in X$, satisfies $0 \leq \psi(x) \leq g(x)$, $\forall x \in X$ and $\psi(x) = 0$, $\forall x \in X \setminus E_1$. Then, $\sum_{\gamma \in \Gamma_s} (\inf_{x \in \bar{A}_\gamma} g(x)) \mathcal{P} \circ \mu(\bar{A}_\gamma) = \int_X \psi d\mathcal{P} \circ \mu = \int_{E_1} \psi d\mathcal{P} \circ \mu \leq \int_{E_1} g d\mathcal{P} \circ \mu < \epsilon/6$, where the equalities follow from Proposition 11.75; and the first inequality follows from Proposition 11.83. Then, $\left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| < \mathcal{P} \circ \mu(E_1) + \epsilon/6 < \delta + \epsilon/6 \leq \epsilon/3$.

This completes the proof of the claim. $\square$

Claim 11.126.2 $\left\| \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right\| < \epsilon/6$.

Proof of claim: Since $\|\phi_n(x)\| \leq g_n(x)$, $\forall x \in X$, then $\|w_i\| \leq g_n(x)$, $\forall x \in A_i$, $i = 1, \dots, \bar{n} + 1$. Then, by Definition 11.79, $\left\| \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right\| \leq \sum_{i=1}^{\bar{n}+1} \|w_i\| \|\mu(A_i \cap E_1)\| \leq \sum_{i=1}^{\bar{n}+1} \|w_i\| \mathcal{P} \circ \mu(A_i \cap E_1) \leq \int_{E_1} g_n d\mathcal{P} \circ \mu < \epsilon/6$. $\square$

Claim 11.126.3 $\forall \gamma \in \Gamma \setminus \bar{\Gamma}$, $\forall i \in \{1, \dots, \bar{n} + 1\}$, we have $\|\hat{w}_\gamma - w_i\| \cdot \|\mu(A_{\gamma,i})\| \leq \frac{3\epsilon}{6\mathcal{P} \circ \mu(X) + 1} \mathcal{P} \circ \mu(A_{\gamma,i})$.

Proof of claim: $\forall \gamma \in \Gamma \setminus \bar{\Gamma}$, $\forall i \in \{1, \dots, \bar{n} + 1\}$, $\hat{U}_\gamma \subseteq V_{i_\gamma}$ where $i_\gamma \in \{1, \dots, \bar{n} + 1\}$. The result is trivial if $A_{\gamma,i} = \emptyset$. If $A_{\gamma,i} \neq \emptyset$, then $\exists x \in A_{\gamma,i} = A_i \cap f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1)$. This implies that $\phi_n(x) = w_i$ and $f(x) \in \mathcal{B}_W(w_i, \bar{\epsilon}) \cap \hat{U}_\gamma \subseteq \mathcal{B}_W(w_i, \bar{\epsilon}) \cap V_{i_\gamma} \subseteq \mathcal{B}_W(w_i, \bar{\epsilon}) \cap \bar{V}_{i_\gamma} = \mathcal{B}_W(w_i, \bar{\epsilon}) \cap \mathcal{B}_W(w_{i_\gamma}, \bar{\epsilon})$. Note that $\hat{w}_\gamma \in \hat{U}_\gamma \subseteq \mathcal{B}_W(w_{i_\gamma}, \bar{\epsilon})$. Therefore, we have $\|\hat{w}_\gamma - w_i\| \leq \|\hat{w}_\gamma - w_{i_\gamma}\| + \|w_{i_\gamma} - f(x)\| + \|f(x) - w_i\| < 3\bar{\epsilon} = \frac{3\epsilon}{6\mathcal{P} \circ \mu(X) + 1}$. Hence, the result holds. $\square$

Note that $F_R = \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma))$ and $\int_X \phi_n d\mu = \sum_{i=1}^{\bar{n}} w_i \mu(A_i) = \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i)$, by Proposition 11.125. Then,

$$\left\| F_R - \int_X \phi_n d\mu \right\| = \left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma)) - \sum_{i=1}^{\bar{n}} w_i \mu(A_i) \right\|$$

$$\begin{aligned}
&= \left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma (\mu(f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1)) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1)) \right. \\
&\quad \left. - \sum_{i=1}^{\bar{n}+1} w_i (\mu(A_i \cap E_1) + \mu(A_i \cap (X \setminus E_1))) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma \left(\sum_{i=1}^{\bar{n}+1} \mu(A_{\gamma,i}) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right) - \sum_{i=1}^{\bar{n}+1} w_i (\mu(A_i \cap E_1) \right. \\
&\quad \left. + \mu(A_i \cap (X \setminus E_1))) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} \hat{w}_\gamma \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right. \\
&\quad \left. - \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) - \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} w_i \mu(A_{\gamma,i}) \right\| \\
&= \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{w}_\gamma - w_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right. \\
&\quad \left. - \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right\| \\
&< \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{w}_\gamma - w_i) \mu(A_{\gamma,i}) \right\| + \frac{\epsilon}{3} + \frac{\epsilon}{6} \\
&= \frac{\epsilon}{2} + \left\| \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} (\hat{w}_\gamma - w_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} (\hat{w}_\gamma - w_i) \mu(A_{\gamma,i}) \right\| \\
&= \frac{\epsilon}{2} + \left\| \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} (\hat{w}_\gamma - w_i) \mu(A_{\gamma,i}) \right\| \\
&\leq \frac{\epsilon}{2} + \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} \|\hat{w}_\gamma - w_i\| \|\mu(A_{\gamma,i})\| \\
&\leq \frac{\epsilon}{2} + \sum_{\gamma \in \Gamma \setminus \bar{\Gamma}} \sum_{i=1}^{\bar{n}+1} \frac{3\epsilon}{6\mathcal{P} \circ \mu(X) + 1} \mathcal{P} \circ \mu(A_{\gamma,i}) \leq \frac{\epsilon}{2} + \frac{3\mathcal{P} \circ \mu(X)\epsilon}{6\mathcal{P} \circ \mu(X) + 1} < \epsilon
\end{aligned}$$

where the first inequality follows from Claims 11.126.1 and 11.126.2; the seventh equality follows from Claim 11.126.1 and the fact that $A_{\gamma,i} = \emptyset$, $\forall \gamma \in \bar{\Gamma}$, $\forall i \in \{1, \dots, \bar{n}+1\}$; and the third inequality follows from Claim 11.126.3.

In summary, we have shown that $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_1 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq n_1$, $\exists \bar{R} \in \mathcal{I}(\mathcal{W})$ such that $\forall R \in \mathcal{I}(\mathcal{W})$ with $\bar{R} \preceq R$, we have

$\|F_R - \int_X \phi_n d\mu\| < \epsilon$. Then, the net $(F_R)_{R \in \mathcal{J}(\mathcal{W})} \subseteq \mathcal{Z}$ is a Cauchy net. By Proposition 4.44, the net admits a limit $\int_X f d\mu \in \mathcal{Z}$. By Propositions 7.21, 7.23, 3.67, and 3.66, we have $\|\int_X f d\mu - \int_X \phi_n d\mu\| = \lim_{R \in \mathcal{J}(\mathcal{W})} \|F_R - \int_X \phi_n d\mu\| \leq \epsilon$, $\forall n \geq n_1$. Hence, we have $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n d\mu$ and f is integrable over $\mathcal{X}$. By Proposition 11.125, $\int_X \phi_n d\mu \in \mathcal{W}$, $\forall n \in \mathbb{N}$. Then, by the completeness of $\bar{\mathcal{W}}$, we have $\int_X f d\mu \in \bar{\mathcal{W}}$.

Note that $0 \leq \|\int_X f d\mu\| = \lim_{n \in \mathbb{N}} \|\int_X \phi_n d\mu\| \leq \limsup_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \phi_n d\mathcal{P} \circ \mu \leq \lim_{n \in \mathbb{N}} \int_X g_n d\mathcal{P} \circ \mu = \int_X g d\mathcal{P} \circ \mu < +\infty$, where the first equality follows from Propositions 7.21 and 3.66; the second inequality follows from Proposition 11.125; and the third inequality follows from Proposition 11.83. This completes the proof of the lemma. $\square$

Lemma 11.127 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, and $g : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. Assume that $\|f(x)\| \leq g(x)$, $\forall x \in X$, and g is integrable over $\bar{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$. Then, there exists a sequence of simple functions $(\varphi_k)_{k=1}^\infty$, $\varphi_k : X \rightarrow \mathcal{W}$, $\forall k \in \mathbb{N}$, such that $\lim_{k \in \mathbb{N}} \varphi_k = f$ a.e. in $\mathcal{X}$, $\|\varphi_k(x)\| \leq g(x)$, $\forall x \in X$, $\forall k \in \mathbb{N}$, f is integrable over $\mathcal{X}$ with $\int_X f d\mu = \lim_{k \in \mathbb{N}} \int_X \varphi_k d\mu \in \mathcal{Z}$, $\lim_{k \in \mathbb{N}} \int_X \mathcal{P} \circ \varphi_k d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu \in [0, \infty) \subset \mathbb{R}$, and $\lim_{k \in \mathbb{N}} \int_X \mathcal{P} \circ (\varphi_k - f) d\mathcal{P} \circ \mu = 0$.*

Proof By Proposition 11.66, there exists a sequence of simple functions $(\varphi_k)_{k=1}^\infty$, $\varphi_k : X \rightarrow \mathcal{W}$, $\forall k \in \mathbb{N}$, such that $\|\varphi_k(x)\| \leq \|f(x)\| \leq g(x)$, $\forall x \in X$, $\forall k \in \mathbb{N}$, and $\lim_{k \in \mathbb{N}} \varphi_k = f$ a.e. in $\mathcal{X}$. By Lemma 11.126, f is integrable over $\mathcal{X}$ and $\int_X f d\mu = \lim_{k \in \mathbb{N}} \int_X \varphi_k d\mu \in \mathcal{Z}$.

Note that $0 \leq \mathcal{P} \circ \varphi_k(x) \leq g(x)$, $\forall x \in X$, $\forall k \in \mathbb{N}$, $\mathcal{P} \circ f$ is $\mathcal{B}$ -measurable, and $\lim_{k \in \mathbb{N}} \mathcal{P} \circ \varphi_k = \mathcal{P} \circ f$ a.e. in $\mathcal{X}$, by Propositions 7.21 and 11.52. Then, by Lebesgue Dominated Convergence Theorem 11.91, we have $0 \leq \lim_{k \in \mathbb{N}} \int_X \mathcal{P} \circ \varphi_k d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu \leq \int_X g d\mathcal{P} \circ \mu < +\infty$, where the second inequality follows from Proposition 11.83. By Lemma 11.43, $\mathcal{P} \circ (\varphi_k - f)$ is $\mathcal{B}$ -measurable. By Propositions 7.21, 7.23, 11.52, and 11.53, $\lim_{k \in \mathbb{N}} \mathcal{P} \circ (\varphi_k - f) = 0$ a.e. in $\mathcal{X}$. Note that $0 \leq \mathcal{P} \circ (\varphi_k - f)(x) \leq 2g(x)$, $\forall x \in X$, $\forall k \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91 and Proposition 11.75, we have $\lim_{k \in \mathbb{N}} \int_X \mathcal{P} \circ (\varphi_k - f) d\mathcal{P} \circ \mu = 0$. This completes the proof of the lemma. $\square$

Theorem 11.128 (Lebesgue Dominated Convergence) *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, $\bar{\mathcal{W}} := \text{span}(\{z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \ni z = wy\}) \subseteq \mathcal{Z}$ be the Banach subspace of $\mathcal{Z}$, $f_n : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $g_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $g : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. Assume that*

- (i) $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$, $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$;

$$(ii) \|f_n(x)\| \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N};$$

$$(iii) \int_X g_n d\mathcal{P} \circ \mu \in \mathbb{R}, \forall n \in \mathbb{N}, \text{ and } \int_X g d\mathcal{P} \circ \mu = \lim_{n \in \mathbb{N}} \int_X g_n d\mathcal{P} \circ \mu \in \mathbb{R}.$$

Then, f is integrable over $\mathcal{X}$, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu \in \bar{\mathcal{W}} \subseteq \mathcal{Z}$, and $0 \leq \|\int_X f d\mu\| \leq \int_X g d\mathcal{P} \circ \mu < +\infty$.

Proof $\forall n \in \mathbb{N}$, by Lemma 11.127, there exists a sequence of simple functions $(\bar{\varphi}_{n,k})_{k=1}^\infty$, $\bar{\varphi}_{n,k} : X \rightarrow \mathcal{W}$, $\forall k \in \mathbb{N}$, such that $\lim_{k \in \mathbb{N}} \bar{\varphi}_{n,k} = f_n$ a.e. in $\mathcal{X}$, $\|\bar{\varphi}_{n,k}(x)\| \leq g_n(x)$, $\forall x \in X$, $\forall k \in \mathbb{N}$, f_n is integrable over $\mathcal{X}$, and $\int_X f_n d\mu = \lim_{k \in \mathbb{N}} \int_X \bar{\varphi}_{n,k} d\mu \in \mathcal{Z}$. By Lemma 11.126, $\int_X f_n d\mu \in \bar{\mathcal{W}}$, $\forall n \in \mathbb{N}$. By Proposition 11.58, $\lim_{k \in \mathbb{N}} \bar{\varphi}_{n,k} = f_n$ in measure in $\mathcal{X}$. Then, $\exists k_n \in \mathbb{N}$ such that $\mathcal{P} \circ \mu(E_n) := \mathcal{P} \circ \mu(\{x \in X \mid \|\bar{\varphi}_{n,k_n}(x) - f_n(x)\| \geq 2^{-n}\}) < 2^{-n}$ and $\|\int_X f_n d\mu - \int_X \bar{\varphi}_{n,k_n} d\mu\| < 2^{-n}$. Denote $\psi_n := \bar{\varphi}_{n,k_n}$.

By Proposition 11.58, $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have $2^{-n} < \epsilon/2$ and $\mathcal{P} \circ \mu(\bar{E}_n) := \mathcal{P} \circ \mu(\{x \in X \mid \|f_n(x) - f(x)\| \geq \epsilon/2\}) < \epsilon/2$. Note that, by Lemma 11.43, $\mathcal{B} \ni F_n := \{x \in X \mid \|\psi_n(x) - f(x)\| \geq \epsilon\} \subseteq \{x \in X \mid \|\psi_n(x) - f_n(x)\| + \|f_n(x) - f(x)\| \geq \epsilon\} \subseteq E_n \cup \bar{E}_n$. Then, we have $\mathcal{P} \circ \mu(F_n) \leq \mathcal{P} \circ \mu(E_n) + \mathcal{P} \circ \mu(\bar{E}_n) < \epsilon$. This implies that $\lim_{n \in \mathbb{N}} \psi_n = f$ in measure in $\mathcal{X}$. By Proposition 11.57, there exists a subsequence $(\psi_{n_i})_{i=1}^\infty$ of $(\psi_n)_{n=1}^\infty$ that converges to f a.e. in $\mathcal{X}$. By Lemma 11.126, f is integrable over $\mathcal{X}$, $\int_X f d\mu = \lim_{i \in \mathbb{N}} \int_X \psi_{n_i} d\mu \in \mathcal{Z}$, and $0 \leq \|\int_X f d\mu\| \leq \int_X g d\mathcal{P} \circ \mu < +\infty$. By the fact that $\|\int_X f_{n_i} d\mu - \int_X \psi_{n_i} d\mu\| < 2^{-n_i}$, $\forall i \in \mathbb{N}$, then we have $\lim_{i \in \mathbb{N}} \int_X f_{n_i} d\mu = \lim_{i \in \mathbb{N}} \int_X \psi_{n_i} d\mu = \int_X f d\mu \in \mathcal{Z}$. So far, we have shown that, for the sequence $(f_n)_{n=1}^\infty$, there exists a subsequence $(f_{n_i})_{i=1}^\infty$ such that $\lim_{i \in \mathbb{N}} \int_X f_{n_i} d\mu = \int_X f d\mu \in \mathcal{Z}$. This then holds for any subsequence of $(f_n)_{n=1}^\infty$. By Proposition 3.71, we have $\lim_{n \in \mathbb{N}} \int_X f_n d\mu = \int_X f d\mu$. By the completeness of $\bar{\mathcal{W}}$, we have $\int_X f d\mu \in \bar{\mathcal{W}} \subseteq \mathcal{Z}$.

This completes the proof of the theorem. $\square$

Definition 11.129 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $\mathcal{W} := \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, and $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable. f is said to be absolutely integrable if $\mathcal{P} \circ f$ is integrable over $\bar{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$, that is, $0 \leq \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu < +\infty$.

Proposition 11.130 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\bar{\mathcal{W}}$ be a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $\mathcal{U}$ be a Banach space over $\mathbb{K}$, $\bar{\mathcal{W}} := \overline{\text{span}(\{z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \ni z = wy\})} \subseteq \mathcal{Z}$ be the Banach subspace of $\mathcal{Z}$, $f_i : X \rightarrow \mathcal{W}$ be absolutely integrable over $\mathcal{X}$, $i = 1, 2$. Then, the following statements hold.

$$(i) f_i \text{ is integrable over } \mathcal{X} \text{ and } \int_X f_i d\mu \in \bar{\mathcal{W}} \subseteq \mathcal{Z}, i = 1, 2.$$

- (ii) $f_1 + f_2$ is absolutely integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu \in \bar{\mathcal{W}}$.
- (iii) $\forall A \in \mathcal{B}(\mathcal{Z}, \mathcal{U})$, Af_1 is absolutely integrable over $\mathcal{X}$ and $\int_X (Af_1) d\mu = A \int_X f_1 d\mu \in \mathcal{U}$.
- (iv) $\forall c \in \mathbb{K}$, cf_1 is absolutely integrable over $\mathcal{X}$ and $\int_X (cf_1) d\mu = c \int_X f_1 d\mu \in \bar{\mathcal{W}}$.
- (v) $\forall H \in \mathcal{B}$, let $\mathcal{H} := (H, \mathcal{B}_H, \mu_H)$ be the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$ as defined in Proposition 11.115. Then, $f_1|_H$ is absolutely integrable over $\mathcal{H}$ and $\int_X (f_1 \chi_{H, X}) d\mu = \int_H f_1|_H d\mu_H \in \bar{\mathcal{W}}$. We will henceforth denote $\int_H f_1|_H d\mu_H$ by $\int_H f_1 d\mu$.
- (vi) If $f_1 = f_2$ a.e. in $\mathcal{X}$ then $\int_X f_1 d\mu = \int_X f_2 d\mu \in \bar{\mathcal{W}}$.
- (vii) $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, $\sum_{i=1}^\infty \int_{E_i} f_1 d\mu = \int_{\bigcup_{i=1}^\infty E_i} f_1 d\mu \in \bar{\mathcal{W}}$.
- (viii) $0 \leq \|\int_X f_1 d\mu\| \leq \int_X \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu < +\infty$.

Proof (i) $\forall i \in \{1, 2\}$, by Lemma 11.127, there exists a sequence of simple functions $(\phi_{i,n})_{n=1}^\infty$, $\phi_{i,n} : X \rightarrow \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\lim_{n \in \mathbb{N}} \phi_{i,n} = f_i$ a.e. in $\mathcal{X}$, $\|\phi_{i,n}(x)\| \leq \mathcal{P} \circ f_i(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, f_i is integrable over $\mathcal{X}$, and $\int_X f_i d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_{i,n} d\mu \in \mathcal{Z}$. By Lebesgue Dominated Convergence Theorem 11.128, $\int_X f_i d\mu \in \bar{\mathcal{W}}$, $i = 1, 2$.

(ii) By Propositions 11.38, 11.39, 7.21, and 7.23, $f_1 + f_2$ and $\mathcal{P} \circ (f_1 + f_2)$ are $\mathcal{B}$ -measurable. Note that $\mathcal{P} \circ (f_1 + f_2)(x) = \|f_1(x) + f_2(x)\| \leq \|f_1(x)\| + \|f_2(x)\| = (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x)$, $\forall x \in X$. By Proposition 11.83, we have $0 \leq \int_X \mathcal{P} \circ (f_1 + f_2) d\mathcal{P} \circ \mu \leq \int_X (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2) d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu + \int_X \mathcal{P} \circ f_2 d\mathcal{P} \circ \mu < +\infty$. Hence, $f_1 + f_2$ is absolutely integrable over $\mathcal{X}$.

Note that $\phi_{1,n} + \phi_{2,n} : X \rightarrow \mathcal{W}$, $f_1 + f_2 : X \rightarrow \mathcal{W}$, and $\|\phi_{1,n}(x) + \phi_{2,n}(x)\| \leq \|\phi_{1,n}(x)\| + \|\phi_{2,n}(x)\| \leq (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Propositions 7.23, 11.52, and 11.53, $\lim_{n \in \mathbb{N}} (\phi_{1,n} + \phi_{2,n}) = f_1 + f_2$ a.e. in $\mathcal{X}$. By Lebesgue Dominated Convergence Theorem 11.128, $f_1 + f_2$ is integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) d\mu = \lim_{n \in \mathbb{N}} \int_X (\phi_{1,n} + \phi_{2,n}) d\mu = \lim_{n \in \mathbb{N}} (\int_X \phi_{1,n} d\mu + \int_X \phi_{2,n} d\mu) = \int_X f_1 d\mu + \int_X f_2 d\mu \in \bar{\mathcal{W}}$, where the second equality follows from Proposition 11.125; and the third equality follows from Propositions 3.66, 3.67, and 7.23.

(iii) Fix any $A \in \mathcal{B}(\mathcal{Z}, \mathcal{U})$. By Propositions 11.38, 7.21, and 7.62, $A\phi_{1,n}$, Af_1 , and $\mathcal{P} \circ (Af_1)$ are $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. Note that $\mathcal{P} \circ (Af_1)(x) = \|Af_1(x)\| \leq \|A\| \|f_1(x)\| = \|A\| \mathcal{P} \circ f_1(x)$, $\forall x \in X$. By Proposition 11.83 and the fact that $0 \leq \int_X \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu < +\infty$, we have $0 \leq \int_X \mathcal{P} \circ (Af_1) d\mathcal{P} \circ \mu \leq \int_X (\|A\| \mathcal{P} \circ f_1) d\mathcal{P} \circ \mu = \|A\| \int_X \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu < +\infty$. Hence, Af_1 is absolutely integrable over $\mathcal{X}$.

Note that $Af_1 : X \rightarrow \hat{\mathcal{W}}$ and $A\phi_{1,n} : X \rightarrow \hat{\mathcal{W}}$, $\forall n \in \mathbb{N}$, where $\hat{\mathcal{W}} := \{F \in \mathcal{B}(\mathcal{Y}, \mathcal{U}) \mid F = AF_1, F_1 \in \mathcal{W}\}$ is a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{U})$.

By Propositions 11.52 and 7.62, $\lim_{n \in \mathbb{N}} A\phi_{1,n} = Af_1$ a.e. in $\mathcal{X}$. By Proposition 11.125, we have $\int_X (A\phi_{1,n}) d\mu = A \int_X \phi_{1,n} d\mu \in \mathcal{U}$, $\forall n \in \mathbb{N}$. Note also that $\|A\phi_{1,n}(x)\| \leq \|A\| \|\phi_{1,n}(x)\| \leq \|A\| \mathcal{P} \circ f_1(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.128, Af_1 is integrable over $\mathcal{X}$ and $\int_X (Af_1) d\mu = \lim_{n \in \mathbb{N}} \int_X (A\phi_{1,n}) d\mu = \lim_{n \in \mathbb{N}} A \int_X \phi_{1,n} d\mu = A \int_X f_1 d\mu \in \mathcal{U}$, where the last equality follows from Propositions 3.66 and 7.62.

(iv) This follows immediately from (iii).

(v) Fix any $H \in \mathcal{B}$. Note that $(\mathcal{P} \circ f_1)|_H = \mathcal{P} \circ (f_1|_H)$. Then, by Proposition 11.115, we have $0 \leq \int_H \mathcal{P} \circ (f_1|_H) d\mathcal{P} \circ \mu_H = \int_H (\mathcal{P} \circ f_1)|_H d(\mathcal{P} \circ \mu)|_{\mathcal{B}_H} = \int_H \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu \leq \int_X \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu < +\infty$, where the second equality and the second inequality follow from Proposition 11.83. Hence, $f_1|_H$ is absolutely integrable over $\mathcal{H}$.

$(\phi_{1,n}|_H)_{n=1}^\infty$ is a sequence of simple functions that converges to $f_1|_H$ a.e. in $\mathcal{H}$. Note that $\|\phi_{1,n}|_H(x)\| \leq (\mathcal{P} \circ f_1)|_H(x) = \mathcal{P} \circ (f_1|_H)(x)$, $\forall x \in H$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.128, we have $\int_H f_1|_H d\mu_H = \lim_{n \in \mathbb{N}} \int_H \phi_{1,n}|_H d\mu_H \in \bar{\mathcal{W}}$. Note that $(\phi_{1,n}\chi_{H,X})_{n=1}^\infty$ is a sequence of simple functions that converges to $f_1\chi_{H,X}$ a.e. in $\mathcal{X}$, and $\|\phi_{1,n}(x)\chi_{H,X}(x)\| \leq \mathcal{P} \circ f_1(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Again by Lebesgue Dominated Convergence Theorem 11.128, we have $\int_X (f_1\chi_{H,X}) d\mu = \lim_{n \in \mathbb{N}} \int_X (\phi_{1,n}\chi_{H,X}) d\mu = \lim_{n \in \mathbb{N}} \int_H \phi_{1,n}|_H d\mu_H = \int_H f_1|_H d\mu_H \in \bar{\mathcal{W}}$, where the second equality follows from Proposition 11.125.

(vi) If $f_1 = f_2$ a.e. in $\mathcal{X}$, then, by Proposition 11.54, $\lim_{n \in \mathbb{N}} \phi_{1,n} = f_2$ a.e. in $\mathcal{X}$. By Lebesgue Dominated Convergence Theorem 11.128, we have $\int_X f_2 d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_{1,n} d\mu = \int_X f_1 d\mu$.

(vii) $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$ and $\bar{E}_n := \bigcup_{i=1}^n E_i \in \mathcal{B}$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} f_1(x)\chi_{\bar{E}_n,X}(x) = f_1(x)\chi_{E,X}(x)$, $\forall x \in X$. Then, we have $\sum_{i=1}^\infty \int_{E_i} f_1 d\mu = \lim_{n \in \mathbb{N}} \sum_{i=1}^n \int_X (f_1\chi_{E_i,X}) d\mu = \lim_{n \in \mathbb{N}} \int_X (f_1\chi_{\bar{E}_n,X}) d\mu = \int_X (f_1\chi_{E,X}) d\mu = \int_E f_1 d\mu \in \bar{\mathcal{W}}$, where the first equality follows from (v); the second equality follows from (ii); the third equality follows from Lebesgue Dominated Convergence Theorem 11.128; the last equality follows from (v); and the last inclusion follows from (v).

(viii) The result follows immediately from Lebesgue Dominated Convergence Theorem 11.128.

This completes the proof of the proposition. $\square$

Theorem 11.131 (Lebesgue Dominated Convergence) *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, $\bar{\mathcal{W}} := \text{span}(\{z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \ni z = wy\}) \subseteq \mathcal{Z}$ be the Banach subspace of $\mathcal{Z}$, $f_n : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $g_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, and $g : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. Assume that*

(i) $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$, $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$;

$$(ii) \|f_n(x)\| \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N};$$

$$(iii) \int_X g_n d\mathcal{P} \circ \mu \in \mathbb{R}, \forall n \in \mathbb{N} \text{ and } \int_X g d\mathcal{P} \circ \mu = \lim_{n \in \mathbb{N}} \int_X g_n d\mathcal{P} \circ \mu \in \mathbb{R}.$$

Then, f is integrable over $\mathcal{X}$, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu \in \bar{\mathcal{W}} \subseteq \mathcal{Z}$, and $0 \leq \|\int_X f d\mu\| \leq \int_X g d\mathcal{P} \circ \mu < +\infty$. Furthermore, f is absolutely integrable over $\mathcal{X}$.

Proof We will show that f is integrable over $\mathcal{X}$, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu \in \bar{\mathcal{W}}$, and $0 \leq \|\int_X f d\mu\| \leq \int_X g d\mathcal{P} \circ \mu < +\infty$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\mathcal{P} \circ \mu(X) < +\infty$; Case 2: $\mathcal{P} \circ \mu(X) = +\infty$.

Case 1: $\mathcal{P} \circ \mu(X) < +\infty$. Then, $\mathcal{X}$ is a finite $\mathcal{Y}$ -valued measure space. By Lebesgue Dominated Convergence Theorem 11.128, f is integrable over $\mathcal{X}$, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu \in \bar{\mathcal{W}}$, and $0 \leq \|\int_X f d\mu\| \leq \int_X g d\mathcal{P} \circ \mu < +\infty$.

Case 2: $\mathcal{P} \circ \mu(X) = +\infty$. Let $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ and $(F_{n,A})_{A \in \mathcal{M}(\mathcal{X})}$ be the nets for the integrals $\int_X f d\mu$ and $\int_X f_n d\mu$ as defined in Definition 11.120, respectively, $\forall n \in \mathbb{N}$. $\forall A \in \mathcal{M}(\mathcal{X})$, we have $A \in \mathcal{B}$ and $\mathcal{P} \circ \mu(A) < \infty$. Let $\mathcal{A} := (A, \mathcal{B}_A, \mu_A)$ be the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$ as defined in Proposition 11.115. By Proposition 11.83, $g|_A$ and $g_n|_A$'s are integrable over $\mathcal{A} := (A, \mathcal{B}_A, \mathcal{P} \circ \mu_A)$, which is the finite measure subspace of $\mathcal{X} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$ as shown in Proposition 11.115. By Proposition 11.83, $\lim_{n \in \mathbb{N}} g_n|_A = g|_A$ a.e. in $\mathcal{A}$ and $\int_A g|_A d\mathcal{P} \circ \mu_A = \lim_{n \in \mathbb{N}} \int_A g_n|_A d\mathcal{P} \circ \mu_A \in \mathbb{R}$. By Lebesgue Dominated Convergence Theorem 11.128, we have $f|_A$ and $f_n|_A$'s are integrable over $\mathcal{A}$, $\int_A f|_A d\mu_A = \lim_{n \in \mathbb{N}} \int_A f_n|_A d\mu_A \in \bar{\mathcal{W}}$, and $0 \leq \|\int_A f|_A d\mu_A\| \leq \int_A g|_A d\mathcal{P} \circ \mu_A < +\infty$. Note that $F_A = \int_A f|_A d\mu_A \in \bar{\mathcal{W}}$ and $F_{n,A} = \int_A f_n|_A d\mu_A \in \bar{\mathcal{W}}$, $\forall n \in \mathbb{N}$. Then, the nets $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ and $(F_{n,A})_{A \in \mathcal{M}(\mathcal{X})}$'s are well defined.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Definition 11.79, $\exists$ a simple function $\phi : X \rightarrow [0, \infty) \subset \mathbb{R}$ such that $0 \leq \phi(x) \leq g(x)$, $\forall x \in X$, and $\int_X g d\mathcal{P} \circ \mu - \epsilon/5 < \int_X \phi d\mathcal{P} \circ \mu \leq \int_X g d\mathcal{P} \circ \mu < +\infty$. Let $A_0 := \{x \in X \mid \phi(x) > 0\}$. Then, $A_0 \in \mathcal{B}$, $\mathcal{P} \circ \mu(A_0) < +\infty$, and $A_0 \in \mathcal{M}(\mathcal{X})$. $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have $0 \leq \int_{A \setminus A_0} g|_{A \setminus A_0} d\mathcal{P} \circ \mu_{A \setminus A_0} = \int_A g|_A d\mathcal{P} \circ \mu_A - \int_{A_0} g|_{A_0} d\mathcal{P} \circ \mu_{A_0} \leq \int_X g d\mathcal{P} \circ \mu - \int_{A_0} \phi|_{A_0} d\mathcal{P} \circ \mu_{A_0} = \int_X g d\mathcal{P} \circ \mu - \int_X \phi d\mathcal{P} \circ \mu < \epsilon/5$, where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Note that $\|F_A - F_{A_0}\| = \left\| \int_A f|_A d\mu_A - \int_{A_0} f|_{A_0} d\mu_{A_0} \right\| = \left\| \int_{A \setminus A_0} f|_{A \setminus A_0} d\mu_{A \setminus A_0} \right\| \leq \int_{A \setminus A_0} g|_{A \setminus A_0} d\mathcal{P} \circ \mu_{A \setminus A_0} < \epsilon/5$, where the second equality and the first inequality follow from Proposition 11.130. Then, the net $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ is a Cauchy net, which admits a limit, by Proposition 4.44, $\int_X f d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_A \in \bar{\mathcal{W}}$. Hence, f is integrable over $\mathcal{X}$. By Propositions 7.21 and 3.66, $0 \leq \|\int_X f d\mu\| = \lim_{A \in \mathcal{M}(\mathcal{X})} \|F_A\| \leq \lim_{A \in \mathcal{M}(\mathcal{X})} \int_A g|_A d\mathcal{P} \circ \mu_A = \int_X g d\mathcal{P} \circ \mu < +\infty$.

$\forall n \in \mathbb{N}$, by Definition 11.79, $\exists$ a simple function $\phi_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ such that $0 \leq \phi_n(x) \leq g_n(x)$, $\forall x \in X$, and $\int_X g_n d\mathcal{P} \circ \mu - \epsilon/5 < \int_X \phi_n d\mathcal{P} \circ \mu \leq \int_X g_n d\mathcal{P} \circ \mu < +\infty$. Let $A_n := \{x \in X \mid \phi_n(x) > 0\}$. Then, $A_n \in \mathcal{B}$, $\mathcal{P} \circ \mu(A_n) < +\infty$, and $A_n \in \mathcal{M}(\mathcal{X})$. $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_n \subseteq A$, we have $0 \leq \int_{A \setminus A_n} g_n|_{A \setminus A_n} d\mathcal{P} \circ \mu_{A \setminus A_n} = \int_A g_n|_A d\mathcal{P} \circ \mu_A - \int_{A_n} g_n|_{A_n} d\mathcal{P} \circ \mu_{A_n} \leq \int_X g_n d\mathcal{P} \circ \mu - \int_{A_n} \phi_n|_{A_n} d\mathcal{P} \circ \mu_{A_n} = \int_X g_n d\mathcal{P} \circ \mu - \int_X \phi_n d\mathcal{P} \circ \mu < \epsilon/5$, where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Note that $\|F_{n,A} - F_{n,A_n}\| = \left\| \int_A f_n|_A d\mu_A - \int_{A_n} f_n|_{A_n} d\mu_{A_n} \right\| = \left\| \int_{A \setminus A_n} f_n|_{A \setminus A_n} d\mu_{A \setminus A_n} \right\| \leq \int_{A \setminus A_n} g_n|_{A \setminus A_n} d\mathcal{P} \circ \mu_{A \setminus A_n} < \epsilon/5$, where the second equality and the first inequality follow from Proposition 11.130. Then, the net $(F_{n,A})_{A \in \mathcal{M}(\mathcal{X})}$ is a Cauchy net, which admits a limit, by Proposition 4.44, $\int_X f_n d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_{n,A} \in \bar{\mathcal{W}}$.

Let $\bar{g}_n := g_n \wedge g$, $\forall n \in \mathbb{N}$. Then, $0 \leq \bar{g}_n(x) \leq g(x)$, $\forall x \in X$, and $\bar{g}_n$ is $\mathcal{B}$ -measurable by Proposition 11.40, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} \bar{g}_n = g$ a.e. in $\mathcal{X}$ by Proposition 11.50. By Fatou's Lemma 11.80, $\int_X g d\mathcal{P} \circ \mu \leq \liminf_{n \in \mathbb{N}} \int_X \bar{g}_n d\mathcal{P} \circ \mu$. By Definition 11.79, $0 \leq \int_X \bar{g}_n d\mathcal{P} \circ \mu \leq \int_X g d\mathcal{P} \circ \mu < +\infty$, $\forall n \in \mathbb{N}$. Then, we have $\int_X g d\mathcal{P} \circ \mu \leq \liminf_{n \in \mathbb{N}} \int_X \bar{g}_n d\mathcal{P} \circ \mu \leq \limsup_{n \in \mathbb{N}} \int_X \bar{g}_n d\mathcal{P} \circ \mu \leq \int_X g d\mathcal{P} \circ \mu < +\infty$. Therefore, by Proposition 3.83, $\int_X g d\mathcal{P} \circ \mu = \lim_{n \in \mathbb{N}} \int_X \bar{g}_n d\mathcal{P} \circ \mu \in \mathbb{R}$. This coupled with (iii) and the fact that $F_{A_0} = \lim_{n \in \mathbb{N}} F_{n,A_0}$, implies that $\exists n_0 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n_0 \leq n$, we have $0 \leq \int_X g d\mathcal{P} \circ \mu - \int_X \bar{g}_n d\mathcal{P} \circ \mu < \epsilon/5$, $|\int_X g d\mathcal{P} \circ \mu - \int_X g_n d\mathcal{P} \circ \mu| < \epsilon/5$, and $\|F_{A_0} - F_{n,A_0}\| < \epsilon/5$. $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have $0 \leq \int_{A \setminus A_0} g_n|_{A \setminus A_0} d\mathcal{P} \circ \mu_{A \setminus A_0} = \int_{A \setminus A_0} (g_n - \bar{g}_n)|_{A \setminus A_0} d\mathcal{P} \circ \mu_{A \setminus A_0} + \int_{A \setminus A_0} \bar{g}_n|_{A \setminus A_0} d\mathcal{P} \circ \mu_{A \setminus A_0} \leq \int_X (g_n - \bar{g}_n) d\mathcal{P} \circ \mu + \int_{A \setminus A_0} g|_{A \setminus A_0} d\mathcal{P} \circ \mu_{A \setminus A_0} < \int_X g_n d\mathcal{P} \circ \mu - \int_X \bar{g}_n d\mathcal{P} \circ \mu + \epsilon/5 \leq |\int_X g_n d\mathcal{P} \circ \mu - \int_X g d\mathcal{P} \circ \mu| + \int_X g d\mathcal{P} \circ \mu - \int_X \bar{g}_n d\mathcal{P} \circ \mu + \epsilon/5 < 3\epsilon/5$, where the first equality, the second inequality, and the third inequality follow from Proposition 11.83. Furthermore, $\|F_{n,A} - F_{n,A_0}\| = \left\| \int_A f_n|_A d\mu_A - \int_{A_0} f_n|_{A_0} d\mu_{A_0} \right\| = \left\| \int_{A \setminus A_0} f_n|_{A \setminus A_0} d\mu_{A \setminus A_0} \right\| \leq \int_{A \setminus A_0} g_n|_{A \setminus A_0} d\mathcal{P} \circ \mu_{A \setminus A_0} < 3\epsilon/5$, where the second equality and the first inequality follow from Proposition 11.130. This implies that $\|\int_X f d\mu - \int_X f_n d\mu\| \leq \|\int_X f d\mu - F_{A_0}\| + \|F_{A_0} - F_{n,A_0}\| + \|F_{n,A_0} - \int_X f_n d\mu\| < \lim_{A \in \mathcal{M}(\mathcal{X})} \|F_A - F_{A_0}\| + \epsilon/5 + \lim_{A \in \mathcal{M}(\mathcal{X})} \|F_{n,A_0} - F_{n,A}\| \leq \epsilon/5 + \epsilon/5 + 3\epsilon/5 = \epsilon$, where the second inequality follows from Propositions 7.21 and 3.66. Hence, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu \in \bar{\mathcal{W}}$.

Then, in both cases, we have shown that f is integrable over $\mathcal{X}$, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu \in \bar{\mathcal{W}}$, and $0 \leq \|\int_X f d\mu\| \leq \int_X g d\mathcal{P} \circ \mu < +\infty$.

Note that $0 \leq \mathcal{P} \circ f_n(x) \leq g_n(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \mathcal{P} \circ f_n = \mathcal{P} \circ f$ a.e. in $\mathcal{X}$ by Propositions 7.21 and 11.52, and $\mathcal{P} \circ f$ and $\mathcal{P} \circ f_n$'s are $\mathcal{B}$ -measurable by Propositions 7.21 and 11.38. Then, by Lebesgue Dominated Convergence Theorem 11.91, $\int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu = \lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ f_n d\mathcal{P} \circ \mu \in \mathbb{R}$. Hence, f is absolutely integrable over $\mathcal{X}$.

This completes the proof of the theorem. $\square$

Proposition 11.132 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, $\mathcal{U}$ be a Banach space over $\mathbb{K}$, $\bar{\mathcal{W}} := \overline{\text{span}(\{z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \ni z = wy\})} \subseteq \mathcal{Z}$ be the Banach subspace of $\mathcal{Z}$, $f_i : X \rightarrow \mathcal{W}$ be absolutely integrable over $\mathcal{X}$, $i = 1, 2$. Then, the following statements hold.*

- (i) f_i is integrable over $\mathcal{X}$ and $\int_X f_i d\mu \in \bar{\mathcal{W}} \subseteq \mathcal{Z}$, $i = 1, 2$.
- (ii) $f_1 + f_2$ is absolutely integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu \in \bar{\mathcal{W}}$.
- (iii) $\forall A \in B(\mathcal{Z}, \mathcal{U})$, Af_1 is absolutely integrable over $\mathcal{X}$ and $\int_X (Af_1) d\mu = A \int_X f_1 d\mu \in \mathcal{U}$.
- (iv) $\forall c \in \mathbb{K}$, cf_1 is absolutely integrable over $\mathcal{X}$ and $\int_X (cf_1) d\mu = c \int_X f_1 d\mu \in \bar{\mathcal{W}}$.
- (v) $\forall H \in \mathcal{B}$, let $\mathcal{H} := (H, \mathcal{B}_H, \mu_H)$ be the $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$ as defined in Proposition 11.115. Then, $f_1|_H$ is absolutely integrable over $\mathcal{H}$ and $\int_X (f_1 \chi_{H,X}) d\mu = \int_H f_1|_H d\mu_H \in \bar{\mathcal{W}}$. We will henceforth denote $\int_H f_1|_H d\mu_H$ by $\int_H f_1 d\mu$.
- (vi) If $f_1 = f_2$ a.e. in $\mathcal{X}$ then $\int_X f_1 d\mu = \int_X f_2 d\mu \in \bar{\mathcal{W}}$.
- (vii) $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, $\sum_{i=1}^\infty \int_{E_i} f_1 d\mu = \int_{\bigcup_{i=1}^\infty E_i} f_1 d\mu \in \bar{\mathcal{W}}$.
- (viii) $0 \leq \|\int_X f_1 d\mu\| \leq \int_X \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu < +\infty$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{P} \circ \mu(X) < +\infty$; Case 2: $\mathcal{P} \circ \mu(X) = +\infty$. Case 1: $\mathcal{P} \circ \mu(X) < +\infty$. The results follow immediately from Proposition 11.130.

Case 2: $\mathcal{P} \circ \mu(X) = +\infty$. $f_1, f_2, \mathcal{P} \circ f_1$, and $\mathcal{P} \circ f_2$ are $\mathcal{B}$ -measurable by Propositions 7.21 and 11.38. Let $(F_{i,E})_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X f_i d\mu$ as defined in Definition 11.120, $i = 1, 2$.

(i) $\forall E \in \mathcal{M}(\mathcal{X})$, by Proposition 11.130, $F_{i,E} = \int_E f_i|_E d\mu_E \in \bar{\mathcal{W}}$, $i = 1, 2$. Then, the net $(F_{i,E})_{E \in \mathcal{M}(\mathcal{X})}$ is well-defined, $i = 1, 2$. By Lebesgue Dominated Convergence Theorem 11.131, f_i is integrable over $\mathcal{X}$ and $\int_X f_i d\mu \in \bar{\mathcal{W}}$, $i = 1, 2$. Then, by Definition 11.120, we have $\lim_{E \in \mathcal{M}(\mathcal{X})} F_{i,E} = \int_X f_i d\mu \in \bar{\mathcal{W}}$, $i = 1, 2$.

(ii) By Propositions 11.38, 11.39, 7.21, and 7.23, $f_1 + f_2$ and $\mathcal{P} \circ (f_1 + f_2)$ are $\mathcal{B}$ -measurable. Note that $\mathcal{P} \circ (f_1 + f_2)(x) = \|f_1(x) + f_2(x)\| \leq \|f_1(x)\| + \|f_2(x)\| = (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x)$, $\forall x \in X$. By Proposition 11.83, $0 \leq \int_X \mathcal{P} \circ (f_1 + f_2) d\mathcal{P} \circ \mu \leq \int_X (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2) d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu + \int_X \mathcal{P} \circ f_2 d\mathcal{P} \circ \mu < +\infty$. Hence, $f_1 + f_2$ is absolutely

integrable over $\mathcal{X}$. Let $(\bar{F}_E)_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X (f_1 + f_2) d\mu$ as defined in Definition 11.120. $\forall E \in \mathcal{M}(\mathcal{X})$, $\bar{F}_E = \int_E (f_1 + f_2)|_E d\mu_E = \int_E f_1|_E d\mu_E + \int_E f_2|_E d\mu_E = F_{1,E} + F_{2,E} \in \bar{\mathcal{W}}$, where the second equality follows from Proposition 11.130. Then, $\int_X (f_1 + f_2) d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} \bar{F}_E = \lim_{E \in \mathcal{M}(\mathcal{X})} (F_{1,E} + F_{2,E}) = \int_X f_1 d\mu + \int_X f_2 d\mu \in \bar{\mathcal{W}}$, where the third equality follows from Propositions 3.66, 3.67, and 7.23. Hence, $f_1 + f_2$ is integrable over $\mathcal{X}$.

(iii) By Propositions 11.38, 7.21, and 7.62, Af_1 and $\mathcal{P} \circ (Af_1)$ are $\mathcal{B}$ -measurable. Note that $\mathcal{P} \circ (Af_1)(x) = \|Af_1(x)\| \leq \|A\| \|f_1(x)\| = \|A\| \mathcal{P} \circ f_1(x)$, $\forall x \in X$. By Proposition 11.83 and the fact that $\int_X \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu \in \mathbb{R}$, we have $0 \leq \int_X \mathcal{P} \circ (Af_1) d\mathcal{P} \circ \mu \leq \int_X (\|A\| \mathcal{P} \circ f_1) d\mathcal{P} \circ \mu = \|A\| \int_X \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu < +\infty$. Hence, Af_1 is absolutely integrable over $\mathcal{X}$. Note also that $Af_1 : X \rightarrow \hat{\mathcal{W}}$, where $\hat{\mathcal{W}} := \{F \in \mathcal{B}(\mathcal{Y}, \mathcal{U}) \mid F = AF_1, F_1 \in \mathcal{W}\}$ is a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{U})$. By (i), $\int_X (Af_1) d\mu \in \mathcal{U}$. Let $(\bar{F}_E)_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X (Af_1) d\mu$ as defined in Definition 11.120. $\forall E \in \mathcal{M}(\mathcal{X})$, $\bar{F}_E = \int_E (Af_1)|_E d\mu_E = A \int_E f_1|_E d\mu_E = AF_{1,E} \in \mathcal{U}$, where the second equality follows from Proposition 11.130. Then, by Propositions 3.66 and 7.62, we have $\int_X (Af_1) d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} \bar{F}_E = \lim_{E \in \mathcal{M}(\mathcal{X})} AF_{1,E} = A \int_X f_1 d\mu \in \mathcal{U}$. Hence, Af_1 is integrable over $\mathcal{X}$.

(iv) This follows immediately from (iii).

(v) Fix any $H \in \mathcal{B}$. By Proposition 11.41, $f_1|_H$ and $\mathcal{P} \circ (f_1|_H) = (\mathcal{P} \circ f_1)|_H$ are $\mathcal{B}_H$ -measurable. By Proposition 11.83, $0 \leq \int_H \mathcal{P} \circ (f_1|_H) d\mathcal{P} \circ \mu_H \leq \int_X (\mathcal{P} \circ f_1) d\mathcal{P} \circ \mu < +\infty$. Hence, $f_1|_H$ is absolutely integrable over $\mathcal{H}$. By (i), $f_1|_H$ is integrable over $\mathcal{H}$. By Propositions 11.38, 11.39, and 7.23, $f_1\chi_{H,X}$ is $\mathcal{B}$ -measurable. By Proposition 11.83, $f_1\chi_{H,X}$ is absolutely integrable over $\mathcal{X}$. Again, by (i), $f_1\chi_{H,X}$ is integrable over $\mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists A_0 \in \mathcal{M}(\mathcal{X})$ such that $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have $\|\int_{A_0} (f_1\chi_{H,X})|_{A_0} d\mu_{A_0} - \int_X (f_1\chi_{H,X}) d\mu\| < \epsilon/2$. We will distinguish two exhaustive and mutually exclusive subcases: Case 2a: $\mathcal{P} \circ \mu(H) = +\infty$; Case 2b: $\mathcal{P} \circ \mu(H) < +\infty$.

Case 2a: $\mathcal{P} \circ \mu(H) = +\infty$. $\exists E_0 \in \mathcal{M}(\mathcal{H})$ such that $\forall E \in \mathcal{M}(\mathcal{H})$ with $E_0 \subseteq E$, we have $\|\int_{E_0} f_1|_{E_0} d\mu_{E_0} - \int_H f_1|_H d\mu_H\| < \epsilon/2$. Let $A_1 := A_0 \cup E_0 \in \mathcal{M}(\mathcal{X})$ and $E_1 := A_1 \cap H \in \mathcal{M}(\mathcal{H})$. By Proposition 11.130, $\int_{A_1} (f_1\chi_{H,X})|_{A_1} d\mu_{A_1} = \int_{H \cap A_1} f_1|_{H \cap A_1} d\mu_{H \cap A_1} = \int_{E_1} f_1|_{E_1} d\mu_{E_1}$. Then, we have $\|\int_X (f_1\chi_{H,X}) d\mu - \int_H f_1|_H d\mu_H\| \leq \left\| \int_X (f_1\chi_{H,X}) d\mu - \int_{A_1} (f_1\chi_{H,X})|_{A_1} d\mu_{A_1} \right\| + \left\| \int_{E_1} f_1|_{E_1} d\mu_{E_1} - \int_H f_1|_H d\mu_H \right\| < \epsilon/2 + \epsilon/2 = \epsilon$. By the arbitrariness of ϵ , we have $\int_X (f_1\chi_{H,X}) d\mu = \int_H f_1|_H d\mu_H \in \bar{\mathcal{W}}$.

Case 2b: $\mathcal{P} \circ \mu(H) < +\infty$. Let $A_1 := A_0 \cup H \in \mathcal{M}(\mathcal{X})$. By Proposition 11.130, $\int_{A_1} (f_1\chi_{H,X})|_{A_1} d\mu_{A_1} = \int_H f_1|_H d\mu_H$. Then, we have $\left\| \int_X (f_1\chi_{H,X}) d\mu - \int_H f_1|_H d\mu_H \right\| = \left\| \int_X (f_1\chi_{H,X}) d\mu - \int_{A_1} (f_1\chi_{H,X})|_{A_1} d\mu_{A_1} \right\| < \epsilon/2$. By the arbitrariness of ϵ , we have $\int_X (f_1\chi_{H,X}) d\mu = \int_H f_1|_H d\mu_H \in \bar{\mathcal{W}}$.

Hence, in both subcases, $\int_X (f_1 \chi_{H,X}) d\mu = \int_H f_1|_H d\mu_H \in \bar{\mathcal{W}}$.

(vi) This follows immediately from Lebesgue Dominated Convergence Theorem 11.131.

(vii) $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$ and $\bar{E}_n := \bigcup_{i=1}^n E_i \in \mathcal{B}$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} f_1(x) \chi_{\bar{E}_n, X}(x) = f_1(x) \chi_{E, X}(x)$, $\forall x \in X$. Then, we have $\sum_{i=1}^\infty \int_{E_i} f_1 d\mu = \lim_{n \in \mathbb{N}} \sum_{i=1}^n \int_X (f_1 \chi_{E_i, X}) d\mu = \lim_{n \in \mathbb{N}} \int_X (f_1 \chi_{\bar{E}_n, X}) d\mu = \int_X (f_1 \chi_{E, X}) d\mu = \int_E f_1 d\mu \in \bar{\mathcal{W}}$, where the first equality follows from (v); the second equality follows from (ii); the third equality follows from Lebesgue Dominated Convergence Theorem 11.131; the last equality follows from (v); and the last inclusion follows from (v).

(viii) This follows immediately from Lebesgue Dominated Convergence Theorem 11.131.

This completes the proof of the proposition. $\square$

Proposition 11.133 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, and $f : X \rightarrow \mathcal{W}$ be absolutely integrable over $\mathcal{X}$. Then, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : X \rightarrow \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\|\varphi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$, $\lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu = \int_X f d\mu \in \mathcal{Z}$, $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \varphi_n d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu \in [0, \infty) \subset \mathbb{R}$, and $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ (\varphi_n - f) d\mathcal{P} \circ \mu = 0$.*

Proof By Proposition 11.66, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : X \rightarrow \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\|\varphi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$. By Lebesgue Dominated Convergence Theorem 11.131, $\lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu = \int_X f d\mu \in \mathcal{Z}$. Note that $\lim_{n \in \mathbb{N}} \mathcal{P} \circ \varphi_n = \mathcal{P} \circ f$ a.e. in $\mathcal{X}$, by Propositions 7.21 and 11.52. By Lebesgue Dominated Convergence Theorem 11.91, we have $\lim_{n \in \mathbb{N}} \int_X (\mathcal{P} \circ \varphi_n) d\mathcal{P} \circ \mu = \int_X (\mathcal{P} \circ f) d\mathcal{P} \circ \mu \in [0, \infty) \subset \mathbb{R}$. By Lemma 11.43, $\mathcal{P} \circ (\varphi_n - f)$ is $\mathcal{B}$ -measurable. By Propositions 7.21, 7.23, 11.53, and 11.52, we have $\lim_{n \in \mathbb{N}} \mathcal{P} \circ (\varphi_n - f) = 0$ a.e. in $\mathcal{X}$. Note that $0 \leq \mathcal{P} \circ (\varphi_n - f)(x) = \|\varphi_n(x) - f(x)\| \leq \|\varphi_n(x)\| + \|f(x)\| \leq 2\mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91 and Proposition 11.75, we have $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ (\varphi_n - f) d\mathcal{P} \circ \mu = 0$. This completes the proof of the proposition. $\square$

Proposition 11.134 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, and $f : X \rightarrow \mathcal{W}$ be absolutely integrable over $\mathcal{X}$. Then, there exists a sequence of simple functions $(\phi_n)_{n=1}^\infty$, $\phi_n : X \rightarrow \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\|\phi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \int_X \phi_n d\mu = \int_X f d\mu \in \mathcal{Z}$, $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \phi_n d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu \in [0, \infty) \subset \mathbb{R}$, and $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ (\phi_n - f) d\mathcal{P} \circ \mu = 0$.*

Proof Let $\bar{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$, which is a measure space. By the assumption, f is absolutely integrable over $\bar{\mathcal{X}}$. By Proposition 11.95, there

exists a sequence of simple functions $(\phi_n)_{n=1}^\infty$, $\phi_n : X \rightarrow \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\|\phi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \phi_n d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu \in [0, \infty) \subset \mathbb{R}$, and $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ (\phi_n - f) d\mathcal{P} \circ \mu = 0$. By Propositions 11.132 and 11.125, we have $\int_X f d\mu \in \mathbb{Z}$ and $\int_X \phi_n d\mu \in \mathbb{Z}$, $\forall n \in \mathbb{N}$. By Proposition 11.132, $\lim_{n \in \mathbb{N}} \|\int_X \phi_n d\mu - \int_X f d\mu\| = \lim_{n \in \mathbb{N}} \|\int_X (\phi_n - f) d\mu\| \leq \lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ (\phi_n - f) d\mathcal{P} \circ \mu = 0$. Then, $\lim_{n \in \mathbb{N}} \int_X \phi_n d\mu = \int_X f d\mu \in \mathbb{Z}$. This completes the proof of the proposition. $\square$

11.7 Calculation With Measures

It is easy to see that a σ -finite measure space $(X, \mathcal{B}, \mu)$ may be identified with a σ -finite $\mathbb{R}$ -valued measure space $(X, \mathcal{B}, \bar{\mu})$ as follows: $\bar{\mu}(E) = \mu(E) \in [0, +\infty) \subset \mathbb{R}$, $\forall E \in \mathcal{B}$ with $\mu(E) < +\infty$; and $\bar{\mu}(E)$ is undefined, $\forall E \in \mathcal{B}$ with $\mu(E) = +\infty$. Then, $\mathcal{P} \circ \bar{\mu} = \mu$. The converse of this identification is proved in the following proposition. Thus, a σ -finite measure on $(X, \mathcal{B})$ is a special kind of σ -finite $\mathbb{R}$ -valued measure on $(X, \mathcal{B})$. We will make use of this identification to imply definitions and results for σ -finite measures from definitions and results for σ -finite $\mathbb{R}$ -valued measures.

Proposition 11.135 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite $\mathbb{R}$ -valued measure space with $\nu := \mathcal{P} \circ \mu$ and $\mu(E) \geq 0$, $\forall E \in \text{dom}(\mu)$. Then, any function $\bar{\nu} : \mathcal{B} \rightarrow \mathbb{R}_e$ satisfying $\bar{\nu}(E) = \sum_{i=1}^\infty \bar{\nu}(E_i)$, $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$; and $\bar{\nu}(E) = \mu(E)$, $\forall E \in \text{dom}(\mu)$, must be such that $\bar{\nu} = \nu$.*

Proof $\forall E \in \mathcal{B}$, we will show that $\bar{\nu}(E) = \nu(E)$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $E \in \text{dom}(\mu)$; Case 2: $E \in \mathcal{B} \setminus \text{dom}(\mu)$.

Case 1: $E \in \text{dom}(\mu)$. Then, $\bar{\nu}(E) = \mu(E) \in [0, \infty) \subset \mathbb{R}$. By Definition 11.108, $\mu(E) \leq \nu(E) < \infty$ and

$$\nu(E) = \sup_{\substack{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E \\ E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n |\mu(E_i)| = \mu(E) = \bar{\nu}(E)$$

Case 2: $E \in \mathcal{B} \setminus \text{dom}(\mu)$. Then, $\nu(E) = \infty$. Since $\mathcal{X}$ is σ -finite, then $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ with $\nu(X_n) < \infty$, $\forall n \in \mathbb{N}$ and $X = \bigcup_{n=1}^\infty X_n$. Without loss of generality, we may assume that $(X_n)_{n=1}^\infty$ is pairwise disjoint. Then, $E = \bigcup_{n=1}^\infty (E \cap X_n)$, $0 \leq \nu(E \cap X_n) \leq \nu(X_n) < \infty$, and $E \cap X_n \in \text{dom}(\mu)$, $\forall n \in \mathbb{N}$. Then, $\bar{\nu}(E) = \sum_{n=1}^\infty \bar{\nu}(E \cap X_n) = \sum_{n=1}^\infty \nu(E \cap X_n) = \nu(E) = \infty$, where the second equality follows from Case 1.

Hence, $\bar{\nu}(E) = \nu(E)$ in both cases. This completes the proof of the proposition. $\square$

Proposition 11.136 *Let $(X, \mathcal{B})$ be a measurable space and $\mathcal{Y}$ and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$. Then, the following statements hold.*

- (i) Let μ_1 and μ_2 be measures on $(X, \mathcal{B})$, $a \in (0, \infty) \subset \mathbb{R}$. Define $\mu := \mu_1 + \mu_2 : \mathcal{B} \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $\mu(E) = \mu_1(E) + \mu_2(E)$, $\forall E \in \mathcal{B}$, then μ is a measure on $(X, \mathcal{B})$. μ is σ -finite if μ_1 and μ_2 are σ -finite. Define $\bar{\mu} := a\mu_1 : \mathcal{B} \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $\bar{\mu}(E) = a\mu_1(E)$, $\forall E \in \mathcal{B}$, then $\bar{\mu}$ is a measure on $(X, \mathcal{B})$. $\bar{\mu}$ is σ -finite if μ_1 is σ -finite.
- (ii) Let μ be a finite measure on $(X, \mathcal{B})$ and $a \in [0, \infty) \subset \mathbb{R}$. Define $\bar{\mu} := a\mu : \mathcal{B} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\bar{\mu}(E) = a\mu(E)$, $\forall E \in \mathcal{B}$. Then, $\bar{\mu}$ is a finite measure on $(X, \mathcal{B})$.
- (iii) Let ν_1 and ν_2 be finite $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$ and $\alpha \in \mathbb{K}$. Define $\nu := \nu_1 + \nu_2 : \mathcal{B} \rightarrow \mathcal{Y}$ by $\nu(E) = \nu_1(E) + \nu_2(E)$, $\forall E \in \mathcal{B}$. Then, ν is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \nu \leq \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2$. Define $\bar{\nu} := \alpha\nu_1 : \mathcal{B} \rightarrow \mathcal{Y}$ by $\bar{\nu}(E) = \alpha\nu_1(E)$, $\forall E \in \mathcal{B}$. Then, $\bar{\nu}$ is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \bar{\nu} = |\alpha| \mathcal{P} \circ \nu_1$.
- (iv) Let μ be a finite $\mathbb{K}$ -valued measure on $(X, \mathcal{B})$ and $y \in \mathcal{Y}$. Define $\nu := \mu y : \mathcal{B} \rightarrow \mathcal{Y}$ by $\nu(E) = \mu(E)y$, $\forall E \in \mathcal{B}$. Then, ν is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \nu = \|y\| \mathcal{P} \circ \mu$.
- (v) Let μ be a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ and $A \in B(\mathcal{Y}, \mathcal{Z})$. Define $\bar{\mu} := A\mu : \mathcal{B} \rightarrow \mathcal{Z}$ by $\bar{\mu}(E) = A\mu(E)$, $\forall E \in \mathcal{B}$. Then, $\bar{\mu}$ is a finite $\mathcal{Z}$ -valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \bar{\mu} \leq \|A\| \mathcal{P} \circ \mu$.

Proof (i) Clearly, $\mu(\emptyset) = \mu_1(\emptyset) + \mu_2(\emptyset) = 0$ and $\bar{\mu}(\emptyset) = a\mu_1(\emptyset) = 0$. $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$. Then, $\sum_{i=1}^\infty (\mu_1 + \mu_2)(E_i) = \sum_{i=1}^\infty (\mu_1(E_i) + \mu_2(E_i)) = \sum_{i=1}^\infty \mu_1(E_i) + \sum_{i=1}^\infty \mu_2(E_i) = \mu_1(E) + \mu_2(E) = (\mu_1 + \mu_2)(E)$ and $\sum_{i=1}^\infty (a\mu_1)(E_i) = \sum_{i=1}^\infty a\mu_1(E_i) = a \sum_{i=1}^\infty \mu_1(E_i) = a\mu_1(E)$. Hence, $\mu_1 + \mu_2$ and $a\mu_1$ are measures on $(X, \mathcal{B})$.

When μ_1 is σ -finite, there exists $(X_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty X_n$ and $\mu_1(X_n) < \infty$, $\forall n \in \mathbb{N}$. Then, $(a\mu_1)(X_n) = a\mu_1(X_n) < \infty$. Hence, $a\mu_1$ is σ -finite.

When μ_2 is also σ -finite, there exists $(\bar{X}_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty \bar{X}_n$ and $\mu_2(\bar{X}_n) < \infty$, $\forall n \in \mathbb{N}$. Then, $X = \bigcup_{n=1}^\infty \bigcup_{m=1}^\infty (X_n \cap \bar{X}_m)$, $X_n \cap \bar{X}_m \in \mathcal{B}$, and $\mu(X_n \cap \bar{X}_m) = \mu_1(X_n \cap \bar{X}_m) + \mu_2(X_n \cap \bar{X}_m) \leq \mu_1(X_n) + \mu_2(\bar{X}_m) < \infty$. Hence, μ is also σ -finite.

(ii) Clearly, $\bar{\mu}(\emptyset) = a\mu(\emptyset) = 0$. $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$. Then, $\sum_{i=1}^\infty (a\mu)(E_i) = \sum_{i=1}^\infty a\mu(E_i) = a \sum_{i=1}^\infty \mu(E_i) = a\mu(E) = (a\mu)(E)$, where the second equality follows from the fact that $\sum_{i=1}^\infty \mu(E_i) = \mu(E) \leq \mu(X) < +\infty$. Hence, $a\mu$ is a finite measure on $(X, \mathcal{B})$.

(iii) Clearly, $(\nu_1 + \nu_2)(\emptyset) = \nu_1(\emptyset) + \nu_2(\emptyset) = \vartheta_{\mathcal{Y}}$ and $(\alpha\nu_1)(\emptyset) = \alpha\nu_1(\emptyset) = \vartheta_{\mathcal{Y}}$. $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$. Then, $\sum_{i=1}^\infty \|(\nu_1 + \nu_2)(E_i)\| = \sum_{i=1}^\infty \|\nu_1(E_i) + \nu_2(E_i)\| \leq \sum_{i=1}^\infty (\|\nu_1(E_i)\| + \|\nu_2(E_i)\|) = \sum_{i=1}^\infty \|\nu_1(E_i)\| + \sum_{i=1}^\infty \|\nu_2(E_i)\| < +\infty$, where the last inequality follows from the fact that ν_1 and ν_2 are finite $\mathcal{Y}$ -valued measures, and $\sum_{i=1}^\infty \|(\alpha\nu_1)(E_i)\| = \sum_{i=1}^\infty |\alpha| \|\nu_1(E_i)\| = |\alpha| \sum_{i=1}^\infty \|\nu_1(E_i)\| < +\infty$.

Note that $\sum_{i=1}^{\infty} (\nu_1 + \nu_2)(E_i) = \sum_{i=1}^{\infty} (\nu_1(E_i) + \nu_2(E_i)) = \sum_{i=1}^{\infty} \nu_1(E_i) + \sum_{i=1}^{\infty} \nu_2(E_i) = \nu_1(E) + \nu_2(E) = (\nu_1 + \nu_2)(E) \in \mathcal{Y}$, where the second equality follows from Propositions 7.23, 3.66, and 3.67, and $\sum_{i=1}^{\infty} (\alpha\nu_1)(E_i) = \sum_{i=1}^{\infty} \alpha\nu_1(E_i) = \alpha\nu_1(E) = (\alpha\nu_1)(E) \in \mathcal{Y}$, where the second equality follows from Propositions 7.23, 3.66, and 3.67. Hence, $\nu_1 + \nu_2$ and $\alpha\nu_1$ are $\mathcal{Y}$ -valued pre-measures on $(X, \mathcal{B})$.

$\forall E \in \mathcal{B}, \forall n \in \mathbb{Z}_+, \forall$ pairwise disjoint $(E_i)_{i=1}^n \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^n E_i$, we have $\sum_{i=1}^n \|(\nu_1 + \nu_2)(E_i)\| \leq \sum_{i=1}^n \|\nu_1(E_i)\| + \sum_{i=1}^n \|\nu_2(E_i)\| \leq \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E)$. Hence, $\mathcal{P} \circ (\nu_1 + \nu_2)(E) \leq \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E)$. Then, $\mathcal{P} \circ \nu \leq \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2$. Since ν_1 and ν_2 are finite, then $\mathcal{P} \circ \nu(X) \leq \mathcal{P} \circ \nu_1(X) + \mathcal{P} \circ \nu_2(X) < +\infty$. Therefore, $\nu_1 + \nu_2$ is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$.

$\forall E \in \mathcal{B}$, we have

$$\begin{aligned} \mathcal{P} \circ (\alpha\nu_1)(E) &= \sup_{\substack{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E \\ E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n \|(\alpha\nu_1)(E_i)\| \\ &= \sup_{\substack{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E \\ E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n}} |\alpha| \sum_{i=1}^n \|\nu_1(E_i)\| = |\alpha| \mathcal{P} \circ \nu_1(E) \end{aligned}$$

where the last equality follows from Proposition 3.81 and the fact that $0 \leq \mathcal{P} \circ \nu_1(E) \leq \mathcal{P} \circ \nu_1(X) < +\infty$. Hence, $\mathcal{P} \circ (\alpha\nu_1) = |\alpha| \mathcal{P} \circ \nu_1$. Therefore, $\alpha\nu_1$ is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$.

(iv) Clearly, $(\mu y)(\emptyset) = \mu(\emptyset)y = \vartheta_{\mathcal{Y}}$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$. Then, $\sum_{i=1}^{\infty} \|(\mu y)(E_i)\| = \sum_{i=1}^{\infty} \|y\| |\mu(E_i)| = \|y\| \sum_{i=1}^{\infty} |\mu(E_i)| < +\infty$. Note that $\sum_{i=1}^{\infty} (\mu y)(E_i) = \sum_{i=1}^{\infty} \mu(E_i)y = \mu(E)y = (\mu y)(E) \in \mathcal{Y}$, where the second equality follows from Propositions 7.23, 3.66, and 3.67. Hence, μy is a $\mathcal{Y}$ -valued pre-measure on $(X, \mathcal{B})$. $\forall E \in \mathcal{B}$,

$$\begin{aligned} \mathcal{P} \circ (\mu y)(E) &= \sup_{\substack{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E \\ E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n \|(\mu y)(E_i)\| \\ &= \sup_{\substack{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E \\ E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n}} \|y\| \sum_{i=1}^n |\mu(E_i)| = \|y\| \mathcal{P} \circ \mu(E) \end{aligned}$$

where the last equality follows from Proposition 3.81 and the fact that $0 \leq \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu(X) < +\infty$. Hence, $\mathcal{P} \circ (\mu y) = \|y\| \mathcal{P} \circ \mu$. Therefore, μy is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$.

(v) Clearly, $\bar{\mu}(\emptyset) = A\mu(\emptyset) = \vartheta_{\mathcal{Z}}$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$. Then, $\mu(E) = \sum_{i=1}^{\infty} \mu(E_i) \in \mathcal{Y}$, and $\sum_{i=1}^{\infty} \|\mu(E_i)\| < \infty$. This implies that $\bar{\mu}(E) = A \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} A\mu(E_i) = \sum_{i=1}^{\infty} \bar{\mu}(E_i) \in \mathcal{Z}$, where the second equality follows from Proposition 3.66,

and $\sum_{i=1}^{\infty} \|\bar{\mu}(E_i)\| \leq \sum_{i=1}^{\infty} \|A\| \|\mu(E_i)\| = \|A\| \sum_{i=1}^{\infty} \|\mu(E_i)\| < \infty$. Hence, $\bar{\mu}$ is a $\mathcal{Z}$ -valued pre-measure on $(X, \mathcal{B})$. $\forall E \in \mathcal{B}$,

$$\begin{aligned} \mathcal{P} \circ \bar{\mu}(E) &= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\bar{\mu}(E_i)\| \\ &\leq \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|A\| \|\mu(E_i)\| \\ &= \|A\| \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(E_i)\| \\ &= \|A\| \mathcal{P} \circ \mu(E) \leq \|A\| \mathcal{P} \circ \mu(X) < \infty \end{aligned}$$

where the second equality follows from Proposition 3.81. Hence, $\mathcal{P} \circ \bar{\mu} \leq \|A\| \mathcal{P} \circ \mu$ and $\bar{\mu}$ is a finite $\mathcal{Z}$ -valued pre-measure on $(X, \mathcal{B})$. Then, $\bar{\mu}$ is a finite $\mathcal{Z}$ -valued measure on $(X, \mathcal{B})$.

This completes the proof of the proposition. $\square$

Proposition 11.137 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and ν_1 and ν_2 be $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$. Assume that, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$, we have $\nu_1(E) = \nu_2(E) \in \mathcal{Y}$. Then, $\nu_1 = \nu_2$ is σ -finite.*

Proof Since μ is σ -finite, then $\exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Fix any $n \in \mathbb{N}$. By the assumption, $\mu(X_n) < \infty$ implies that $X_n \in \text{dom}(\nu_1) \cap \text{dom}(\nu_2)$ and $\mathcal{P} \circ \nu_i(X_n) < \infty$, $i = 1, 2$. Hence, ν_1 and ν_2 are σ -finite. Let $\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite measure subspace of $\mathcal{X}$, $(X_n, \mathcal{B}_n, \nu_{n,i})$ be the finite $\mathcal{Y}$ -valued measure subspace of $(X, \mathcal{B}, \nu_i)$, $i = 1, 2$. $\forall E \in \mathcal{B}_n$, we have $\mu(E) \leq \mu(X_n) < \infty$. Then, $\nu_{n,1}(E) = \nu_1(E) = \nu_2(E) = \nu_{n,2}(E)$. Thus, $\nu_{n,1} = \nu_{n,2}$. By Proposition 11.117, we have $\nu_1 = \nu_2$. $\square$

Proposition 11.138 *Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $\mathcal{Z}$ be a Banach space over $\mathbb{K}$. Then, the following statements hold.*

- (i) *Let μ_1 and μ_2 be σ -finite $\mathcal{Z}$ -valued measures on $(X, \mathcal{B})$. Then, there exists a unique $\mathcal{Z}$ -valued measure μ on $(X, \mathcal{B})$ such that $\mu(E) = \mu_1(E) + \mu_2(E) \in \mathcal{Z}$, $\forall E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2)$. Furthermore, μ is σ -finite. We will denote this σ -finite $\mathcal{Z}$ -valued measure by $\mu := \mu_1 + \mu_2$. Then, $\mathcal{P} \circ \mu \leq \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2$.*
- (ii) *Let μ be a σ -finite measure on $(X, \mathcal{B})$ and $\alpha \in [0, \infty) \subset \mathbb{R}$. Then, there exists a unique measure $\bar{\mu}$ on $(X, \mathcal{B})$ such that $\bar{\mu}(E) = \alpha\mu(E)$, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$. Furthermore, $\bar{\mu}$ is σ -finite. We will denote this σ -finite measure by $\bar{\mu} := \alpha\mu$. The σ -finite measure $\bar{\mu}$ satisfies: if $\alpha > 0$, then $\bar{\mu}(E) = \alpha\mu(E)$, $\forall E \in \mathcal{B}$; if $\alpha = 0$, then $\bar{\mu}(E) = 0$, $\forall E \in \mathcal{B}$.*

- (iii) Let μ be a σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ and $\alpha \in \mathbb{K}$. Then, there exists a unique $\mathcal{Y}$ -valued measure $\bar{\mu}$ on $(X, \mathcal{B})$ such that $\bar{\mu}(E) = \alpha\mu(E)$, $\forall E \in \text{dom}(\mu)$. Furthermore, $\bar{\mu}$ is σ -finite. We will denote this σ -finite $\mathcal{Y}$ -valued measure by $\bar{\mu} := \alpha\mu$. The σ -finite $\mathcal{Y}$ -valued measure $\bar{\mu}$ satisfies: if $\alpha \neq 0$, then $\bar{\mu}(E) = \alpha\mu(E)$, $\forall E \in \text{dom}(\mu)$, and $\bar{\mu}(E)$ is undefined, $\forall E \in \mathcal{B} \setminus \text{dom}(\mu)$; if $\alpha = 0$, then $\bar{\mu}(E) = \vartheta_{\mathcal{Y}}$, $\forall E \in \mathcal{B}$; and $\mathcal{P} \circ (\alpha\mu) = |\alpha| \mathcal{P} \circ \mu$.
- (iv) Let μ be a σ -finite $\mathbb{K}$ -valued measure on $(X, \mathcal{B})$ and $y \in \mathcal{Y}$. Then, there exists a unique $\mathcal{Y}$ -valued measure $\bar{\mu}$ on $(X, \mathcal{B})$ such that $\bar{\mu}(E) = \mu(E)y$, $\forall E \in \text{dom}(\mu)$. Furthermore, $\bar{\mu}$ is σ -finite. We will denote this σ -finite $\mathcal{Y}$ -valued measure by $\bar{\mu} := \mu y$. The σ -finite $\mathcal{Y}$ -valued measure $\bar{\mu}$ satisfies: if $y \neq \vartheta_{\mathcal{Y}}$, then $\bar{\mu}(E) = \mu(E)y$, $\forall E \in \text{dom}(\mu)$, and $\bar{\mu}(E)$ is undefined, $\forall E \in \mathcal{B} \setminus \text{dom}(\mu)$; if $y = \vartheta_{\mathcal{Y}}$, then $\bar{\mu}(E) = \vartheta_{\mathcal{Y}}$, $\forall E \in \mathcal{B}$; and $\mathcal{P} \circ (\mu y) = \|y\| \mathcal{P} \circ \mu$.
- (v) Let μ be a σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ and $A \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$. Then, there is a unique $\mathcal{Z}$ -valued measure $\bar{\mu}$ on $(X, \mathcal{B})$ such that $\bar{\mu}(E) = A\mu(E)$, $\forall E \in \text{dom}(\mu)$. Furthermore, $\bar{\mu}$ is σ -finite and $\mathcal{P} \circ \bar{\mu} \leq \|A\| \mathcal{P} \circ \mu$. We will denote this σ -finite $\mathcal{Z}$ -valued measure by $\bar{\mu} := A\mu$.

Proof (i) Since μ_1 and μ_2 are σ -finite, then there exists $(X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mathcal{P} \circ \mu_i(X_n) < \infty$, $\forall n \in \mathbb{N}$, $i = 1, 2$. Without loss of generality, we may assume that $(X_n)_{n=1}^{\infty}$ is pairwise disjoint. Let $(X_n, \mathcal{B}_n, \mu_{i,n})$ be the finite $\mathcal{Z}$ -valued measure subspace of $(X, \mathcal{B}, \mu_i)$, $\forall n \in \mathbb{N}$, $i = 1, 2$. By Proposition 11.136, $\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_{1,n} + \mu_{2,n})$ is a finite $\mathcal{Z}$ -valued measure space, $\forall n \in \mathbb{N}$. By Proposition 11.118, the generation process on $(\mathcal{X}_n)_{n=1}^{\infty}$ yields a unique σ -finite $\mathcal{Z}$ -valued measure space $\mathcal{X} := (X, \mathcal{B}, \mu)$ on the set X .

We will show that $\mathcal{P} \circ \mu \leq \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2$. $\forall E \in \mathcal{B}$, $\mathcal{P} \circ \mu(E) \in [0, \infty] \subset \mathbb{R}_e$. We will show that $\mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E)$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) = \infty$; Case 2: $\mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) < \infty$. Case 1: $\mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) = \infty$. Then, it is trivially true that $\mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) = \infty$. Case 2: $\mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) < \infty$. Then, $\mathcal{P} \circ \mu_i(E) \in [0, \infty) \subset \mathbb{R}$, $i = 1, 2$. Hence, $E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2)$. Let $E_n := E \cap X_n \in \mathcal{B}_n$, $\forall n \in \mathbb{N}$. Then, $\mathcal{P} \circ \mu(E) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu(E_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ (\mu_{1,n} + \mu_{2,n})(E_n) \leq \sum_{n=1}^{\infty} (\mathcal{P} \circ \mu_{1,n}(E_n) + \mathcal{P} \circ \mu_{2,n}(E_n)) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_1(E_n) + \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_2(E_n) = \mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) < \infty$, where the second equality follows from Propositions 11.118 and 11.115; the first inequality follows Proposition 11.136; and the third equality follows from Proposition 11.115. Hence, $\mathcal{P} \circ \mu \leq \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2$.

$\forall E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2)$. Then, $\mathcal{P} \circ \mu_i(E) < \infty$ and $\mu_i(E) \in \mathcal{Z}$, $i = 1, 2$. This implies that $\mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) < \infty$, $E \in \text{dom}(\mu)$, and $\mu(E) \in \mathcal{Z}$. Let $E_n := E \cap X_n \in \mathcal{B}_n$, $\forall n \in \mathbb{N}$. Then, $\mu(E) = \sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{\infty} (\mu_{1,n} + \mu_{2,n})(E_n) = \sum_{n=1}^{\infty} (\mu_1(E_n) +$

$\mu_2(E_n)) = \sum_{n=1}^{\infty} \mu_1(E_n) + \sum_{n=1}^{\infty} \mu_2(E_n) = \mu_1(E) + \mu_2(E) \in \mathbb{Z}$, where the first equality follows from Definition 11.108; the second equality follows from Proposition 11.118; the third equality follows from Propositions 11.115 and 11.136; the fourth equality follows from Propositions 7.23, 3.66, and 3.67; and the last equality follows from Definition 11.108. Hence, μ is the σ -finite $\mathbb{Z}$ -valued measure we seek.

(ii) Define $\bar{\mu} : \mathcal{B} \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $\bar{\mu}(E) = \alpha\mu(E)$, $\forall E \in \mathcal{B}$, if $\alpha > 0$; and $\bar{\mu}(E) = 0$, $\forall E \in \mathcal{B}$, if $\alpha = 0$. When $\alpha \in (0, \infty) \subset \mathbb{R}$, $\bar{\mu}$ is a σ -finite measure on $(X, \mathcal{B})$ by Proposition 11.136. When $\alpha = 0$, clearly $\bar{\mu}$ is a finite measure on $(X, \mathcal{B})$. Hence, $\bar{\mu}$ is the σ -finite measure we seek. Let $\hat{\mu}$ be any measure on $(X, \mathcal{B})$ such that $\hat{\mu}(E) = \alpha\mu(E)$, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$. Since μ is σ -finite, then there exists $(X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $(X_n)_{n=1}^{\infty}$ is pairwise disjoint. $\forall E \in \mathcal{B}$, let $E_n := E \cap X_n$, $\forall n \in \mathbb{N}$. Then, $\mu(E_n) < \infty$, $\forall n \in \mathbb{N}$. This implies that $\hat{\mu}(E) = \sum_{n=1}^{\infty} \hat{\mu}(E_n) = \sum_{n=1}^{\infty} \alpha\mu(E_n) = \sum_{n=1}^{\infty} \bar{\mu}(E_n) = \bar{\mu}(E) \in [0, \infty] \subset \mathbb{R}_e$. Hence, $\bar{\mu}$ is unique.

(iii) Define $\nu := |\alpha| \mathcal{P} \circ \mu$ as prescribed in (ii), which is a σ -finite measure on $(X, \mathcal{B})$. Define $\bar{\mu}$ to be a function from $\mathcal{B}$ to $\mathcal{Y}$ by, if $\alpha \neq 0$, then $\bar{\mu}(E) = \alpha\mu(E) \in \mathcal{Y}$, $\forall E \in \text{dom}(\mu)$, and $\bar{\mu}(E)$ is undefined, $\forall E \in \mathcal{B} \setminus \text{dom}(\mu)$; if $\alpha = 0$, then $\bar{\mu}(E) = \vartheta_{\mathcal{Y}}$, $\forall E \in \mathcal{B}$. We will show that $\bar{\mu}$ is a σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ satisfying: $\bar{\mu}(E) = \alpha\mu(E)$, $\forall E \in \text{dom}(\mu)$, and $\mathcal{P} \circ \bar{\mu} = \nu$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\alpha = 0$; Case 2: $\alpha \neq 0$.

Case 1: $\alpha = 0$. Then, $\bar{\mu}(E) = \vartheta_{\mathcal{Y}}$ and $\nu(E) = 0$, $\forall E \in \mathcal{B}$. It is easy to show that $\bar{\mu}$ is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \bar{\mu} = \nu$. It is straightforward to check that $\bar{\mu}(E) = 0\mu(E) = \vartheta_{\mathcal{Y}}$, $\forall E \in \text{dom}(\mu)$. Hence, this case is proved.

Case 2: $\alpha \neq 0$. $\forall E \in \mathcal{B}$ with $\nu(E) = \infty$, then $\mathcal{P} \circ \mu(E) = \infty$. This leads to that $\mu(E)$ is undefined and $\bar{\mu}(E)$ is also undefined. $\forall E \in \mathcal{B}$ with $\nu(E) < \infty$, then $\mathcal{P} \circ \mu(E) = \nu(E)/|\alpha| < \infty$ and $E \in \text{dom}(\mu)$. Then, $\bar{\mu}(E) = \alpha\mu(E) \in \mathcal{Y}$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^{\infty} E_i$, we have $\bar{\mu}(E) = \alpha\mu(E) = \alpha \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \alpha\mu(E_i) = \sum_{i=1}^{\infty} \bar{\mu}(E_i) \in \mathcal{Y}$ and $\sum_{i=1}^{\infty} \|\bar{\mu}(E_i)\| = \sum_{i=1}^{\infty} |\alpha| \|\mu(E_i)\| \leq \sum_{i=1}^{\infty} |\alpha| \mathcal{P} \circ \mu(E_i) = |\alpha| \mathcal{P} \circ \mu(E) = \nu(E) < \infty$. $\forall E \in \mathcal{B}$ with $\nu(E) < \infty$,

$$\begin{aligned}
\nu(E) &= |\alpha| \mathcal{P} \circ \mu(E) \\
&= |\alpha| \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(E_i)\| \\
&= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n |\alpha| \|\mu(E_i)\| \\
&= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\bar{\mu}(E_i)\|
\end{aligned}$$

where the third equality follows from Proposition 3.81. Hence, $\bar{\mu}$ is a σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \bar{\mu} = \nu = |\alpha| \mathcal{P} \circ \mu$. This case is proved.

Hence, $\bar{\mu}$ is the σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ we seek. Let $\hat{\mu}$ be any $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ such that $\hat{\mu}(E) = \alpha\mu(E)$, $\forall E \in \text{dom}(\mu)$. $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu(E) < \infty$, then $E \in \text{dom}(\mu)$ and $\hat{\mu}(E) = \bar{\mu}(E) \in \mathcal{Y}$. By Proposition 11.137, $\bar{\mu} = \hat{\mu}$. Hence, $\bar{\mu}$ is unique.

(iv) Define $\nu := \|y\| \mathcal{P} \circ \mu$ as prescribed in (ii), which is a σ -finite measure on $(X, \mathcal{B})$. Define $\bar{\mu}$ to be a function from $\mathcal{B}$ to $\mathcal{Y}$ by, if $y \neq \vartheta_{\mathcal{Y}}$, then $\bar{\mu}(E) = \mu(E)y \in \mathcal{Y}$, $\forall E \in \text{dom}(\mu)$, and $\bar{\mu}(E)$ is undefined, $\forall E \in \mathcal{B} \setminus \text{dom}(\mu)$; if $y = \vartheta_{\mathcal{Y}}$, then $\bar{\mu}(E) = \vartheta_{\mathcal{Y}}$, $\forall E \in \mathcal{B}$. We will show that $\bar{\mu}$ is a σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ satisfying: $\bar{\mu}(E) = \mu(E)y$, $\forall E \in \text{dom}(\mu)$, and $\mathcal{P} \circ \bar{\mu} = \nu$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $y = \vartheta_{\mathcal{Y}}$; Case 2: $y \neq \vartheta_{\mathcal{Y}}$.

Case 1: $y = \vartheta_{\mathcal{Y}}$. Then, $\bar{\mu}(E) = \vartheta_{\mathcal{Y}}$ and $\nu(E) = 0$, $\forall E \in \mathcal{B}$. It is easy to show that $\bar{\mu}$ is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \bar{\mu} = \nu$. It is straightforward to check that $\bar{\mu}(E) = \mu(E)y = \vartheta_{\mathcal{Y}}$, $\forall E \in \text{dom}(\mu)$. Hence, this case is proved.

Case 2: $y \neq \vartheta_{\mathcal{Y}}$. $\forall E \in \mathcal{B}$ with $\nu(E) = \infty$, then $\mathcal{P} \circ \mu(E) = \infty$. This leads to that $\mu(E)$ is undefined and $\bar{\mu}(E)$ is also undefined. $\forall E \in \mathcal{B}$ with $\nu(E) < \infty$, then $\mathcal{P} \circ \mu(E) = \nu(E)/\|y\| < \infty$ and $E \in \text{dom}(\mu)$. Then, $\bar{\mu}(E) = \mu(E)y \in \mathcal{Y}$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^{\infty} E_i$, we have $\bar{\mu}(E) = \mu(E)y = (\sum_{i=1}^{\infty} \mu(E_i))y = \sum_{i=1}^{\infty} \mu(E_i)y = \sum_{i=1}^{\infty} \bar{\mu}(E_i) \in \mathcal{Y}$ and $\sum_{i=1}^{\infty} \|\bar{\mu}(E_i)\| = \sum_{i=1}^{\infty} \|y\| |\mu(E_i)| \leq \sum_{i=1}^{\infty} \|y\| \mathcal{P} \circ \mu(E_i) = \|y\| \mathcal{P} \circ \mu(E) = \nu(E) < \infty$. $\forall E \in \mathcal{B}$ with $\nu(E) < \infty$,

$$\begin{aligned} \nu(E) &= \|y\| \mathcal{P} \circ \mu(E) \\ &= \|y\| \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n |\mu(E_i)| \\ &= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|y\| |\mu(E_i)| \\ &= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\bar{\mu}(E_i)\| \end{aligned}$$

where the third equality follows from Proposition 3.81. Hence, $\bar{\mu}$ is a σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \bar{\mu} = \nu = \|y\| \mathcal{P} \circ \mu$. This case is proved.

Hence, $\bar{\mu}$ is the σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ we seek. Let $\hat{\mu}$ be any $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ such that $\hat{\mu}(E) = \mu(E)y$, $\forall E \in \text{dom}(\mu)$. $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu(E) < \infty$, then $E \in \text{dom}(\mu)$ and $\hat{\mu}(E) = \bar{\mu}(E) \in \mathcal{Y}$. By Proposition 11.137, $\bar{\mu} = \hat{\mu}$. Hence, $\bar{\mu}$ is unique.

(v) Since μ is σ -finite, then $\exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mathcal{P} \circ \mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that

$(X_n)_{n=1}^\infty$ is pairwise disjoint. Fix any $n \in \mathbb{N}$. Let $\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X} := (X, \mathcal{B}, \mu)$. By Proposition 11.136, $\bar{\mathcal{X}}_n := (X_n, \mathcal{B}_n, \bar{\mu}_n := A\mu_n)$ is a finite $\mathcal{Z}$ -valued measure space with $\mathcal{P} \circ \bar{\mu}_n \leq \|A\| \mathcal{P} \circ \mu_n$. By Proposition 11.118, the generation process on $(\bar{\mathcal{X}}_n)_{n=1}^\infty$ yields a unique σ -finite $\mathcal{Z}$ -valued measure space $\bar{\mathcal{X}} := (X, \mathcal{B}, \bar{\mu})$ on the set X .

$\forall E \in \mathcal{B}$, $\mathcal{P} \circ \bar{\mu}(E) = \sum_{n=1}^\infty \mathcal{P} \circ \bar{\mu}(E \cap X_n) = \sum_{n=1}^\infty \mathcal{P} \circ \bar{\mu}_n(E \cap X_n) \leq \sum_{n=1}^\infty \|A\| \mathcal{P} \circ \mu_n(E \cap X_n)$. If $\|A\| = 0$, then $\mathcal{P} \circ \bar{\mu}(E) \leq \sum_{n=1}^\infty 0 = 0 = (\|A\| \mathcal{P} \circ \mu)(E)$. If $\|A\| > 0$, then $\mathcal{P} \circ \bar{\mu}(E) \leq \sum_{n=1}^\infty \|A\| \mathcal{P} \circ \mu_n(E \cap X_n) = \|A\| \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(E \cap X_n) = \|A\| \mathcal{P} \circ \mu(E) = (\|A\| \mathcal{P} \circ \mu)(E)$. Hence, we have $\mathcal{P} \circ \bar{\mu} \leq \|A\| \mathcal{P} \circ \mu$.

$\forall E \in \text{dom}(\mu)$, $\mathcal{P} \circ \mu(E) < \infty$. Then, $\mathcal{P} \circ \bar{\mu}(E) \leq (\|A\| \mathcal{P} \circ \mu)(E) < \infty$ and $E \in \text{dom}(\bar{\mu})$. This leads to $\bar{\mu}(E) = \sum_{n=1}^\infty \bar{\mu}(E \cap X_n) = \sum_{n=1}^\infty \bar{\mu}_n(E \cap X_n) = \sum_{n=1}^\infty A\mu_n(E \cap X_n) = \sum_{n=1}^\infty A\mu(E \cap X_n) = A \sum_{n=1}^\infty \mu(E \cap X_n) = A\mu(E)$, where the first equality follows from Definition 11.108; the second and the fourth equalities follow from Propositions 11.115 and 11.136; the fifth equality follows from Proposition 3.66; and the last equality follows from Definition 11.108. Hence, $\bar{\mu}$ is the σ -finite $\mathcal{Z}$ -valued measure we seek.

This completes the proof of the proposition. $\square$

Note that, for σ -finite measures μ_1 and μ_2 , by Proposition 11.135, they are identified with σ -finite $\mathbb{R}$ -valued measures $\bar{\mu}_1$ and $\bar{\mu}_2$, respectively. Then, the σ -finite measure $\mu_1 + \mu_2$ defined in Proposition 11.136 is identified with the σ -finite $\mathbb{R}$ -valued measure $\bar{\mu}_1 + \bar{\mu}_2$ defined in Proposition 11.138. Similar statement can be made for $\alpha\mu_1$ and $\alpha\bar{\mu}_1$, where $\alpha \in [0, \infty) \subset \mathbb{R}$. Hence, the definitions for $\mu_1 + \mu_2$ and $\alpha\mu_1$ in Propositions 11.136 and 11.138 are unambiguous.

Proposition 11.139 *Let $m, n \in \mathbb{Z}_+$, $\mathcal{Y}_i$ be a normed linear space over $\mathbb{K}$, $i = 1, \dots, m$, $\mathcal{Z}_j$ be a normed linear space over $\mathbb{K}$, $j = 1, \dots, n$, $\mathcal{Y} := \prod_{i=1}^m \mathcal{Y}_i$, $\mathcal{Z} := \prod_{j=1}^n \mathcal{Z}_j$, $h : \prod_{i=1}^m \prod_{j=1}^n \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \rightarrow \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be defined by,*

$$h(A_{1,1}, A_{2,1}, \dots, A_{n,1}, \dots, A_{1,m}, \dots, A_{n,m}) = \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix}$$

$\forall A_{j,i} \in \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j)$, $i = 1, \dots, m$, $j = 1, \dots, n$, and $p_{j,i} : \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \rightarrow \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j)$ defined by

$$p_{j,i} \left(\begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right) = A_{j,i}$$

$\forall \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $i = 1, \dots, m$, $j = 1, \dots, n$. Then, h

and $p_{j,i}$, $i = 1, \dots, m$, $j = 1, \dots, n$, are bound linear functions with norm less than or equal to 1.

Proof It is easy to show that h is a linear function. Then, we have

$$\begin{aligned}
 \|h\| &= \sup_{\sum_{i=1}^m \sum_{j=1}^n \|A_{j,i}\|^2 \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| \\
 &= \sup_{\sum_{i=1}^m \sum_{j=1}^n \|A_{j,i}\|^2 \leq 1} \sup_{\sum_{i=1}^m \|y_i\|^2 \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \right\| \\
 &= \sup_{\sum_{i=1}^m \sum_{j=1}^n \|A_{j,i}\|^2 \leq 1} \sup_{\sum_{i=1}^m \|y_i\|^2 \leq 1} \left(\sum_{j=1}^n \left\| \sum_{i=1}^m A_{j,i} y_i \right\|^2 \right)^{1/2} \\
 &\leq \sup_{\sum_{i=1}^m \sum_{j=1}^n \|A_{j,i}\|^2 \leq 1} \sup_{\sum_{i=1}^m \|y_i\|^2 \leq 1} \left(\sum_{j=1}^n \left(\sum_{i=1}^m \|A_{j,i}\| \|y_i\| \right)^2 \right)^{1/2} \\
 &\leq \sup_{\sum_{i=1}^m \sum_{j=1}^n \|A_{j,i}\|^2 \leq 1} \sup_{\sum_{i=1}^m \|y_i\|^2 \leq 1} \left(\sum_{j=1}^n \left(\sum_{i=1}^m \|A_{j,i}\|^2 \right) \left(\sum_{i=1}^m \|y_i\|^2 \right) \right)^{1/2} \\
 &\leq \sup_{\sum_{i=1}^m \sum_{j=1}^n \|A_{j,i}\|^2 \leq 1} \left(\sum_{j=1}^n \sum_{i=1}^m \|A_{j,i}\|^2 \right)^{1/2} \leq 1
 \end{aligned}$$

where the second inequality follows from the Cauchy-Schwarz Inequality.

Fix any $i \in \{1, \dots, m\}$ and any $j \in \{1, \dots, n\}$. It is easy to show that $p_{j,i}$ is a linear function. Then, we have

$$\begin{aligned}
 \|p_{j,i}\| &= \sup_{\left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| \leq 1} \|A_{j,i}\| \\
 &= \sup_{\left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| \leq 1} \sup_{\|y_i\| \leq 1} \|A_{j,i} y_i\| \\
 &\leq \sup_{\left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| \leq 1} \sup_{\|y_i\| \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \begin{bmatrix} \vartheta_{y_1} \\ \vdots \\ y_i \\ \vdots \\ \vartheta_{y_m} \end{bmatrix} \right\|
 \end{aligned}$$

$$\begin{aligned}
&\leq \left\| \left[\begin{array}{ccc} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{array} \right] \right\| \leq 1 \sup_{\|y_i\| \leq 1} \left\| \left[\begin{array}{ccc} A_{1,1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{array} \right] \right\| \|y_i\| \\
&\leq 1
\end{aligned}$$

This completes the proof of the proposition. $\square$

Proposition 11.140 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathbb{K}$ -valued measure space, $\mathcal{Y}_i$ be a normed linear space over $\mathbb{K}$, $\mathcal{Z}_j$ be a Banach space over $\mathbb{K}$, $\mathcal{W}_{j,i} \subseteq \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j)$ be a separable subspace, $A_{j,i} : X \rightarrow \mathcal{W}_{j,i}$, $i = 1, 2$, $j = 1, 2$, and $A : X \rightarrow \mathcal{B}(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$ be defined by $A(x) = \begin{bmatrix} A_{1,1}(x) & A_{1,2}(x) \\ A_{2,1}(x) & A_{2,2}(x) \end{bmatrix}$, $\forall x \in X$. Then, A is absolutely integrable over $\mathcal{X}$ if, and only if, $A_{j,i}$, $i = 1, 2$, $j = 1, 2$, are absolutely integrable over $\mathcal{X}$. In this case, we have $\int_X A d\mu = \begin{bmatrix} \int_X A_{1,1} d\mu & \int_X A_{1,2} d\mu \\ \int_X A_{2,1} d\mu & \int_X A_{2,2} d\mu \end{bmatrix} \in \mathcal{B}(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$.*

Proof Let $A_{j,i}$, $i = 1, 2$, $j = 1, 2$, be absolutely integrable over $\mathcal{X}$. By Proposition 11.39, $a : X \rightarrow \mathcal{W}_{1,1} \times \mathcal{W}_{2,1} \times \mathcal{W}_{1,2} \times \mathcal{W}_{2,2}$ defined by $a(x) = (A_{1,1}(x), A_{2,1}(x), A_{1,2}(x), A_{2,2}(x))$, $\forall x \in X$, is $\mathcal{B}$ -measurable. By Propositions 11.139 and 11.38, $A = ha$ is $\mathcal{B}$ -measurable, where h is the bounded linear function defined in Proposition 11.139 for $m = n = 2$. By Propositions 11.38 and 7.21, $\mathcal{P} \circ A$ is $\mathcal{B}$ -measurable. $\forall x \in X$, $\mathcal{P} \circ A(x) = \|ha(x)\| \leq \|h\| \|a(x)\| \leq \|A_{1,1}(x)\| + \|A_{2,1}(x)\| + \|A_{1,2}(x)\| + \|A_{2,2}(x)\| = \sum_{i=1}^2 \sum_{j=1}^2 \mathcal{P} \circ A_{j,i}(x)$. By Proposition 11.83, $0 \leq \int_X (\mathcal{P} \circ A) d\mathcal{P} \circ \mu \leq \int_X (\sum_{i=1}^2 \sum_{j=1}^2 \mathcal{P} \circ A_{j,i}) d\mathcal{P} \circ \mu = \sum_{i=1}^2 \sum_{j=1}^2 \int_X (\mathcal{P} \circ A_{j,i}) d\mathcal{P} \circ \mu < \infty$. Hence, A is absolutely integrable over $\mathcal{X}$. By Propositions 4.31, 7.66, and 7.22, $\mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j)$, $i = 1, 2$, $j = 1, 2$, $\prod_{i=1}^2 \prod_{j=1}^2 \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j)$, and $\mathcal{B}(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$ are Banach spaces over $\mathbb{K}$. By Propositions 7.22, 3.28, and 4.4, $\prod_{i=1}^2 \prod_{j=1}^2 \mathcal{W}_{j,i} \subseteq \prod_{i=1}^2 \prod_{j=1}^2 \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j)$ is a separable subspace. By Proposition 11.132, $A_{j,i}$, $i = 1, 2$, $j = 1, 2$, a , and A are integrable over $\mathcal{X}$ and $\int_X A d\mu = h \int_X a d\mu \in \mathcal{B}(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$.

On the other hand, let A be absolutely integrable over $\mathcal{X}$. Fix any $i \in \{1, 2\}$ and any $j \in \{1, 2\}$. Then, by Propositions 11.139, 11.38, $A_{j,i} = p_{j,i}A$ is $\mathcal{B}$ -measurable, where $p_{j,i}$ is the bounded linear function as defined in Proposition 11.139. By Propositions 7.21 and 11.38, $\mathcal{P} \circ A_{j,i}$ is $\mathcal{B}$ -measurable. $\forall x \in X$, $\mathcal{P} \circ A_{j,i}(x) \leq \|p_{j,i}\| \mathcal{P} \circ A(x) \leq \mathcal{P} \circ A(x)$. Then, by Proposition 11.83, $0 \leq \int_X \mathcal{P} \circ A_{j,i} d\mathcal{P} \circ \mu \leq \int_X \mathcal{P} \circ A d\mathcal{P} \circ \mu < \infty$. Hence, $A_{j,i}$ is absolutely integrable over $\mathcal{X}$. By Proposition 11.132, $A_{j,i}$ is integrable over $\mathcal{X}$ and $\int_X A_{j,i} d\mu \in \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j)$.

Hence, A is absolutely integrable over $\mathcal{X}$ if, and only if, $A_{j,i}$, $i = 1, 2$, $j = 1, 2$, are absolutely integrable over $\mathcal{X}$.

Let $A, A_{j,i}, i = 1, 2, j = 1, 2$, be absolutely integrable over $\mathcal{X}$. We will show that $\int_X A d\mu = \begin{bmatrix} \int_X A_{1,1} d\mu & \int_X A_{1,2} d\mu \\ \int_X A_{2,1} d\mu & \int_X A_{2,2} d\mu \end{bmatrix} \in B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$ by distinguishing two exhaustive and mutually exclusives cases: Case 1: $\mathcal{P} \circ \mu(X) < \infty$; Case 2: $\mathcal{P} \circ \mu(X) = \infty$.

Case 1: $\mathcal{P} \circ \mu(X) < \infty$. By Proposition 11.133, $\forall i \in \{1, 2\}, \forall j \in \{1, 2\}, \exists$ sequence of simple functions $(\varphi_{j,i,n})_{n=1}^\infty, \varphi_{j,i,n} : X \rightarrow \mathcal{W}_{j,i}, \forall n \in \mathbb{N}$, such that $\lim_{n \in \mathbb{N}} \varphi_{j,i,n} = A_{j,i}$ a.e. in $\mathcal{X}, \|\varphi_{j,i,n}(x)\| \leq \mathcal{P} \circ A_{j,i}(x), \forall x \in X, \forall n \in \mathbb{N}$, and $\int_X A_{j,i} d\mu = \lim_{n \in \mathbb{N}} \int_X \varphi_{j,i,n} d\mu \in B(\mathcal{Y}_i, \mathcal{Z}_j)$. By Proposition 11.53, we have $\lim_{n \in \mathbb{N}} \psi_n = a$ a.e. in $\mathcal{X}$, where $\psi_n : X \rightarrow \prod_{i=1}^2 \prod_{j=1}^2 \mathcal{W}_{j,i}$ is defined by $\psi_n(x) = (\varphi_{1,1,n}(x), \varphi_{2,1,n}(x), \varphi_{1,2,n}(x), \varphi_{2,2,n}(x)), \forall x \in X, \forall n \in \mathbb{N}$. Clearly, ψ_n is a simple function, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.131, $\int_X a d\mu = \lim_{n \in \mathbb{N}} \int_X \psi_n d\mu = \lim_{n \in \mathbb{N}} (\int_X \varphi_{1,1,n} d\mu, \int_X \varphi_{2,1,n} d\mu, \int_X \varphi_{1,2,n} d\mu, \int_X \varphi_{2,2,n} d\mu) = (\int_X A_{1,1} d\mu, \int_X A_{2,1} d\mu, \int_X A_{1,2} d\mu, \int_X A_{2,2} d\mu) \in \prod_{i=1}^2 \prod_{j=1}^2 B(\mathcal{Y}_i, \mathcal{Z}_j)$, where the second equality follows from Proposition 11.75; and the third equality follows from Proposition 3.67. Then, by Proposition 11.132, $\int_X A d\mu = h \int_X a d\mu = \begin{bmatrix} \int_X A_{1,1} d\mu & \int_X A_{1,2} d\mu \\ \int_X A_{2,1} d\mu & \int_X A_{2,2} d\mu \end{bmatrix} \in B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$.

Case 2: $\mathcal{P} \circ \mu(X) = \infty$. Let $(A_E)_{E \in \mathcal{M}(\mathcal{X})}$ and $(A_{j,i,E})_{E \in \mathcal{M}(\mathcal{X})}$ be the nets for $\int_X A d\mu$ and $\int_X A_{j,i} d\mu$ as defined in Definition 11.71, respectively, $i = 1, 2, j = 1, 2. \forall E \in \mathcal{M}(\mathcal{X})$, by Case 1, $A_E = \int_E A|_E d\mu_E = \begin{bmatrix} \int_E A_{1,1}|_E d\mu_E & \int_E A_{1,2}|_E d\mu_E \\ \int_E A_{2,1}|_E d\mu_E & \int_E A_{2,2}|_E d\mu_E \end{bmatrix} = \begin{bmatrix} A_{1,1,E} & A_{1,2,E} \\ A_{2,1,E} & A_{2,2,E} \end{bmatrix} \in B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$, where $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$ is the finite $\mathbb{K}$ -valued measure subspace of $\mathcal{X}$. Since $A, A_{j,i}, i = 1, 2, j = 1, 2$ are integrable over $\mathcal{X}$, we have $\int_X A d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} A_E = \lim_{E \in \mathcal{M}(\mathcal{X})} \begin{bmatrix} A_{1,1,E} & A_{1,2,E} \\ A_{2,1,E} & A_{2,2,E} \end{bmatrix} = h \lim_{E \in \mathcal{M}(\mathcal{X})} (A_{1,1,E}, A_{2,1,E}, A_{1,2,E}, A_{2,2,E}) = h(\int_X A_{1,1} d\mu, \int_X A_{2,1} d\mu, \int_X A_{1,2} d\mu, \int_X A_{2,2} d\mu) = \begin{bmatrix} \int_X A_{1,1} d\mu & \int_X A_{1,2} d\mu \\ \int_X A_{2,1} d\mu & \int_X A_{2,2} d\mu \end{bmatrix} \in B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$, where the first equality follows from Definition 11.119; the third equality follows from Propositions 11.139 and 3.66; and the fourth equality follows from Proposition 3.67 and Definition 11.119.

Hence, in both cases, we have $\int_X A d\mu = \begin{bmatrix} \int_X A_{1,1} d\mu & \int_X A_{1,2} d\mu \\ \int_X A_{2,1} d\mu & \int_X A_{2,2} d\mu \end{bmatrix} \in B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$. This completes the proof of the proposition. $\square$

Proposition 11.141 *Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, and Z be the set of σ -finite $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$. Define vector addition and scalar multiplication on Z according to Proposition 11.138 and $\vartheta \in Z$ by $\vartheta(E) = \vartheta_{\mathcal{Y}}, \forall E \in \mathcal{B}$. Then, $(Z, +, \cdot, \vartheta) =: \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ is a vector space over $\mathbb{K}$.*

Proof We will first show that $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ is a vector space over $\mathbb{K}$. Fix any $\mu_1, \mu_2, \mu_3 \in Z$ and fix any $\alpha, \beta \in \mathbb{K}$.

(i) $\forall E \in \mathcal{B}$ with $(\mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2)(E) < \infty$, then $E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2)$. Then, we have $(\mu_1 + \mu_2)(E) = \mu_1(E) + \mu_2(E) = (\mu_2 + \mu_1)(E) \in \mathcal{Y}$. By Proposition 11.137, $\mu_1 + \mu_2 = \mu_2 + \mu_1$.

(ii) $\forall E \in \mathcal{B}$ with $(\mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2 + \mathcal{P} \circ \mu_3)(E) < \infty$, then $E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2) \cap \text{dom}(\mu_3)$ and, by Proposition 11.138, $E \in \text{dom}(\mu_1 + \mu_2) \cap \text{dom}(\mu_2 + \mu_3)$. Then, we have $((\mu_1 + \mu_2) + \mu_3)(E) = (\mu_1 + \mu_2)(E) + \mu_3(E) = (\mu_1(E) + \mu_2(E)) + \mu_3(E) = \mu_1(E) + (\mu_2(E) + \mu_3(E)) = \mu_1(E) + (\mu_2 + \mu_3)(E) = (\mu_1 + (\mu_2 + \mu_3))(E) \in \mathcal{Y}$. By Proposition 11.137, $(\mu_1 + \mu_2) + \mu_3 = \mu_1 + (\mu_2 + \mu_3)$.

(iii) $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu_1(E) < \infty$, we have $(\mu_1 + \vartheta)(E) = \mu_1(E) + \vartheta(E) = \mu_1(E) \in \mathcal{Y}$. By Proposition 11.137, $\mu_1 + \vartheta = \mu_1$.

(iv) $\forall E \in \mathcal{B}$ with $(\mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2)(E) < \infty$, we have $E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2)$ and $E \in \text{dom}(\mu_1 + \mu_2) \cap \text{dom}(\alpha\mu_1) \cap \text{dom}(\alpha\mu_2)$. Then, $(\alpha(\mu_1 + \mu_2))(E) = \alpha(\mu_1 + \mu_2)(E) = \alpha(\mu_1(E) + \mu_2(E)) = \alpha\mu_1(E) + \alpha\mu_2(E) = (\alpha\mu_1)(E) + (\alpha\mu_2)(E) = (\alpha\mu_1 + \alpha\mu_2)(E) \in \mathcal{Y}$. By Proposition 11.137, $\alpha(\mu_1 + \mu_2) = \alpha\mu_1 + \alpha\mu_2$.

(v) $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu_1(E) < \infty$, we have $((\alpha + \beta)\mu_1)(E) = (\alpha + \beta)\mu_1(E) = \alpha\mu_1(E) + \beta\mu_1(E) = (\alpha\mu_1)(E) + (\beta\mu_1)(E) = (\alpha\mu_1 + \beta\mu_1)(E) \in \mathcal{Y}$. By Proposition 11.137, $(\alpha + \beta)\mu_1 = \alpha\mu_1 + \beta\mu_1$.

(vi) $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu_1(E) < \infty$, we have $((\alpha\beta)\mu_1)(E) = (\alpha\beta)\mu_1(E) = \alpha(\beta\mu_1(E)) = \alpha(\beta\mu_1)(E) = (\alpha(\beta\mu_1))(E) \in \mathcal{Y}$. By Proposition 11.137, $(\alpha\beta)\mu_1 = \alpha(\beta\mu_1)$.

(vii) $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu_1(E) < \infty$, we have $(0\mu_1)(E) = 0\mu_1(E) = \vartheta\mu_1(E) = \vartheta(E) \in \mathcal{Y}$ and $(1\mu_1)(E) = 1\mu_1(E) = \mu_1(E) \in \mathcal{Y}$. By Proposition 11.137, $0\mu_1 = \vartheta$ and $1\mu_1 = \mu_1$.

Therefore, $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) = (Z, +, \cdot, \vartheta)$ is a vector space over $\mathbb{K}$. This completes the proof of the proposition. $\square$

Proposition 11.142 *Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and Z be the set of finite $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$. Define vector addition and scalar multiplication on Z according to Proposition 11.136 and $\vartheta \in Z$ by $\vartheta(E) = \vartheta_{\mathcal{Y}}, \forall E \in \mathcal{B}$. Then, $\mathcal{Z} := (Z, +, \cdot, \vartheta)$ is a vector space over $\mathbb{K}$. Define $\|\cdot\| : Z \rightarrow [0, \infty) \subset \mathbb{R}$ by $\|\mu\| = \mathcal{P} \circ \mu(X)$, $\forall \mu \in Z$. Then, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y}) := (\mathcal{Z}, \mathbb{K}, \|\cdot\|)$ is a normed linear space.*

Furthermore, if $\mathcal{Y}$ is a Banach space, then $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ is a subspace of $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ and is a Banach space over $\mathbb{K}$.

Proof We will first show that $\mathcal{Z}$ is a vector space over $\mathbb{K}$. Fix any $\mu_1, \mu_2, \mu_3 \in Z$ and fix any $\alpha, \beta \in \mathbb{K}$.

(i) By Proposition 11.136, $\mu_1 + \mu_2, \mu_2 + \mu_1 \in Z$ and, $\forall E \in \mathcal{B}$, $(\mu_1 + \mu_2)(E) = \mu_1(E) + \mu_2(E) = (\mu_2 + \mu_1)(E) \in \mathcal{Y}$. Hence, $\mu_1 + \mu_2 = \mu_2 + \mu_1$.

(ii) By Proposition 11.136, $(\mu_1 + \mu_2) + \mu_3, \mu_1 + (\mu_2 + \mu_3) \in Z$ and, $\forall E \in \mathcal{B}$, $((\mu_1 + \mu_2) + \mu_3)(E) = (\mu_1 + \mu_2)(E) + \mu_3(E) = (\mu_1(E) + \mu_2(E)) + \mu_3(E) =$

$\mu_1(E) + (\mu_2(E) + \mu_3(E)) = \mu_1(E) + (\mu_2 + \mu_3)(E) = (\mu_1 + (\mu_2 + \mu_3))(E) \in \mathcal{Y}$.
Hence, $(\mu_1 + \mu_2) + \mu_3 = \mu_1 + (\mu_2 + \mu_3)$.

(iii) By Proposition 11.136, $\mu_1 + \vartheta \in Z$ and, $\forall E \in \mathcal{B}$, $(\mu_1 + \vartheta)(E) = \mu_1(E) + \vartheta(E) = \mu_1(E) \in \mathcal{Y}$. Hence, $\mu_1 + \vartheta = \mu_1$.

(iv) By Proposition 11.136, $\alpha(\mu_1 + \mu_2), \alpha\mu_1 + \alpha\mu_2 \in Z$ and, $\forall E \in \mathcal{B}$, $(\alpha(\mu_1 + \mu_2))(E) = \alpha(\mu_1 + \mu_2)(E) = \alpha(\mu_1(E) + \mu_2(E)) = \alpha\mu_1(E) + \alpha\mu_2(E) = (\alpha\mu_1)(E) + (\alpha\mu_2)(E) = (\alpha\mu_1 + \alpha\mu_2)(E) \in \mathcal{Y}$. Hence, $\alpha(\mu_1 + \mu_2) = \alpha\mu_1 + \alpha\mu_2$.

(v) By Proposition 11.136, $(\alpha + \beta)\mu_1, \alpha\mu_1 + \beta\mu_1 \in Z$ and, $\forall E \in \mathcal{B}$, $((\alpha + \beta)\mu_1)(E) = (\alpha + \beta)\mu_1(E) = \alpha\mu_1(E) + \beta\mu_1(E) = (\alpha\mu_1)(E) + (\beta\mu_1)(E) = (\alpha\mu_1 + \beta\mu_1)(E) \in \mathcal{Y}$. Hence, $(\alpha + \beta)\mu_1 = \alpha\mu_1 + \beta\mu_1$.

(vi) By Proposition 11.136, $(\alpha\beta)\mu_1, \alpha(\beta\mu_1) \in Z$ and, $\forall E \in \mathcal{B}$, $((\alpha\beta)\mu_1)(E) = (\alpha\beta)\mu_1(E) = \alpha(\beta\mu_1(E)) = \alpha(\beta\mu_1)(E) = (\alpha(\beta\mu_1))(E) \in \mathcal{Y}$. Then, $(\alpha\beta)\mu_1 = \alpha(\beta\mu_1)$.

(vii) By Proposition 11.136, $0\mu_1, 1\mu_1 \in Z$ and, $\forall E \in \mathcal{B}$, $(0\mu_1)(E) = 0\mu_1(E) = \vartheta_{\mathcal{Y}} = \vartheta(E) \in \mathcal{Y}$ and $(1\mu_1)(E) = 1\mu_1(E) = \mu_1(E) \in \mathcal{Y}$. Hence, $0\mu_1 = \vartheta$ and $1\mu_1 = \mu_1$.

Therefore, Z is a vector space over $\mathbb{K}$. In the following, we show that $\|\cdot\|$ defines a norm on Z . Fix any $\mu_1, \mu_2 \in Z$ and fix any $\alpha \in \mathbb{K}$.

(i) $\|\mu_1\| = \mathcal{P} \circ \mu_1(X) \in [0, \infty) \subset \mathbb{R}$ since $\mu_1 \in Z$. Clearly, $\|\vartheta\| = \mathcal{P} \circ \vartheta(X) = 0$. If $\|\mu_1\| = 0$, we have, $\forall E \in \mathcal{B}$, $0 \leq \|\mu_1(E)\| \leq \mathcal{P} \circ \mu_1(E) \leq \mathcal{P} \circ \mu_1(X) = \|\mu_1\| = 0$, which implies that $\mu_1(E) = \vartheta_{\mathcal{Y}}$, $\forall E \in \mathcal{B}$. Then, $\mu_1 = \vartheta$. This shows that $\|\mu_1\| = 0 \Leftrightarrow \mu_1 = \vartheta$.

(ii) $\|\mu_1 + \mu_2\| = \mathcal{P} \circ (\mu_1 + \mu_2)(X) \leq (\mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2)(X) = \mathcal{P} \circ \mu_1(X) + \mathcal{P} \circ \mu_2(X) = \|\mu_1\| + \|\mu_2\|$, where the inequality follows from Proposition 11.136.

(iii) By Proposition 11.136, $\|\alpha\mu_1\| = \mathcal{P} \circ (\alpha\mu_1)(X) = (|\alpha| \mathcal{P} \circ \mu_1)(X) = |\alpha| \mathcal{P} \circ \mu_1(X) = |\alpha| \|\mu_1\|$.

Hence, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ is a normed linear space.

Let $\mathcal{Y}$ be a Banach space. Clearly, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y}) \subseteq \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$. Therefore, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ is a subspace of $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$. Finally, we show that $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ is a Banach space when $\mathcal{Y}$ is a Banach space. Let $(\mu_n)_{n=1}^\infty \subseteq \mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ be a Cauchy sequence. $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$, $\forall m, n \in \mathbb{N}$ with $N \leq m$ and $N \leq n$, we have $\|\mu_n - \mu_m\| = \mathcal{P} \circ (\mu_n - \mu_m)(X) < \epsilon$. $\forall E \in \mathcal{B}$, $\|\mu_n(E) - \mu_m(E)\| = \|(\mu_n - \mu_m)(E)\| \leq \mathcal{P} \circ (\mu_n - \mu_m)(E) \leq \mathcal{P} \circ (\mu_n - \mu_m)(X) < \epsilon$. Hence, $(\mu_n(E))_{n=1}^\infty$ is a uniform Cauchy sequence uniform in $E \in \mathcal{B}$. By $\mathcal{Y}$ being complete, $\forall E \in \mathcal{B}$, there exists a unique $\mu(E) \in \mathcal{Y}$ such that $\lim_{n \in \mathbb{N}} \mu_n(E) = \mu(E)$. This defines a function $\mu : \mathcal{B} \rightarrow \mathcal{Y}$. Clearly, $\mu(\emptyset) = \lim_{n \in \mathbb{N}} \mu_n(\emptyset) = \vartheta_{\mathcal{Y}}$. $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$. Then, $\sum_{i=1}^\infty \|\mu(E_i)\| = \sum_{i=1}^\infty \lim_{n \in \mathbb{N}} \|\mu_n(E_i)\| \leq \liminf_{n \in \mathbb{N}} \sum_{i=1}^\infty \|\mu_n(E_i)\| \leq \liminf_{n \in \mathbb{N}} \sum_{i=1}^\infty \mathcal{P} \circ \mu_n(E_i) = \liminf_{n \in \mathbb{N}} \mathcal{P} \circ \mu_n(E) \leq \liminf_{n \in \mathbb{N}} \mathcal{P} \circ \mu_n(X) = \lim_{n \in \mathbb{N}} \|\mu_n\| \in \mathbb{R}$, where the first equality follows from Propositions 7.21 and 3.66; the first inequality follows from Fatou's Lemma 11.80; the second inequality follows from Definition 11.108; and the last step

follows from Propositions 4.23 and 7.21 and the fact that $(\mu_n)_{n=1}^\infty$ is a Cauchy sequence. This implies that $\sum_{i=1}^\infty \mu(E_i) = \sum_{i=1}^\infty \lim_{n \in \mathbb{N}} \mu_n(E_i) = \lim_{m \in \mathbb{N}} \lim_{n \in \mathbb{N}} \sum_{i=1}^m \mu_n(E_i) = \lim_{m \in \mathbb{N}} \lim_{n \in \mathbb{N}} \mu_n(\bigcup_{i=1}^m E_i) = \lim_{n \in \mathbb{N}} \lim_{m \in \mathbb{N}} \mu_n(\bigcup_{i=1}^m E_i) = \lim_{n \in \mathbb{N}} \mu_n(E) = \mu(E) \in \mathcal{Y}$, where fourth equality follows from Iterated Limit Theorem 4.56. Hence, μ is a $\mathcal{Y}$ -valued pre-measure on $(X, \mathcal{B})$.

Note that

$$\begin{aligned}
\mathcal{P} \circ \mu(X) &= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, X = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(E_i)\| \\
&= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, X = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \lim_{m \in \mathbb{N}} \sum_{i=1}^n \|\mu_m(E_i)\| \\
&\leq \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, X = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \lim_{m \in \mathbb{N}} \sum_{i=1}^n \mathcal{P} \circ \mu_m(E_i) \\
&= \lim_{m \in \mathbb{N}} \mathcal{P} \circ \mu_m(X) = \lim_{m \in \mathbb{N}} \|\mu_m\| \in \mathbb{R}
\end{aligned}$$

where the second equality follows from Propositions 7.21 and 3.66; and the last step follows from Propositions 4.23 and 7.21 and the fact that $(\mu_n)_{n=1}^\infty$ is a Cauchy sequence. Hence, μ is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ and $\mu \in \mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$.

Note that, $\forall n \in \mathbb{N}$ with $N \leq n$,

$$\begin{aligned}
\|\mu_n - \mu\| &= \mathcal{P} \circ (\mu_n - \mu)(X) \\
&= \sup_{m \in \mathbb{Z}_+, (E_i)_{i=1}^m \subseteq \mathcal{B}, X = \bigcup_{i=1}^m E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq m} \sum_{i=1}^m \|(\mu_n - \mu)(E_i)\| \\
&= \sup_{m \in \mathbb{Z}_+, (E_i)_{i=1}^m \subseteq \mathcal{B}, X = \bigcup_{i=1}^m E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq m} \sum_{i=1}^m \|\mu_n(E_i) - \mu(E_i)\| \\
&= \sup_{\substack{m \in \mathbb{Z}_+, (E_i)_{i=1}^m \subseteq \mathcal{B}, X = \bigcup_{i=1}^m E_i \\ E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq m}} \sum_{i=1}^m \lim_{k \in \mathbb{N}} \|\mu_n(E_i) - \mu_k(E_i)\| \\
&= \sup_{\substack{m \in \mathbb{Z}_+, (E_i)_{i=1}^m \subseteq \mathcal{B}, X = \bigcup_{i=1}^m E_i \\ E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq m}} \lim_{k \in \mathbb{N}} \sum_{i=1}^m \|(\mu_n - \mu_k)(E_i)\| \\
&\leq \sup_{\substack{m \in \mathbb{Z}_+, (E_i)_{i=1}^m \subseteq \mathcal{B}, X = \bigcup_{i=1}^m E_i \\ E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq m}} \liminf_{k \in \mathbb{N}} \sum_{i=1}^m \mathcal{P} \circ (\mu_n - \mu_k)(E_i) \\
&= \liminf_{k \in \mathbb{N}} \mathcal{P} \circ (\mu_n - \mu_k)(X) = \liminf_{k \in \mathbb{N}} \|\mu_n - \mu_k\| \leq \epsilon
\end{aligned}$$

where the fourth equality follows from Propositions 7.21, 7.23, 3.66, and 3.67. Hence, $\lim_{n \in \mathbb{N}} \|\mu_n - \mu\| = 0$ and $\lim_{n \in \mathbb{N}} \mu_n = \mu \in \mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$.

Thus, we have shown that every Cauchy sequence in $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ converges in $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$. Hence, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ is a Banach space.

This completes the proof of the proposition. $\square$

In the above proposition, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ admits a topology $\mathcal{O}_1$ that is induced by the norm. We may define a topology on the vector space $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ when $\mathcal{Y}$ is a Banach space, whose subset topology on $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ is exactly $\mathcal{O}_1$. Let $\mathcal{O}_B$ be a collection of subsets of $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ such that, $\forall B \in \mathcal{O}_B$, $\exists E \in \mathcal{B}$, let $\mathcal{B}_E := \{A \in \mathcal{B} \mid A \subseteq E\}$, $\exists \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \nu_E \in \mathcal{M}_f(E, \mathcal{B}_E, \mathcal{Y})$ such that

$$B = \{ \mu \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) \mid (E, \mathcal{B}_E, \mu_E) \text{ is the finite } \mathcal{Y}\text{-valued measure subspace of } (X, \mathcal{B}, \mu) \text{ and } \|\mu_E - \nu_E\|_{\mathcal{M}_f(E, \mathcal{B}_E, \mathcal{Y})} < \epsilon \}$$

It is easy to show that $\mathcal{O}_B$ is a basis for a topology $\mathcal{O}$ on $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ by Proposition 3.18. We will abuse the notation and denote the topological space $(\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}), \mathcal{O})$ by $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$. With this topology in place, we may define convergence for nets in $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ according to Section 3.9. For clarity, we summarize the concept of convergence for sequences in $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ in the following definition.

Definition 11.143 Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $(\mu_n)_{n=1}^\infty \subseteq \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$, and $\nu \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$. We will say the sequence $(\mu_n)_{n=1}^\infty$ converges to ν and write $\lim_{n \rightarrow \mathbb{N}} \mu_n = \nu$ if, $\forall E \in \text{dom}(\nu)$, $\lim_{n \in \mathbb{N}} \mathcal{P} \circ (\mu_n - \nu)(E) = 0$.

This definition means that if $(\mu_n)_{n=1}^\infty$ converges to ν , $\forall E \in \text{dom}(\nu)$, the subspace $(E, \mathcal{B}_E, \mu_{n,E})$ of $(X, \mathcal{B}, \mu_n)$ is finite for sufficiently large $n \in \mathbb{N}$ and $\lim_{n \in \mathbb{N}} \mu_{n,E} = \nu_E$ in $\mathcal{M}_f(E, \mathcal{B}_E, \mathcal{Y})$, where $(E, \mathcal{B}_E, \nu_E)$ is the finite $\mathcal{Y}$ -valued measure subspace of $(X, \mathcal{B}, \nu)$. Furthermore, by Proposition 11.142 and its proof and Propositions 7.21 and 3.66, we have $\lim_{n \in \mathbb{N}} \mu_n(E) = \nu(E)$ and $\lim_{n \in \mathbb{N}} \mathcal{P} \circ \mu_n(E) = \mathcal{P} \circ \nu(E)$, $\forall E \in \text{dom}(\nu)$.

Proposition 11.144 Let $(X, \mathcal{B})$ be a measurable space, μ_1 and μ_2 be σ -finite measures on $(X, \mathcal{B})$, $a \in [0, \infty) \subset \mathbb{R}$, $f : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, and $\mathcal{X}_i := (X, \mathcal{B}, \mu_i)$, $i = 1, 2$. Then, the following statements hold.

- (i) If $\mu_1 \leq \mu_2$ (that is $\mu_1(E) \leq \mu_2(E)$, $\forall E \in \mathcal{B}$), then $0 \leq \int_X f d\mu_1 \leq \int_X f d\mu_2 \leq \infty$.
- (ii) Let $\mu := \mu_1 + \mu_2$ (as defined in Proposition 11.136), then $0 \leq \int_X f d\mu = \int_X f d\mu_1 + \int_X f d\mu_2 \leq \infty$.
- (iii) Let $\nu := a\mu_1$ (as defined in Proposition 11.138), then $0 \leq \int_X f d\nu = \int_X (af) d\mu_1 \leq \infty$.

Proof (i) We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\int_X f d\mu_2 = \infty$; Case 2: $\int_X f d\mu_2 < \infty$. Case 1: $\int_X f d\mu_2 = \infty$. The result holds trivially in this case. Case 2: $\int_X f d\mu_2 < \infty$. By Definition 11.79, $0 \leq \int_X f d\mu_1 = \sup_{0 \leq \phi \leq f} \int_X \phi d\mu_1 \leq \infty$, where the supremum is over all simple functions $\phi : \mathcal{X}_1 \rightarrow [0, \infty) \subset \mathbb{R}$ such that $0 \leq \phi(x) \leq f(x)$, $\forall x \in X$. Fix any such simple function ϕ . Let ϕ admit the canonical representation $\phi = \sum_{i=1}^n a_i \chi_{A_i, X}$, where $n \in \mathbb{Z}_+$, $a_1, \dots, a_n \in (0, \infty) \subset \mathbb{R}$ are distinct, and $A_1, \dots, A_n \in \mathcal{B}$ are nonempty, pairwise disjoint, and $\mu_1(A_i) < \infty$, $i = 1, \dots, n$. Then, by Proposition 11.75, $\int_X \phi d\mu_1 = \sum_{i=1}^n a_i \mu_1(A_i) \in [0, +\infty) \subset \mathbb{R}$. By Proposition 11.83, $\int_X \phi d\mu_2 \leq \int_X f d\mu_2 < \infty$.

Claim 11.144.1 $\mu_2(A_i) < \infty$, $i = 1, \dots, n$.

Proof of claim: Suppose $\exists i_0 \in \{1, \dots, n\}$ such that $\mu_2(A_{i_0}) = \infty$. Since $\mathcal{X}_2$ is σ -finite, then, by Proposition 11.7, $\exists E \in \mathcal{B}$ with $E \subseteq A_{i_0}$ such that $\frac{1}{a_{i_0}} \int_X f d\mu_2 < \mu_2(E) < \infty$. Choose $\varphi : \mathcal{X}_2 \rightarrow [0, \infty) \subset \mathbb{R}$ by $\varphi = a_{i_0} \chi_{E, X}$. Clearly, φ is a simple function of $\mathcal{X}_2$ and $0 \leq \varphi(x) \leq f(x)$, $\forall x \in X$. By Proposition 11.75, $\int_X \varphi d\mu_2 = a_{i_0} \mu_2(E) > \int_X f d\mu_2$, which contradicts Definition 11.79. Hence, the claim is true. $\square$

By the claim, ϕ is a simple function of $\mathcal{X}_2$. By Proposition 11.75, $\int_X \phi d\mu_2 = \sum_{i=1}^n a_i \mu_2(A_i) \geq \sum_{i=1}^n a_i \mu_1(A_i) = \int_X \phi d\mu_1$. This implies that $\int_X f d\mu_2 \geq \int_X \phi d\mu_2 \geq \int_X \phi d\mu_1$. By the arbitrariness of ϕ , we have $\int_X f d\mu_2 \geq \int_X f d\mu_1$. Hence, in both cases, we have $0 \leq \int_X f d\mu_1 \leq \int_X f d\mu_2 \leq \infty$.

(ii) By Proposition 11.136, μ is a σ -finite measure on $(X, \mathcal{B})$. By Proposition 11.66, there exists a sequence of simple functions $(\phi_n)_{n=1}^\infty$, $\phi_n : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, such that $0 \leq \phi_n(x) \leq f(x) < \infty$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \phi_n = f$ a.e. in $\mathcal{X}$, where $\mathcal{X} := (X, \mathcal{B}, \mu)$. Since $\mu = \mu_1 + \mu_2$, then $\mu_i \leq \mu$, $i = 1, 2$. This implies that $(\phi_n)_{n=1}^\infty$ is sequence of simple functions of $\mathcal{X}_i$ and $\lim_{n \in \mathbb{N}} \phi_n = f$ a.e. in $\mathcal{X}_i$, $i = 1, 2$. By Fatou's Lemma 11.80 and Propositions 3.83 and 11.83, we have $0 \leq \int_X f d\mu \leq \liminf_{n \in \mathbb{N}} \int_X \phi_n d\mu \leq \limsup_{n \in \mathbb{N}} \int_X \phi_n d\mu \leq \sup_{n \in \mathbb{N}} \int_X \phi_n d\mu \leq \int_X f d\mu \leq \infty$. Then, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n d\mu \in [0, \infty] \subset \mathbb{R}_e$. By an argument that is similar to the above, we have $\int_X f d\mu_i = \lim_{n \in \mathbb{N}} \int_X \phi_n d\mu_i \in [0, \infty] \subset \mathbb{R}_e$, $i = 1, 2$.

Fix any $n \in \mathbb{N}$. Let ϕ_n admit the canonical representation $\phi_n = \sum_{i=1}^{m_n} a_{i,n} \chi_{A_{i,n}, X}$. By Proposition 11.75, $\int_X \phi_n d\mu = \sum_{i=1}^{m_n} a_{i,n} \mu(A_{i,n}) = \sum_{i=1}^{m_n} a_{i,n} \mu_1(A_{i,n}) + \sum_{i=1}^{m_n} a_{i,n} \mu_2(A_{i,n}) = \int_X \phi_n d\mu_1 + \int_X \phi_n d\mu_2 \in [0, \infty) \subset \mathbb{R}$. Then, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n d\mu = \lim_{n \in \mathbb{N}} (\int_X \phi_n d\mu_1 + \int_X \phi_n d\mu_2) = \int_X f d\mu_1 + \int_X f d\mu_2 \in [0, \infty] \subset \mathbb{R}_e$.

(iii) We will distinguish two exhaustive and mutually exclusive cases: Case 1: $a = 0$; Case 2: $a > 0$. Case 1: $a = 0$. Then, by Proposition 11.138, $\nu(E) = 0$, $\forall E \in \mathcal{B}$. This implies that, by Definition 11.70 Proposition 11.75, $0 = \int_X f d\nu = \int_X (af) d\mu_1$. Hence, the result holds. Case 2: $a > 0$. By Definition 11.79, $\int_X (af) d\mu_1 = \sup_{0 \leq \phi \leq af} \int_X \phi d\mu_1$,

where the supremum is over all simple functions $\phi : \mathcal{X}_1 \rightarrow [0, \infty) \subset \mathbb{R}$ such that $0 \leq \phi(x) \leq af(x) < \infty$, $\forall x \in X$. Clearly, ϕ is such a simple function if, and only if, $\phi/a : \tilde{\mathcal{X}} \rightarrow [0, \infty) \subset \mathbb{R}$ is a simple function of $\tilde{\mathcal{X}} := (X, \mathcal{B}, \nu)$ with $0 \leq \phi(x)/a \leq f(x)$, $\forall x \in X$. Then, $\int_X (af) d\mu_1 = \sup_{0 \leq \phi \leq af} \int_X \phi d\mu_1 = \sup_{0 \leq \phi/a \leq f} \int_X \phi d\mu_1 = \sup_{0 \leq \phi \leq f} \int_X (a\phi) d\mu_1 = \sup_{0 \leq \phi \leq f} \int_X \phi d\nu = \int_X f d\nu \in [0, \infty] \subset \mathbb{R}_e$, where the fourth equality follows from Proposition 11.75; and the fifth equality follows from Definition 11.79.

This completes the proof of the proposition. $\square$

Proposition 11.145 *Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $\mu_1, \mu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, and $\mu := \mu_1 + \mu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$. Assume that f is absolutely integrable over $\mathcal{X}_i := (X, \mathcal{B}, \mu_i)$, $i = 1, 2$. Then, f is absolutely integrable over $\mathcal{X} := (X, \mathcal{B}, \mu)$ and $\int_X f d\mu = \int_X f d\mu_1 + \int_X f d\mu_2 \in \mathcal{Z}$.*

Proof By Propositions 11.138 and 11.136, $\mathcal{P} \circ \mu \leq \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2 =: \nu$ and ν is σ -finite. By Proposition 11.144, $0 \leq \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu \leq \int_X \mathcal{P} \circ f d\nu = \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu_1 + \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu_2 < \infty$. Hence, f is absolutely integrable over $\mathcal{X}$.

By Proposition 11.66, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : \hat{\mathcal{X}} \rightarrow \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\|\varphi_n(x)\| \leq \|f(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\hat{\mathcal{X}}$, where $\hat{\mathcal{X}} := (X, \mathcal{B}, \nu)$. Since $\mathcal{P} \circ \mu \leq \nu$ and $\mathcal{P} \circ \mu_i \leq \nu$, $i = 1, 2$, then $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$, $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}_i$, $i = 1, 2$, and φ_n is a simple function of $\mathcal{X}$, $\mathcal{X}_1$, or $\mathcal{X}_2$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.131, we have $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu \in \mathcal{Z}$ and $\int_X f d\mu_i = \lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu_i \in \mathcal{Z}$, $i = 1, 2$.

$\forall n \in \mathbb{N}$, let φ_n admit canonical representation $\varphi_n = \sum_{i=1}^{m_n} w_{i,n} \chi_{A_{i,n}, X}$. By Proposition 11.125, we have $\int_X \varphi_n d\mu = \sum_{i=1}^{m_n} w_{i,n} \mu(A_{i,n}) = \sum_{i=1}^{m_n} w_{i,n} (\mu_1(A_{i,n}) + \mu_2(A_{i,n})) = \sum_{i=1}^{m_n} w_{i,n} \mu_1(A_{i,n}) + \sum_{i=1}^{m_n} w_{i,n} \mu_2(A_{i,n}) = \int_X \varphi_n d\mu_1 + \int_X \varphi_n d\mu_2 \in \mathcal{Z}$. Then, by Propositions 7.23, 3.66, and 3.67, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu = \lim_{n \in \mathbb{N}} (\int_X \varphi_n d\mu_1 + \int_X \varphi_n d\mu_2) = \lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu_1 + \lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu_2 = \int_X f d\mu_1 + \int_X f d\mu_2 \in \mathcal{Z}$. This completes the proof of the proposition. $\square$

Proposition 11.146 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite $\mathcal{Y}_1$ -valued measure space, where $\mathcal{Y}_1$ is a normed linear space over $\mathbb{K}$, $\mathcal{Y}_2$ and $\mathcal{Z}$ be Banach spaces over $\mathbb{K}$, $A \in B(\mathcal{Y}_1, \mathcal{Y}_2)$, $\nu := A\mu \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}_2)$ as defined in Proposition 11.138, $\mathcal{W}_2$ be a separable subspace of $B(\mathcal{Y}_2, \mathcal{Z})$, and $f : X \rightarrow \mathcal{W}_2$ be $\mathcal{B}$ -measurable. Assume that $\|A\| \mathcal{P} \circ f$ is integrable over $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. Then, f is absolutely integrable over $\tilde{\mathcal{X}} := (X, \mathcal{B}, \nu)$, fA is absolutely integrable over $\mathcal{X}$, and $\int_X f d\nu = \int_X (fA) d\mu \in \mathcal{Z}$.*

Proof By Propositions 11.38, 7.64, and 7.21, fA and $\mathcal{P} \circ (fA)$ are $\mathcal{B}$ -measurable. By Proposition 11.83, $0 \leq \int_X \mathcal{P} \circ (fA) d\mathcal{P} \circ \mu \leq$

$\int_X (\|A\| \mathcal{P} \circ f) d\mathcal{P} \circ \mu < \infty$. Thus, fA is absolutely integrable over $\mathcal{X}$. By Proposition 11.138, $\mathcal{P} \circ \nu \leq \|A\| \mathcal{P} \circ \mu$. By Proposition 11.144, $0 \leq \int_X \mathcal{P} \circ f d\mathcal{P} \circ \nu \leq \int_X \mathcal{P} \circ f d(\|A\| \mathcal{P} \circ \mu) = \int_X (\|A\| \mathcal{P} \circ f) d\mathcal{P} \circ \mu < \infty$. Hence, f is absolutely integrable over $\mathcal{X}$.

By Proposition 11.66, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : \mathcal{X} \rightarrow \mathcal{W}_2$, $\forall n \in \mathbb{N}$, such that $\|\varphi_n(x)\| \leq \|f(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$. Since $\mathcal{P} \circ \nu \leq \|A\| \mathcal{P} \circ \mu$, then $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\tilde{\mathcal{X}}$ and φ_n is a simple function of $\tilde{\mathcal{X}}$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.131, we have $\int_X f d\nu = \lim_{n \in \mathbb{N}} \int_X \varphi_n d\nu \in \mathcal{Z}$. Clearly, $\|\varphi_n(x)A\| \leq \|A\| \|\varphi_n(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Note that $fA : X \rightarrow \mathcal{W}_1$ and $\varphi_n A : X \rightarrow \mathcal{W}_1$, $\forall n \in \mathbb{N}$, where $\mathcal{W}_1 := \{F \in B(\mathcal{Y}_1, \mathcal{Z}) \mid F = F_1 A, F_1 \in \mathcal{W}_2\}$ is a separable subspace of $B(\mathcal{Y}_1, \mathcal{Z})$. By Proposition 11.52, we have $\lim_{n \in \mathbb{N}} \varphi_n A = fA$ a.e. in $\mathcal{X}$. By Lebesgue Dominated Convergence Theorem 11.131, $\int_X (fA) d\mu = \lim_{n \in \mathbb{N}} \int_X (\varphi_n A) d\mu \in \mathcal{Z}$. By Proposition 11.125, we have $\int_X (\varphi_n A) d\mu = \int_X \varphi_n d\nu \in \mathcal{Z}$, $\forall n \in \mathbb{N}$. Then, $\int_X (fA) d\mu = \int_X f d\nu \in \mathcal{Z}$. This completes the proof of the proposition. $\square$

Proposition 11.147 *Let $n, m \in \mathbb{Z}_+$, $(X, \mathcal{B})$ be a measurable space, and $\mu_{j,i}$ be a σ -finite $B(\mathcal{Y}_i, \mathcal{Z}_j)$ -valued measure on $(X, \mathcal{B})$, where $\mathcal{Y}_i$ is a normed linear space over $\mathbb{K}$ and $\mathcal{Z}_j$ is a Banach space over $\mathbb{K}$, $i = 1, \dots, m$, $j = 1, \dots, n$. Then, there exists a unique σ -finite $B(\mathcal{Y}, \mathcal{Z})$ -valued measure*

$$\mu \text{ on } (X, \mathcal{B}) \text{ such that } \mu(E) = \begin{bmatrix} \mu_{1,1}(E) & \cdots & \mu_{1,m}(E) \\ \vdots & & \vdots \\ \mu_{n,1}(E) & \cdots & \mu_{n,m}(E) \end{bmatrix} \in B(\mathcal{Y}, \mathcal{Z}),$$

$\forall E \in \bigcap_{i=1}^m \bigcap_{j=1}^n \text{dom}(\mu_{j,i})$, where $\mathcal{Y} := \prod_{i=1}^m \mathcal{Y}_i$ and $\mathcal{Z} := \prod_{j=1}^n \mathcal{Z}_j$. Furthermore, $\max_{\substack{i \in \{1, \dots, m\} \\ j \in \{1, \dots, n\}}} \mathcal{P} \circ \mu_{j,i} \leq \mathcal{P} \circ \mu \leq \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}$. μ will be

$$\text{called the vector measure and denoted by } \begin{bmatrix} \mu_{1,1} & \cdots & \mu_{1,m} \\ \vdots & & \vdots \\ \mu_{n,1} & \cdots & \mu_{n,m} \end{bmatrix}.$$

Proof Consider the special case where $\mu_{j,i}$'s are finite. Then, μ is uniquely defined since $\text{dom}(\mu_{j,i}) = \mathcal{B}$, $i = 1, \dots, m$, $j = 1, \dots, n$.

$$\mu(\emptyset) = \begin{bmatrix} \mu_{1,1}(\emptyset) & \cdots & \mu_{1,m}(\emptyset) \\ \vdots & & \vdots \\ \mu_{n,1}(\emptyset) & \cdots & \mu_{n,m}(\emptyset) \end{bmatrix} = \begin{bmatrix} \vartheta_{B(\mathcal{Y}_1, \mathcal{Z}_1)} & \cdots & \vartheta_{B(\mathcal{Y}_m, \mathcal{Z}_1)} \\ \vdots & & \vdots \\ \vartheta_{B(\mathcal{Y}_1, \mathcal{Z}_n)} & \cdots & \vartheta_{B(\mathcal{Y}_m, \mathcal{Z}_n)} \end{bmatrix} = \vartheta_{B(\mathcal{Y}, \mathcal{Z})}.$$

$\forall$ pairwise disjoint $(E_k)_{k=1}^\infty \subseteq \mathcal{B}$, we have $\sum_{k=1}^\infty \|\mu(E_k)\| \leq \sum_{k=1}^\infty \|h\| (\sum_{i=1}^m \sum_{j=1}^n \|\mu_{j,i}(E_k)\|^2)^{1/2} \leq \sum_{k=1}^\infty \sum_{i=1}^m \sum_{j=1}^n \|\mu_{j,i}(E_k)\| = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^\infty \|\mu_{j,i}(E_k)\| < \infty$, where the first inequality follows from Proposition 11.139 with h being the bounded linear function defined in Proposition 11.139; and the last inequality follows from the fact that $\mu_{j,i}$ is a finite $B(\mathcal{Y}_i, \mathcal{Z}_j)$ -valued measure on $(X, \mathcal{B})$ and therefore is a finite $B(\mathcal{Y}_i, \mathcal{Z}_j)$ -valued pre-measure on $(X, \mathcal{B})$, $i = 1, \dots, m$, $j = 1, \dots, n$. We

also have $\sum_{k=1}^{\infty} \mu(E_k) = \sum_{k=1}^{\infty} \begin{bmatrix} \mu_{1,1}(E_k) & \cdots & \mu_{1,m}(E_k) \\ \vdots & & \vdots \\ \mu_{n,1}(E_k) & \cdots & \mu_{n,m}(E_k) \end{bmatrix} =$
 $\begin{bmatrix} \sum_{k=1}^{\infty} \mu_{1,1}(E_k) & \cdots & \sum_{k=1}^{\infty} \mu_{1,m}(E_k) \\ \vdots & & \vdots \\ \sum_{k=1}^{\infty} \mu_{n,1}(E_k) & \cdots & \sum_{k=1}^{\infty} \mu_{n,m}(E_k) \end{bmatrix} = \mu(E) \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, where
 $E := \bigcup_{k=1}^{\infty} E_k \in \mathcal{B}$; the second equality follows from Propositions 3.66, 3.67, and Proposition 11.139; and the third equality follows from the fact that $\mu_{j,i}$ is a finite $\mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j)$ -valued pre-measure on $(X, \mathcal{B})$, $i = 1, \dots, m$, $j = 1, \dots, n$. This shows that μ is a $\mathcal{Y}$ -valued pre-measure on $(X, \mathcal{B})$. Observe that, $\forall s \in \{1, \dots, m\}$, $\forall t \in \{1, \dots, n\}$, $\forall E \in \mathcal{B}$,

$$\begin{aligned} \mathcal{P} \circ \mu_{t,s}(E) &= \sup_{k \in \mathbb{Z}_+, (E_i)_{i=1}^k \subseteq \mathcal{B}, E = \bigcup_{i=1}^k E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq k} \sum_{i=1}^k \|\mu_{t,s}(E_i)\| \\ &= \sup_{k \in \mathbb{Z}_+, (E_i)_{i=1}^k \subseteq \mathcal{B}, E = \bigcup_{i=1}^k E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq k} \sum_{i=1}^k \|p_{t,s} \mu(E_i)\| \\ &\leq \sup_{k \in \mathbb{Z}_+, (E_i)_{i=1}^k \subseteq \mathcal{B}, E = \bigcup_{i=1}^k E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq k} \sum_{i=1}^k \|\mu(E_i)\| \\ &= \mathcal{P} \circ \mu(E) \end{aligned}$$

where the first equality follows from Definition 11.108; the second equality and the first inequality follow from Propositions 11.139 and 7.64; and the last equality follows from Definition 11.99. Observe also that, $\forall E \in \mathcal{B}$,

$$\begin{aligned} \mathcal{P} \circ \mu(E) &= \sup_{k \in \mathbb{Z}_+, (E_i)_{i=1}^k \subseteq \mathcal{B}, E = \bigcup_{i=1}^k E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq k} \sum_{i=1}^k \|\mu(E_i)\| \\ &\leq \sup_{k \in \mathbb{Z}_+, (E_i)_{i=1}^k \subseteq \mathcal{B}, E = \bigcup_{i=1}^k E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq k} \sum_{i=1}^k \sum_{s=1}^m \sum_{t=1}^n \|\mu_{t,s}(E_i)\| \\ &\leq \sum_{s=1}^m \sum_{t=1}^n \sup_{k \in \mathbb{Z}_+, (E_i)_{i=1}^k \subseteq \mathcal{B}, E = \bigcup_{i=1}^k E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq k} \sum_{i=1}^k \|\mu_{t,s}(E_i)\| \\ &= \sum_{s=1}^m \sum_{t=1}^n \mathcal{P} \circ \mu_{t,s}(E) \leq \sum_{s=1}^m \sum_{t=1}^n \mathcal{P} \circ \mu_{t,s}(X) < \infty \end{aligned}$$

where the first equality follows from Definition 11.99; the first inequality follows from Propositions 11.139, 7.64, and 7.22; the second inequality follows from Proposition 3.81; and the second equality follows from Definition 11.108. Then, μ is a finite $\mathcal{Y}$ -valued pre-measure on $(X, \mathcal{B})$ and $\max_{\substack{i \in \{1, \dots, m\} \\ j \in \{1, \dots, n\}}} \mathcal{P} \circ \mu_{j,i} \leq \mathcal{P} \circ \mu \leq \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}$. Then, $(X, \mathcal{B}, \mu)$ is a finite $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$ -valued measure space. This completes the proof for the special case.

Consider the general case where $\mu_{j,i}$'s are σ -finite. Then, $\exists (X_k)_{k=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{k=1}^\infty X_k$ and $\mathcal{P} \circ \mu_{j,i}(X_k) < \infty$, $\forall i \in \{1, \dots, m\}$, $\forall j \in \{1, \dots, n\}$, $\forall k \in \mathbb{N}$. Without loss of generality, assume that $(X_k)_{k=1}^\infty$ is pairwise disjoint. Let $\mathcal{X}_{j,i,k} := (X_k, \mathcal{B}_k, \mu_{j,i,k})$ be the finite $B(\mathcal{Y}_i, \mathcal{Z}_j)$ -valued measure subspace of $(X, \mathcal{B}, \mu_{j,i})$, $\forall k \in \mathbb{N}$, $\forall i \in \{1, \dots, m\}$, $\forall j \in \{1, \dots, n\}$. Then, by the special case, there exists a unique finite $B(\mathcal{Y}, \mathcal{Z})$ -valued measure $\bar{\mu}_k$ on $(X_k, \mathcal{B}_k)$ such that $\bar{\mu}_k(E) = \begin{bmatrix} \mu_{1,1,k}(E) & \cdots & \mu_{1,m,k}(E) \\ \vdots & & \vdots \\ \mu_{n,1,k}(E) & \cdots & \mu_{n,m,k}(E) \end{bmatrix} = \begin{bmatrix} \mu_{1,1}(E) & \cdots & \mu_{1,m}(E) \\ \vdots & & \vdots \\ \mu_{n,1}(E) & \cdots & \mu_{n,m}(E) \end{bmatrix} \in B(\mathcal{Y}, \mathcal{Z})$, $\forall E \in \mathcal{B}_k$. Let $\bar{\mathcal{X}}_k := (X_k, \mathcal{B}_k, \bar{\mu}_k)$ be the finite $B(\mathcal{Y}, \mathcal{Z})$ -valued measure space, $\forall k \in \mathbb{N}$. By Propositions 4.31, 7.22, and 7.66, $B(\mathcal{Y}, \mathcal{Z})$ is a Banach space over $\mathbb{K}$. By Proposition 11.118, there is a unique σ -finite $B(\mathcal{Y}, \mathcal{Z})$ -valued measure space $\mathcal{X} := (X, \mathcal{B}, \mu)$ on X such that $\bar{\mathcal{X}}_k$ is the finite $B(\mathcal{Y}, \mathcal{Z})$ -valued measure subspace of $\mathcal{X}$, $\forall k \in \mathbb{N}$. $\forall s \in \{1, \dots, m\}$, $\forall t \in \{1, \dots, n\}$, $\forall E \in \mathcal{B}$, we have $\mathcal{P} \circ \mu_{t,s}(E) = \sum_{k=1}^\infty \mathcal{P} \circ \mu_{t,s}(E \cap X_k) = \sum_{k=1}^\infty \mathcal{P} \circ \mu_{t,s,k}(E \cap X_k) \leq \sum_{k=1}^\infty \mathcal{P} \circ \bar{\mu}_k(E \cap X_k) = \sum_{k=1}^\infty \mathcal{P} \circ \mu(E \cap X_k) = \mathcal{P} \circ \mu(E) \leq \sum_{k=1}^\infty \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i,k}(E \cap X_k) = \sum_{k=1}^\infty \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}(E \cap X_k) = \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}(E)$, where the second equality follows from Proposition 11.115; the first inequality follows from the special case; the third equality follows from Proposition 11.115; and the second inequality follows from the special case; and the fifth equality follows from Proposition 11.115. Then, $\max_{\substack{i \in \{1, \dots, m\} \\ j \in \{1, \dots, n\}}} \mathcal{P} \circ \mu_{j,i} \leq \mathcal{P} \circ \mu \leq \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}$. $\forall E \in \bigcap_{i=1}^m \bigcap_{j=1}^n \text{dom}(\mu_{j,i})$, $\mathcal{P} \circ \mu(E) \leq \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}(E) < \infty$ and $\mu(E) \in B(\mathcal{Y}, \mathcal{Z})$. This implies that

$$\begin{aligned}
\mu(E) &= \sum_{k=1}^\infty \mu(E \cap X_k) = \sum_{k=1}^\infty \bar{\mu}_k(E \cap X_k) \\
&= \sum_{k=1}^\infty \begin{bmatrix} \mu_{1,1}(E \cap X_k) & \cdots & \mu_{1,m}(E \cap X_k) \\ \vdots & & \vdots \\ \mu_{n,1}(E \cap X_k) & \cdots & \mu_{n,m}(E \cap X_k) \end{bmatrix} \\
&= \begin{bmatrix} \sum_{k=1}^\infty \mu_{1,1}(E \cap X_k) & \cdots & \sum_{k=1}^\infty \mu_{1,m}(E \cap X_k) \\ \vdots & & \vdots \\ \sum_{k=1}^\infty \mu_{n,1}(E \cap X_k) & \cdots & \sum_{k=1}^\infty \mu_{n,m}(E \cap X_k) \end{bmatrix} \\
&= \begin{bmatrix} \mu_{1,1}(E) & \cdots & \mu_{1,m}(E) \\ \vdots & & \vdots \\ \mu_{n,1}(E) & \cdots & \mu_{n,m}(E) \end{bmatrix}
\end{aligned}$$

where the first equality follows from Definition 11.108; the second equality follows from Proposition 11.115; the third equality follows from the special case; the fourth equality follows from Propositions 11.139, 3.66, and 3.67; the fifth equality follows from Definition 11.108. This completes the proof

of the proposition. $\square$

Proposition 11.148 *Let $\mathcal{X}_i := (X, \mathcal{B}, \mu_i)$ be a σ -finite $\mathcal{Y}_i$ -valued measure space, where $\mathcal{Y}_i$ is a Banach space over $\mathbb{K}$, $\mathcal{Z}_j$ be a Banach space over $\mathbb{K}$, $\mathcal{W}_{j,i} \subseteq B(\mathcal{Y}_i, \mathcal{Z}_j)$ be a separable subspace, $f_{j,i} : X \rightarrow \mathcal{W}_{j,i}$, $i = 1, 2$, $j = 1, 2$, $\mu := \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ be the vector measure as defined in Proposition 11.147, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be the σ -finite $\mathcal{Y}_1 \times \mathcal{Y}_2$ -valued measure space, and $F : X \rightarrow B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$ be defined by $F(x) = \begin{bmatrix} f_{1,1}(x) & f_{1,2}(x) \\ f_{2,1}(x) & f_{2,2}(x) \end{bmatrix}$, $\forall x \in X$. Assume that F is absolutely integrable over $\mathcal{X}$. Then, $f_{j,i}$, $j = 1, 2$ are absolutely integrable over $\mathcal{X}_i$, $i = 1, 2$. In this case, we have $\int_X F d\mu = \begin{bmatrix} \int_X f_{1,1} d\mu_1 + \int_X f_{1,2} d\mu_2 \\ \int_X f_{2,1} d\mu_1 + \int_X f_{2,2} d\mu_2 \end{bmatrix} \in \mathcal{Z}_1 \times \mathcal{Z}_2$.*

Proof Let F be absolutely integrable over $\mathcal{X}$. Fix any $i \in \{1, 2\}$ and any $j \in \{1, 2\}$. Then, by Propositions 11.139 and 11.38, $f_{j,i} = p_{j,i}F$ is $\mathcal{B}$ -measurable, where $p_{j,i}$ is the bounded linear function as defined in Proposition 11.139. By Propositions 7.21 and 11.38, $\mathcal{P} \circ f_{j,i}$ is $\mathcal{B}$ -measurable. $\forall x \in X$, $\mathcal{P} \circ f_{j,i}(x) \leq \|p_{j,i}\| \mathcal{P} \circ F(x) \leq \mathcal{P} \circ F(x)$. Then, by Propositions 11.83, 11.147, and 11.144, $0 \leq \int_X \mathcal{P} \circ f_{j,i} d\mathcal{P} \circ \mu_i \leq \int_X \mathcal{P} \circ f_{j,i} d\mathcal{P} \circ \mu \leq \int_X \mathcal{P} \circ F d\mathcal{P} \circ \mu < \infty$. Hence, $f_{j,i}$ is absolutely integrable over $\mathcal{X}_i$. By Proposition 11.132, $f_{j,i}$ is integrable over $\mathcal{X}_i$ and $\int_X f_{j,i} d\mu_i \in B(\mathcal{Y}_i, \mathcal{Z}_j)$.

By Propositions 7.22, 4.4, and 3.28, $\prod_{i=1}^2 \prod_{j=1}^2 \mathcal{W}_{j,i} \subseteq \prod_{i=1}^2 \prod_{j=1}^2 B(\mathcal{Y}_i, \mathcal{Z}_j)$ is a separable subspace. Let h be the bounded linear function as defined in Proposition 11.139 for $n = m = 2$. Then, $\mathcal{W} := h(\prod_{i=1}^2 \prod_{j=1}^2 \mathcal{W}_{j,i}) \subseteq B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2)$ is a separable subspace. By Proposition 11.66, $\exists$ sequence of simple functions $(\psi_n)_{n=1}^\infty$, $\psi_n : \mathcal{X} \rightarrow \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\lim_{n \in \mathbb{N}} \psi_n = F$ a.e. in $\mathcal{X}$, $\|\psi_n(x)\| \leq \|F(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Fix any $i \in \{1, 2\}$ and any $j \in \{1, 2\}$. Let $\varphi_{j,i,n} := p_{j,i}\psi_n$. Then, $\varphi_{j,i,n} : \mathcal{X} \rightarrow p_{j,i}(\mathcal{W}) \subseteq B(\mathcal{Y}_i, \mathcal{Z}_j)$ is a simple function. By Proposition 11.52, $\lim_{n \in \mathbb{N}} \varphi_{j,i,n} = p_{j,i}F = f_{j,i}$ a.e. in $\mathcal{X}$. Since $\mathcal{P} \circ \mu \geq \mathcal{P} \circ \mu_i$, we have $\varphi_{j,i,n} : \mathcal{X}_i \rightarrow p_{j,i}(\mathcal{W}) \subseteq B(\mathcal{Y}_i, \mathcal{Z}_j)$ is a simple function, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \varphi_{j,i,n} = f_{j,i}$ a.e. in $\mathcal{X}_i$, and F is absolutely integrable over $\mathcal{X}_i$, by Proposition 11.144. Note that $\|\varphi_{j,i,n}(x)\| \leq \|p_{j,i}\| \|\psi_n(x)\| \leq \|F(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.131, $\int_X f_{j,i} d\mu_i = \lim_{n \in \mathbb{N}} \int_X \varphi_{j,i,n} d\mu_i \in \mathcal{Z}_j$. By Lebesgue Dominated Convergence Theorem 11.131,

$$\begin{aligned} \int_X F d\mu &= \lim_{n \in \mathbb{N}} \int_X \psi_n d\mu = \lim_{n \in \mathbb{N}} \begin{bmatrix} \int_X \varphi_{1,1,n} d\mu_1 + \int_X \varphi_{1,2,n} d\mu_2 \\ \int_X \varphi_{2,1,n} d\mu_1 + \int_X \varphi_{2,2,n} d\mu_2 \end{bmatrix} \\ &= \begin{bmatrix} \int_X f_{1,1} d\mu_1 + \int_X f_{1,2} d\mu_2 \\ \int_X f_{2,1} d\mu_1 + \int_X f_{2,2} d\mu_2 \end{bmatrix} \in \mathcal{Z}_1 \times \mathcal{Z}_2 \end{aligned}$$

where the second equality follows from Proposition 11.125; and the third equality follows from Propositions 3.67, 3.66, and 7.23.

This completes the proof of the proposition. $\square$

Theorem 11.149 (Fatou's Lemma) *Let $(X, \mathcal{B})$ be a measurable space, $(\mu_n)_{n=1}^\infty$ be a sequence of measures on $(X, \mathcal{B})$, μ be a measure on $(X, \mathcal{B})$, $f_n : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $f : X \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable. Assume that $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$, where $\mathcal{X} := (X, \mathcal{B}, \mu)$, and $\lim_{n \in \mathbb{N}} \mu_n(E) = \mu(E)$, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$. Then, $0 \leq \int_X f \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \leq \infty$.*

Proof By Definition 11.79, $\int_X f \, d\mu = \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu$, where the supremum is over all simple functions $\phi : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$ that satisfies $0 \leq \phi(x) \leq f(x)$, $\forall x \in X$. Fix any such simple function ϕ . Then, $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\phi(x) \leq M$, $\forall x \in X$ and $\mu(\bar{E}) := \mu(\{x \in X \mid \phi(x) > 0\}) < \infty$. By $\lim_{n \in \mathbb{N}} \mu_n(\bar{E}) = \mu(\bar{E}) < \infty$, $\exists n_0 \in \mathbb{N}$ such that $\mu_n(\bar{E}) < \infty$, $\forall n \in \mathbb{N}$ with $n_0 \leq n$. Thus, $\forall n \in \mathbb{N}$ with $n_0 \leq n$, ϕ is a simple function of $\mathcal{X}_n := (X, \mathcal{B}, \mu_n)$. By Proposition 11.75, $\int_X \phi \, d\mu =: c_1 \in [0, \infty) \subset \mathbb{R}$. Let $\hat{E} := \{x \in X \mid (f_n(x))_{n=1}^\infty \text{ does not converge to } f(x)\}$. By the assumption $\hat{E} \in \mathcal{B}$ and $\mu(\hat{E}) = 0$.

Fix any $\epsilon \in (0, 1) \subset \mathbb{R}$. Define $E_n := \left\{ x \in X \setminus \hat{E} \mid f_k(x) \geq (1 - \frac{\epsilon}{2+2c_1})\phi(x), \forall k \geq n, k \in \mathbb{N} \right\}$, $\forall n \in \mathbb{N}$. By Propositions 7.23, 11.38, and 11.39, $E_n \in \mathcal{B}$, $\forall n \in \mathbb{N}$. Clearly, $\forall x \in X \setminus (\hat{E} \cup \bar{E})$, $x \in E_n$, $\forall n \in \mathbb{N}$. Then, $X \setminus E_n \subseteq \hat{E} \cup \bar{E}$, $\forall n \in \mathbb{N}$. $\mu(X \setminus E_n) \leq \mu(\hat{E} \cup \bar{E}) \leq \mu(\bar{E}) < \infty$, $\forall n \in \mathbb{N}$. $\forall x \in \bar{E} \setminus \hat{E}$, then $\lim_{n \in \mathbb{N}} f_n(x) = f(x) \geq \phi(x) > 0$ and there exists $n_x \in \mathbb{N}$ such that $x \in E_{n_x}$. Then, $\bigcup_{n=1}^\infty E_n = X \setminus \hat{E}$. Clearly, $E_n \subseteq E_{n+1}$, $\forall n \in \mathbb{N}$. Then, $X \setminus E_n \supseteq X \setminus E_{n+1}$, $\forall n \in \mathbb{N}$, and $\bigcap_{n=1}^\infty (X \setminus E_n) = X \setminus (\bigcup_{n=1}^\infty E_n) = \hat{E}$. By Proposition 11.5, we have $\lim_{n \in \mathbb{N}} \mu(X \setminus E_n) = \mu(\hat{E}) = 0$. Then, $\exists n_1 \in \mathbb{N}$ with $n_0 \leq n_1$ such that $\mu(X \setminus E_{n_1}) < \frac{\epsilon}{4+4M}$. By the assumption that $\lim_{i \in \mathbb{N}} \mu_i(X \setminus E_{n_1}) = \mu(X \setminus E_{n_1}) < \frac{\epsilon}{4+4M}$. Then, there exists $n_2 \in \mathbb{N}$ with $n_1 \leq n_2$, $\forall n \in \mathbb{N}$ with $n_2 \leq n$, we have $\mu_n(X \setminus E_{n_1}) < \frac{\epsilon}{2+2M}$. Fix any $n \in \mathbb{N}$ with $n_2 \leq n$, $X \setminus E_{n_1} \supseteq X \setminus E_n$, which implies that $0 \leq \mu_n(X \setminus E_n) \leq \mu_n(X \setminus E_{n_1}) < \frac{\epsilon}{2+2M}$. Thus, by Proposition 11.83, $\int_X f_n \, d\mu_n \geq \int_{E_n} f_n \, d\mu_n \geq \int_{E_n} (1 - \frac{\epsilon}{2+2c_1})\phi \, d\mu_n = (1 - \frac{\epsilon}{2+2c_1})(\int_X \phi \, d\mu_n - \int_{X \setminus E_n} \phi \, d\mu_n) \geq (1 - \frac{\epsilon}{2+2c_1})(\int_X \phi \, d\mu_n - M\mu_n(X \setminus E_n)) > (1 - \frac{\epsilon}{2+2c_1})\int_X \phi \, d\mu_n - \epsilon/2$. By the assumption that $\lim_{n \in \mathbb{N}} \mu_n(E) = \mu(E)$, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$, and Proposition 11.75, we have $\lim_{n \in \mathbb{N}} \int_X \phi \, d\mu_n = \int_X \phi \, d\mu = c_1$. Then, $\liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \geq (1 - \frac{\epsilon}{2+2c_1})c_1 - \epsilon/2 > c_1 - \epsilon = \int_X \phi \, d\mu - \epsilon$. By the arbitrariness of ϵ , we have $\liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \geq \int_X \phi \, d\mu$. By the arbitrariness of ϕ , we have $\liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \geq \int_X f \, d\mu$.

This completes the proof of the lemma. $\square$

Theorem 11.150 (Lebesgue Dominated Convergence) *Let $(X, \mathcal{B})$ be a measurable space, $(\mu_n)_{n=1}^\infty$ be a sequence of measures on $(X, \mathcal{B})$, μ be a measure on $(X, \mathcal{B})$, $f_n : X \rightarrow \mathbb{R}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $f : X \rightarrow \mathbb{R}$*

be $\mathcal{B}$ -measurable, $g_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $g : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. Assume that

(i) $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$ and $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$, where $\mathcal{X} := (X, \mathcal{B}, \mu)$;

(ii) $|f_n(x)| \leq g_n(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$;

(iii) $\int_X g_n d\mu_n \in \mathbb{R}$, $\forall n \in \mathbb{N}$, and $\int_X g d\mu = \lim_{n \in \mathbb{N}} \int_X g_n d\mu_n \in \mathbb{R}$;

(iv) $\lim_{n \in \mathbb{N}} \mu_n(E) = \mu(E)$, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$.

Then, f is absolutely integrable over $\mathcal{X}$, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu_n \in \mathbb{R}$, and $0 \leq \left| \int_X f d\mu \right| \leq \int_X g d\mu < \infty$.

Proof By Proposition 11.92, f_n is absolutely integrable over $\mathcal{X}_n := (X, \mathcal{B}, \mu_n)$ and $\int_X f_n d\mu_n \in \mathbb{R}$. Clearly, $|f(x)| \leq g(x)$ a.e. in $\mathcal{X}$. By Proposition 11.83, f is absolutely integrable over $\mathcal{X}$. By Proposition 11.92, $\int_X f d\mu \in \mathbb{R}$ and $0 \leq \left| \int_X f d\mu \right| \leq \int_X \mathcal{P} \circ f d\mu \leq \int_X g d\mu < \infty$.

By Propositions 7.23, 11.52, and 11.53, $\lim_{n \in \mathbb{N}} (g_n - f_n) = g - f$ a.e. in $\mathcal{X}$ and $\lim_{n \in \mathbb{N}} (g_n + f_n) = g + f$ a.e. in $\mathcal{X}$. Clearly, $(g_n - f_n)(x) \geq 0$ and $(g_n + f_n)(x) \geq 0$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Propositions 7.23, 11.52, and 11.53, $\lim_{n \in \mathbb{N}} (g_n - f_n) = \lim_{n \in \mathbb{N}} (g_n - f_n) \vee 0 = (g - f) \vee 0 = g - f$ a.e. in $\mathcal{X}$ and $\lim_{n \in \mathbb{N}} (g_n + f_n) = \lim_{n \in \mathbb{N}} (g_n + f_n) \vee 0 = (g + f) \vee 0 = g + f$ a.e. in $\mathcal{X}$. By Fatou's Lemma 11.149 and Propositions 11.92 and 3.83, we have $0 \leq \int_X ((g - f) \vee 0) d\mu = \int_X (g - f) d\mu = \int_X g d\mu - \int_X f d\mu \leq \liminf_{n \in \mathbb{N}} \int_X (g_n - f_n) d\mu_n = \liminf_{n \in \mathbb{N}} (\int_X g_n d\mu_n - \int_X f_n d\mu_n) = \int_X g d\mu - \limsup_{n \in \mathbb{N}} \int_X f_n d\mu_n$ and $0 \leq \int_X ((g + f) \vee 0) d\mu = \int_X (g + f) d\mu = \int_X g d\mu + \int_X f d\mu \leq \liminf_{n \in \mathbb{N}} \int_X (g_n + f_n) d\mu_n = \liminf_{n \in \mathbb{N}} (\int_X g_n d\mu_n + \int_X f_n d\mu_n) = \int_X g d\mu + \liminf_{n \in \mathbb{N}} \int_X f_n d\mu_n$. Then, by Proposition 3.83, we have $\int_X f d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n d\mu_n \leq \limsup_{n \in \mathbb{N}} \int_X f_n d\mu_n \leq \int_X f d\mu \in \mathbb{R}$. This further implies that $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu_n \in \mathbb{R}$. This completes the proof of the theorem. $\square$

In the preceding two results, we work with measure spaces which may not be σ -finite, and we assume a mode of convergence for the measures to be set-wise convergence rather than the stronger "norm" convergence alluded to in Definition 11.143. The next result shows that Lebesgue Dominated convergence Theorem holds in general under the "norm" convergence assumption on the Banach space valued measures.

Theorem 11.151 (Lebesgue Dominated Convergence) *Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $(\mu_n)_{n=1}^\infty \subseteq \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$, $\mu \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, $\bar{\mathcal{W}} := \overline{\text{span}(\{z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \ni z = wy\})} \subseteq \mathcal{Z}$ be the Banach subspace of $\mathcal{Z}$, $f_n : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $g_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$, $g : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. Assume that*

- (i) $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$ and $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$, where $\mathcal{X} := (X, \mathcal{B}, \mu)$;
- (ii) $\|f_n(x)\| \leq g_n(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$;
- (iii) $\int_X g_n d\mathcal{P} \circ \mu_n \in \mathbb{R}$, $\forall n \in \mathbb{N}$, and $\int_X g d\mathcal{P} \circ \mu = \lim_{n \in \mathbb{N}} \int_X g_n d\mathcal{P} \circ \mu_n \in \mathbb{R}$;
- (iv) $\lim_{n \in \mathbb{N}} \mu_n = \mu$ as defined in Definition 11.143.

Then, f is absolutely integrable over $\mathcal{X}$, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu_n \in \bar{\mathcal{W}} \subseteq \mathcal{Z}$, and $0 \leq \|\int_X f d\mu\| \leq \int_X g d\mathcal{P} \circ \mu < \infty$.

Proof By Proposition 11.132, f_n is absolutely integrable over $\mathcal{X}_n := (X, \mathcal{B}, \mu_n)$ and $\int_X f_n d\mu_n \in \bar{\mathcal{W}}$. Clearly, $\|f(x)\| \leq g(x)$ a.e. in $\mathcal{X}$. By Proposition 11.83, f is absolutely integrable over $\mathcal{X}$. By Proposition 11.132, $\int_X f d\mu \in \bar{\mathcal{W}}$ and $0 \leq \|\int_X f d\mu\| \leq \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu \leq \int_X g d\mathcal{P} \circ \mu < \infty$.

We will distinguish two exhaustive and mutually exclusives cases: Case 1: $\mathcal{P} \circ \mu(X) < \infty$; Case 2: $\mathcal{P} \circ \mu(X) = \infty$.

Case 1: $\mathcal{P} \circ \mu(X) < \infty$. By $\lim_{n \in \mathbb{N}} \mu_n = \mu$, we have $\lim_{n \in \mathbb{N}} \mathcal{P} \circ \mu_n(X) = \mathcal{P} \circ \mu(X) < \infty$. Without loss of generality, we may assume that $\mathcal{P} \circ \mu_n(X) < \infty$, $\forall n \in \mathbb{N}$. Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. By Proposition 11.84, $\exists \delta \in (0, \epsilon/6] \subset \mathbb{R}$ such that $\forall A \in \mathcal{B}$ with $\mathcal{P} \circ \mu(A) < \delta$, we have $\int_A g d\mathcal{P} \circ \mu < \epsilon/6$. By Egoroff's Theorem 11.55, $\exists \bar{E} \in \mathcal{B}$ with $\mathcal{P} \circ \mu(\bar{E}) < \delta/2$ such that $\left(f_n|_{X \setminus \bar{E}}\right)_{n=1}^{\infty}$ converges uniformly to $f|_{X \setminus \bar{E}}$. Then, $\exists n_0 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n_0 \leq n$, $\|f_n(x) - f(x)\| < \frac{\epsilon}{4\mathcal{P} \circ \mu(X) + 1}$, $\forall x \in X \setminus \bar{E}$. By Proposition 11.5, there exists $M \in \mathbb{N}$ such that $E_1 := \{x \in X \mid g(x) \geq M\} \in \mathcal{B}$ and $\mathcal{P} \circ \mu(E_1) < \delta/2$. Let $E := \bar{E} \cup E_1 \in \mathcal{B}$. Then, $\mathcal{P} \circ \mu(E) < \delta$. Hence, $\int_E g d\mathcal{P} \circ \mu < \epsilon/6$. By Lebesgue Dominated Convergence Theorem 11.150, we have $\lim_{n \in \mathbb{N}} \int_E g_n d\mathcal{P} \circ \mu_n = \lim_{n \in \mathbb{N}} \int_X (g_n \chi_{E, X}) d\mathcal{P} \circ \mu_n = \int_X (g \chi_{E, X}) d\mathcal{P} \circ \mu = \int_E g d\mathcal{P} \circ \mu < \epsilon/6$. Then, $\exists n_1 \in \mathbb{N}$ with $n_0 \leq n_1$, $\forall n \in \mathbb{N}$ with $n_1 \leq n$, we have $0 \leq \int_E g_n d\mathcal{P} \circ \mu_n < \epsilon/3$. By the assumption (iv), $\lim_{n \in \mathbb{N}} \mathcal{P} \circ (\mu_n - \mu)(X) = 0$. This implies that $\exists n_2 \in \mathbb{N}$ with $n_1 \leq n_2$, $\forall n \in \mathbb{N}$ with $n_2 \leq n$, we have $\mathcal{P} \circ (\mu_n - \mu)(X) < \frac{\epsilon}{4M}$ and $\mathcal{P} \circ \mu_n(X) < \mathcal{P} \circ \mu(X) + 1/4$.

Fix any $n \in \mathbb{N}$ with $n_2 \leq n$. Clearly, $f|_{X \setminus E}$ is absolutely integrable over $\mathcal{E}_n$ and $\int_{X \setminus E} f d\mu_n \in \mathcal{Z}$, since $\|f(x)\| \leq M$ a.e. in $\mathcal{E}_n$, where $\mathcal{E}_n := (X \setminus E, \mathcal{B}_{X \setminus E}, \mu_{n, X \setminus E})$ is the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}_n$. We have

$$\begin{aligned}
 & \left\| \int_X f_n d\mu_n - \int_X f d\mu \right\| \\
 &= \left\| \int_E f_n d\mu_n + \int_{X \setminus E} f_n d\mu_n - \int_{X \setminus E} f d\mu_n + \int_{X \setminus E} f d\mu_n \right. \\
 & \quad \left. - \int_{X \setminus E} f d\mu - \int_E f d\mu \right\|
 \end{aligned}$$

$$\begin{aligned}
&\leq \left\| \int_E f_n d\mu_n \right\| + \left\| \int_E f d\mu \right\| + \left\| \int_{X \setminus E} (f_n - f) d\mu_n \right\| \\
&\quad + \left\| \int_{X \setminus E} f d(\mu_n - \mu) \right\| \\
&\leq \int_E g_n d\mathcal{P} \circ \mu_n + \int_E g d\mathcal{P} \circ \mu + \int_{X \setminus E} \mathcal{P} \circ (f_n - f) d\mathcal{P} \circ \mu_n \\
&\quad + \int_{X \setminus E} \mathcal{P} \circ f d\mathcal{P} \circ (\mu_n - \mu) \\
&< \epsilon/3 + \epsilon/6 + \frac{\mathcal{P} \circ \mu_n(X \setminus E)\epsilon}{4\mathcal{P} \circ \mu(X) + 1} + M\mathcal{P} \circ (\mu_n - \mu)(X \setminus E) < \epsilon
\end{aligned}$$

where the first equality follows from Proposition 11.132; the first inequality follows from Propositions 11.132, 11.145, and 11.146; the second inequality follows from Proposition 11.132; and the third inequality follows from Proposition 11.83. Hence, we have $\lim_{n \in \mathbb{N}} \int_X f_n d\mu_n = \int_X f d\mu \in \bar{\mathcal{W}}$. This case is proved.

Case 2: $\mathcal{P} \circ \mu(X) = \infty$. Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. By Definition 11.79, $\exists$ a simple function $\phi : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$ such that $0 \leq \phi(x) \leq g(x)$, $\forall x \in X$ and $\int_X g d\mathcal{P} \circ \mu - \epsilon/4 < \int_X \phi d\mathcal{P} \circ \mu \leq \int_X g d\mathcal{P} \circ \mu < \infty$. Let $A_0 := \{x \in X \mid \phi(x) > 0\}$. Then, $A_0 \in \mathcal{B}$ and $\mathcal{P} \circ \mu(A_0) < \infty$. Note that, by Propositions 11.83 and 11.75, $0 \leq \int_{X \setminus A_0} g d\mathcal{P} \circ \mu = \int_X g d\mathcal{P} \circ \mu - \int_{A_0} g d\mathcal{P} \circ \mu \leq \int_X g d\mathcal{P} \circ \mu - \int_{A_0} \phi d\mathcal{P} \circ \mu = \int_X g d\mathcal{P} \circ \mu - \int_X \phi d\mathcal{P} \circ \mu < \epsilon/4$. This implies that $0 \leq \left\| \int_X f d\mu - \int_{A_0} f d\mu \right\| = \left\| \int_{X \setminus A_0} f d\mu \right\| \leq \int_{X \setminus A_0} \mathcal{P} \circ f d\mathcal{P} \circ \mu \leq \int_{X \setminus A_0} g d\mathcal{P} \circ \mu < \epsilon/4$, where the first equality and the second inequality follow from Proposition 11.132; and the third inequality follows from Proposition 11.83. By Lebesgue Dominated Convergence Theorem 11.150, we have $\int_{X \setminus A_0} g d\mathcal{P} \circ \mu = \lim_{n \in \mathbb{N}} \int_{X \setminus A_0} g_n d\mathcal{P} \circ \mu_n$. By Case 1, we have $\int_{A_0} f d\mu = \lim_{n \in \mathbb{N}} \int_{A_0} f_n d\mu_n \in \bar{\mathcal{W}}$. Then, $\exists n_0 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n_0 \leq n$, we have $\left\| \int_{A_0} f d\mu - \int_{A_0} f_n d\mu_n \right\| < \epsilon/4$ and $\int_{X \setminus A_0} g_n d\mathcal{P} \circ \mu_n < \epsilon/2$. This leads to $\left\| \int_X f d\mu - \int_X f_n d\mu_n \right\| \leq \left\| \int_X f d\mu - \int_{A_0} f d\mu \right\| + \left\| \int_{A_0} f d\mu - \int_{A_0} f_n d\mu_n \right\| + \left\| \int_{A_0} f_n d\mu_n - \int_X f_n d\mu_n \right\| < \epsilon/4 + \epsilon/4 + \left\| \int_{X \setminus A_0} f_n d\mu_n \right\| \leq \epsilon/2 + \int_{X \setminus A_0} g_n d\mathcal{P} \circ \mu_n < \epsilon$, where the second and the third inequalities follow from Proposition 11.132. Hence, $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X f_n d\mu_n \in \bar{\mathcal{W}}$. This case is proved.

This completes the proof of the theorem. $\square$

Proposition 11.152 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, $f : X \rightarrow \mathcal{W}$ be absolutely integrable over $\mathcal{X}$. Define $\nu : \mathcal{B} \rightarrow \mathcal{Z}$ by $\nu(E) = \int_E f d\mu$, $\forall E \in \mathcal{B}$. Then, $\bar{\mathcal{X}} := (X, \mathcal{B}, \nu)$ is a finite $\mathcal{Z}$ -valued measure space and $0 \leq \mathcal{P} \circ \nu(E) \leq \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu \leq$*

$$\int_X \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty, \forall E \in \mathcal{B}.$$

Proof By Proposition 11.125, $\nu(\emptyset) = \int_{\emptyset} f \, d\mu = \vartheta_{\mathcal{Z}}$. $\forall E \in \mathcal{B}$, by Proposition 11.132, $\nu(E) = \int_E f \, d\mu \in \mathcal{Z}$ and $0 \leq \|\nu(E)\| \leq \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \leq \int_X \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty$. $\forall$ pairwise disjoint $(E_n)_{n=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{n=1}^{\infty} E_n \in \mathcal{B}$. Then, $\sum_{n=1}^{\infty} \|\nu(E_n)\| \leq \sum_{n=1}^{\infty} \int_{E_n} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty$, where the equality follows from Proposition 11.83. By Proposition 11.132, we have $\nu(E) = \int_E f \, d\mu = \sum_{n=1}^{\infty} \int_{E_n} f \, d\mu = \sum_{n=1}^{\infty} \nu(E_n) \in \mathcal{Z}$. Hence, ν is a $\mathcal{Z}$ -valued pre-measure on $(X, \mathcal{B})$. $\forall E \in \mathcal{B}$, $\mathcal{P} \circ \nu(E) = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\nu(E_i)\| \leq \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \int_{E_i} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \leq \int_X \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty$. Thus, ν is a finite $\mathcal{Z}$ -valued pre-measure, and hence a finite $\mathcal{Z}$ -valued measure on $(X, \mathcal{B})$. This completes the proof of the proposition. $\square$

Proposition 11.153 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, and $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable. Then, there exists a unique $\mathcal{Z}$ -valued measure ν on $(X, \mathcal{B})$ such that $\nu(E) = \int_E f \, d\mu \in \mathcal{Z}$, $\forall E \in \mathcal{B}$ with $\int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty$. Furthermore, ν is σ -finite and $0 \leq \mathcal{P} \circ \nu(E) \leq \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu$, $\forall E \in \mathcal{B}$.*

Proof Since $\mathcal{X}$ is σ -finite, then $\exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mathcal{P} \circ \mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $(X_n)_{n=1}^{\infty}$ is pairwise disjoint. $\forall m \in \mathbb{N}$, let $\bar{X}_m := \{x \in X \mid m-1 \leq \|f(x)\| < m\} \in \mathcal{B}$. Clearly, $(\bar{X}_m)_{m=1}^{\infty} \subseteq \mathcal{B}$ is pairwise disjoint and $X = \bigcup_{m=1}^{\infty} \bar{X}_m$. $\forall m, n \in \mathbb{N}$, let $\mathcal{X}_{m,n} := (\bar{X}_m \cap X_n, \mathcal{B}_{m,n}, \mu_{m,n})$ be the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$. Then, $f|_{\bar{X}_m \cap X_n}$ is absolutely integrable over $\mathcal{X}_{m,n}$. By Propositions 11.152 and 11.132, we may define a finite $\mathcal{Z}$ -valued measure space $\hat{\mathcal{X}}_{m,n} := (\bar{X}_m \cap X_n, \mathcal{B}_{m,n}, \nu_{m,n})$ by $\nu_{m,n}(E) = \int_E f \, d\mu \in \mathcal{Z}$, $\forall E \in \mathcal{B}_{m,n}$. Clearly, the sets $\bar{X}_m \cap X_n$, $m \in \mathbb{N}$, $n \in \mathbb{N}$, are pairwise disjoint. Then, by Proposition 11.118, the generation process on $(\hat{\mathcal{X}}_{m,n})_{m \in \mathbb{N}, n \in \mathbb{N}}$ yields the unique σ -finite $\mathcal{Z}$ -valued measure ν on $(X, \mathcal{B})$ such that $\hat{\mathcal{X}}_{m,n}$ is a finite $\mathcal{Z}$ -valued measure subspace of $\hat{\mathcal{X}} := (X, \mathcal{B}, \nu)$, $\forall m, n \in \mathbb{N}$.

Fix any $E \in \mathcal{B}$ with $\int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty$. $\forall m, n \in \mathbb{N}$, let $E_{m,n} := E \cap \bar{X}_m \cap X_n \in \mathcal{B}$. Then, $\mathcal{P} \circ \nu(E) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mathcal{P} \circ \nu_{m,n}(E_{m,n}) \leq \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \int_{E_{m,n}} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty$, where the first equality follows from Proposition 11.118; the first inequality follows from Proposition 11.152; and the second equality follows from Proposition 11.83. This implies that $E \in \text{dom}(\nu)$ and $\nu(E) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \nu_{m,n}(E_{m,n}) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \int_{E_{m,n}} f \, d\mu = \int_E f \, d\mu \in \mathcal{Z}$, where the first equality follows from Proposition 11.118; the second equality follows from Proposition 11.152;

and the third equality follows from Proposition 11.132. Hence, ν is the σ -finite $\mathbb{Z}$ -valued measure we seek.

This completes the proof of the proposition. $\square$

Proposition 11.154 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite $\mathbb{K}$ -valued measure space, $\mathbb{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathbb{K}, \mathbb{Z}) = \mathbb{Z}$, and $f : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable. Then, there exists a unique $\mathbb{Z}$ -valued measure ν on $(X, \mathcal{B})$ such that $\mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu \in [0, \infty] \subset \mathbb{R}_e, \forall E \in \mathcal{B}$, and $\nu(E) = \int_E f d\mu \in \mathbb{Z}, \forall E \in \mathcal{B}$ with $\mathcal{P} \circ \nu(E) < \infty$. Furthermore, ν is σ -finite.*

Proof By Proposition 11.153, there exists a unique $\mathbb{Z}$ -valued measure ν on $(X, \mathcal{B})$ such that $\nu(E) = \int_E f d\mu \in \mathbb{Z}, \forall E \in \mathcal{B}$ with $\int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu < \infty$. Furthermore, ν is σ -finite and $0 \leq \mathcal{P} \circ \nu(E) \leq \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu, \forall E \in \mathcal{B}$. All we need to show is that $\mathcal{P} \circ \nu(E) \geq \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu, \forall E \in \mathcal{B}$.

Define $\bar{\nu} : \mathcal{B} \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $\bar{\nu}(E) = \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu, \forall E \in \mathcal{B}$. By Proposition 11.116, $\bar{\nu}$ is a σ -finite measure on $(X, \mathcal{B})$. Fix any $E \in \mathcal{B}$. We will distinguish two exhaustive and mutually exclusives cases: Case 1: $E \in \mathcal{B} \setminus \text{dom}(\nu)$; Case 2: $E \in \text{dom}(\nu)$. Case 1: $E \in \mathcal{B} \setminus \text{dom}(\nu)$. Clearly, we have $\mathcal{P} \circ \nu(E) = \infty \geq \bar{\nu}(E)$. This case is proved. Case 2: $E \in \text{dom}(\nu)$. Since $\bar{\nu}$ is σ -finite, then $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty X_n$ and $\bar{\nu}(X_n) < \infty, \forall n \in \mathbb{N}$. Without loss of generality, we may assume that $(X_n)_{n=1}^\infty$ is pairwise disjoint. Let $E_n := E \cap X_n \in \mathcal{B}, \forall n \in \mathbb{N}$. Then, $\forall n \in \mathbb{N}$, we have $\int_{E_n} \mathcal{P} \circ f d\mathcal{P} \circ \mu = \bar{\nu}(E_n) \leq \bar{\nu}(X_n) < \infty. \forall n \in \mathbb{N}, \forall F \in \mathcal{B}$ with $F \subseteq E_n$, we have $0 \leq \mathcal{P} \circ \nu(F) \leq \int_F \mathcal{P} \circ f d\mathcal{P} \circ \mu = \bar{\nu}(F) \leq \bar{\nu}(E_n) < \infty$ and $\nu(F) = \int_F f d\mu \in \mathbb{Z}$. Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. Fix any $n \in \mathbb{N}$. Let $\mathcal{E}_n := (E_n, \mathcal{B}_{E_n}, \mu_{E_n})$ be the σ -finite $\mathbb{K}$ -valued measure subspace of $\mathcal{X}$. By Proposition 11.94, $\exists$ simple function $\varphi_n : \mathcal{E}_n \rightarrow \mathcal{W}$ such that $\|\varphi_n(x)\| \leq \mathcal{P} \circ f(x), \forall x \in E_n, \left| \int_{E_n} \mathcal{P} \circ \varphi_n d\mathcal{P} \circ \mu - \int_{E_n} \mathcal{P} \circ f d\mathcal{P} \circ \mu \right| < 2^{-n-2}\epsilon$, and $\int_{E_n} \mathcal{P} \circ (f|_{E_n} - \varphi_n) d\mathcal{P} \circ \mu < 2^{-n-1}\epsilon$. Let φ_n admit the canonical representation $\varphi_n = \sum_{i=1}^{m_n} w_{i,n} \chi_{E_{i,n}, E_n}$, where $m_n \in \mathbb{Z}_+, w_{1,n}, \dots, w_{m_n,n} \in \mathcal{W}$ are distinct and none equals to $\vartheta_{\mathbb{Z}}$, and $E_{1,n}, \dots, E_{m_n,n} \in \mathcal{B}$ are subsets of E_n , nonempty, pairwise disjoint, and $\mathcal{P} \circ \mu(E_{i,n}) < \infty, i = 1, \dots, m_n. \forall i \in \{1, \dots, m_n\}, \exists m_{i,n} \in \mathbb{Z}_+, \exists$ pairwise disjoint $(E_{j,i,n})_{j=1}^{m_{i,n}} \subseteq \mathcal{B}$ with $E_{i,n} = \bigcup_{j=1}^{m_{i,n}} E_{j,i,n}$ such that $\mathcal{P} \circ \mu(E_{i,n}) < \sum_{j=1}^{m_{i,n}} |\mu(E_{j,i,n})| + \frac{2^{-n-2}\epsilon}{(1+m_n)\|w_{i,n}\|}$. This leads to

$$\begin{aligned} \bar{\nu}(E_n) &= \int_{E_n} \mathcal{P} \circ f d\mathcal{P} \circ \mu < \int_{E_n} \mathcal{P} \circ \varphi_n d\mathcal{P} \circ \mu + 2^{-n-2}\epsilon \\ &= \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \|\nu(E_{j,i,n})\| + \sum_{i=1}^{m_n} \|w_{i,n}\| \mathcal{P} \circ \mu(E_{i,n}) + 2^{-n-2}\epsilon \\ &\quad - \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \|\nu(E_{j,i,n})\| \end{aligned}$$

$$\begin{aligned}
&\leq \mathcal{P} \circ \nu(E_n) + \sum_{i=1}^{m_n} \|w_{i,n}\| \left(\sum_{j=1}^{m_{i,n}} |\mu(E_{j,i,n})| + \frac{2^{-n-2}\epsilon}{(1+m_n)\|w_{i,n}\|} \right) \\
&\quad + 2^{-n-2}\epsilon - \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \|\nu(E_{j,i,n})\| \\
&< \mathcal{P} \circ \nu(E_n) + \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \|w_{i,n}\mu(E_{j,i,n})\| + 2^{-n-1}\epsilon - \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \|\nu(E_{j,i,n})\| \\
&= \mathcal{P} \circ \nu(E_n) + 2^{-n-1}\epsilon + \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \left(\left\| \int_{E_{j,i,n}} \varphi_n d\mu \right\| - \left\| \int_{E_{j,i,n}} f d\mu \right\| \right) \\
&\leq \mathcal{P} \circ \nu(E_n) + 2^{-n-1}\epsilon + \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \left\| \int_{E_{j,i,n}} (\varphi_n - f|_{E_n}) d\mu \right\| \\
&\leq \mathcal{P} \circ \nu(E_n) + 2^{-n-1}\epsilon + \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \int_{E_{j,i,n}} \mathcal{P} \circ (\varphi_n - f|_{E_n}) d\mathcal{P} \circ \mu \\
&= \mathcal{P} \circ \nu(E_n) + 2^{-n-1}\epsilon + \int_{\bigcup_{i=1}^{m_n} \bigcup_{j=1}^{m_{i,n}} E_{j,i,n}} \mathcal{P} \circ (\varphi_n - f|_{E_n}) d\mathcal{P} \circ \mu \\
&\leq \mathcal{P} \circ \nu(E_n) + 2^{-n-1}\epsilon + \int_{E_n} \mathcal{P} \circ (\varphi_n - f|_{E_n}) d\mathcal{P} \circ \mu \\
&< \mathcal{P} \circ \nu(E_n) + 2^{-n}\epsilon
\end{aligned}$$

where the second equality follows from Proposition 11.75; the third equality follows from Proposition 11.125; the fourth and the fifth inequalities follow from Proposition 11.132; and the fourth equality and the sixth inequality follow from Proposition 11.83. Then, $\bar{\nu}(E) = \sum_{n=1}^{\infty} \bar{\nu}(E_n) < \sum_{n=1}^{\infty} (\mathcal{P} \circ \nu(E_n) + 2^{-n}\epsilon) = \mathcal{P} \circ \nu(E) + \epsilon$. By the arbitrariness of ϵ , we have $\bar{\nu}(E) \leq \mathcal{P} \circ \nu(E)$. This case is proved.

Hence, $\mathcal{P} \circ \nu(E) = \bar{\nu}(E) = \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu$, $\forall E \in \mathcal{B}$. This completes the proof of the proposition. $\square$

Definition 11.155 Let $(X, \mathcal{B})$ be a measurable space, μ_1 and μ_2 be measures on $(X, \mathcal{B})$. We will say that μ_1 is absolutely continuous with respect to μ_2 if $\forall A \in \mathcal{B}$ with $\mu_2(A) = 0$, we have $\mu_1(A) = 0$. In this case, we will denote $\mu_1 \ll \mu_2$. μ_1 and μ_2 are said to be mutually singular if $\exists A, B \in \mathcal{B}$ with $A \cap B = \emptyset$ and $X = A \cup B$, such that $\mu_1(A) = \mu_2(B) = 0$. In this case, we will denote $\mu_1 \perp \mu_2$.

Proposition 11.156 Let μ , λ_1 , and λ_2 be measures on a measurable space $(X, \mathcal{B})$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, ν_1 and ν_2 be σ -finite $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$, $a, b \in (0, \infty) \subset \mathbb{R}$, and $\alpha, \beta \in \mathbb{K}$. Then, the following statements hold.

- (i) $\lambda_1 \ll \mu$ and $\lambda_2 \ll \mu$ if, and only if, $a\lambda_1 + b\lambda_2 \ll \mu$.

(ii) $\lambda_1 \perp \mu$ and $\lambda_2 \perp \mu$ if, and only if, $a\lambda_1 + b\lambda_2 \perp \mu$.

(iii) If $\lambda_1 \ll \mu$ and $\lambda_2 \perp \mu$, then $\lambda_1 \perp \lambda_2$.

(iv) If $\lambda_1 \perp \lambda_1$, then $\lambda_1(E) = 0, \forall E \in \mathcal{B}$.

(v) If $\mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_2$, then $\mathcal{P} \circ (\alpha\nu_1) \perp \mathcal{P} \circ (\beta\nu_2)$.

(vi) If $\mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_1$, then $\nu_1(E) = \vartheta_{\mathcal{Y}}, \forall E \in \mathcal{B}$.

If, in addition, $\mathcal{Y}$ is complete, then the following statements hold.

(vii) If $\mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_2$, then $\mathcal{P} \circ (\nu_1 + \nu_2) = \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2$.

(viii) If $\mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_2$, then $\mathcal{P} \circ (\alpha\nu_1 + \beta\nu_2) = |\alpha| \mathcal{P} \circ \nu_1 + |\beta| \mathcal{P} \circ \nu_2$.

(ix) If $\mathcal{P} \circ \nu_1 \ll \mu$ and $\mathcal{P} \circ \nu_2 \ll \mu$, then $\mathcal{P} \circ (\alpha\nu_1 + \beta\nu_2) \ll \mu$.

(x) If $\mathcal{P} \circ \nu_1 \perp \mu$ and $\mathcal{P} \circ \nu_2 \perp \mu$, then $\mathcal{P} \circ (\alpha\nu_1 + \beta\nu_2) \perp \mu$.

Proof (i) By Proposition 11.136, $a\lambda_1 + b\lambda_2$ is a measure on $(X, \mathcal{B})$. “Sufficiency” $\forall E \in \mathcal{B}$ with $\mu(E) = 0$, by $a\lambda_1 + b\lambda_2 \ll \mu$, we have $a\lambda_1(E) + b\lambda_2(E) = 0$. Then, $\lambda_1(E) = \lambda_2(E) = 0$. Hence, $\lambda_1 \ll \mu$ and $\lambda_2 \ll \mu$. “Necessity” $\forall E \in \mathcal{B}$ with $\mu(E) = 0$, by $\lambda_1 \ll \mu$ and $\lambda_2 \ll \mu$, we have $\lambda_1(E) = \lambda_2(E) = 0$. Then, $a\lambda_1(E) + b\lambda_2(E) = 0$. Hence, $a\lambda_1 + b\lambda_2 \ll \mu$.

(ii) By Proposition 11.136, $a\lambda_1 + b\lambda_2$ is a measure on $(X, \mathcal{B})$. “Sufficiency” Let $a\lambda_1 + b\lambda_2 \perp \mu$. Then, $\exists A, B \in \mathcal{B}$ with $A = X \setminus B$ such that $(a\lambda_1 + b\lambda_2)(A) = 0 = \mu(B)$. Then, $\lambda_1(A) = \lambda_2(A) = \mu(B) = 0$. Hence, $\lambda_1 \perp \mu$ and $\lambda_2 \perp \mu$. “Necessity” Let $\lambda_1 \perp \mu$ and $\lambda_2 \perp \mu$. Then, $\exists A_i, B_i \in \mathcal{B}$ with $A_i = X \setminus B_i$ such that $\lambda_i(A_i) = \mu(B_i) = 0, i = 1, 2$. Let $A := A_1 \cap A_2 \in \mathcal{B}$ and $B := B_1 \cup B_2 \in \mathcal{B}$. Then, $A = X \setminus B$ and $0 \leq a\lambda_1(A) + b\lambda_2(A) = (a\lambda_1 + b\lambda_2)(A) \leq a\lambda_1(A_1) + b\lambda_2(A_2) = 0 \leq \mu(B) \leq \mu(B_1) + \mu(B_2) = 0$. Hence, $a\lambda_1 + b\lambda_2 \perp \mu$.

(iii) By $\lambda_2 \perp \mu$, $\exists A, B \in \mathcal{B}$ with $A = X \setminus B$ such that $\lambda_2(A) = 0 = \mu(B)$. By $\lambda_1 \ll \mu$, we have $\lambda_1(B) = 0$. Hence, $\lambda_1 \perp \lambda_2$.

(iv) By $\lambda_1 \perp \lambda_1$, $\exists A, B \in \mathcal{B}$ with $A = X \setminus B$ such that $\lambda_1(A) = \lambda_1(B) = 0$. Then, $\lambda_1(X) = \lambda_1(A) + \lambda_1(B) = 0$. Hence, the result holds.

(v) By Proposition 11.138, $\alpha\nu_1$ and $\beta\nu_2$ are σ -finite $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$ with $\mathcal{P} \circ (\alpha\nu_1) = |\alpha| \mathcal{P} \circ \nu_1$ and $\mathcal{P} \circ (\beta\nu_2) = |\beta| \mathcal{P} \circ \nu_2$. By $\mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_2$, $\exists A, B \in \mathcal{B}$ with $A = X \setminus B$ such that $\mathcal{P} \circ \nu_1(A) = 0 = \mathcal{P} \circ \nu_2(B)$. Then, by Proposition 11.138, $\mathcal{P} \circ (\alpha\nu_1)(A) = |\alpha| \mathcal{P} \circ \nu_1(A) = 0 = |\beta| \mathcal{P} \circ \nu_2(B) = \mathcal{P} \circ (\beta\nu_2)(B)$. Hence, $\mathcal{P} \circ (\alpha\nu_1) \perp \mathcal{P} \circ (\beta\nu_2)$.

(vi) By (iv), $\mathcal{P} \circ \nu_1(E) = 0, \forall E \in \mathcal{B}$. Then, $\nu_1(E) = \vartheta_{\mathcal{Y}}, \forall E \in \mathcal{B}$.

Let $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, then $\mathcal{M}_{\sigma}(X, \mathcal{B}, \mathcal{Y})$ is well-defined by Proposition 11.142.

(vii) By Proposition 11.138, $\nu_1 + \nu_2 \in \mathcal{M}_{\sigma}(X, \mathcal{B}, \mathcal{Y})$ and $\mathcal{P} \circ (\nu_1 + \nu_2) \leq \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2$. By $\mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_2$, $\exists A, B \in \mathcal{B}$ with $A = X \setminus B$ such that $\mathcal{P} \circ \nu_1(A) = 0 = \mathcal{P} \circ \nu_2(B)$. Fix any $E \in \mathcal{B}$. We will show that

$\mathcal{P} \circ (\nu_1 + \nu_2)(E) \geq \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E)$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) < \infty$; Case 2: $\mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) = \infty$.

Case 1: $\mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) < \infty$. Then, $\mathcal{P} \circ \nu_i(E) < \infty$ and $E \in \text{dom}(\nu_i)$, $i = 1, 2$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_1 \in \mathbb{Z}_+$, $\exists$ pairwise disjoint $(E_i)_{i=1}^{n_1} \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^{n_1} E_i$ such that $\mathcal{P} \circ \nu_1(E) < \sum_{i=1}^{n_1} \|\nu_1(E_i)\| + \epsilon/2 < +\infty$. $\exists n_2 \in \mathbb{Z}_+$, $\exists$ pairwise disjoint $(F_i)_{i=1}^{n_2} \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^{n_2} F_i$ such that $\mathcal{P} \circ \nu_2(E) < \sum_{i=1}^{n_2} \|\nu_2(F_i)\| + \epsilon/2 < +\infty$. Then,

$$\begin{aligned}
\mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) &< \sum_{i=1}^{n_1} \|\nu_1(E_i)\| + \sum_{i=1}^{n_2} \|\nu_2(F_i)\| + \epsilon \\
&= \sum_{i=1}^{n_1} \|\nu_1(E_i \cap A) + \nu_1(E_i \cap B)\| + \sum_{i=1}^{n_2} \|\nu_2(F_i \cap A) + \nu_2(F_i \cap B)\| + \epsilon \\
&= \sum_{i=1}^{n_1} \|\nu_2(E_i \cap B) + \nu_1(E_i \cap B)\| + \sum_{i=1}^{n_2} \|\nu_2(F_i \cap A) + \nu_1(F_i \cap A)\| + \epsilon \\
&= \sum_{i=1}^{n_1} \|(\nu_1 + \nu_2)(E_i \cap B)\| + \sum_{i=1}^{n_2} \|(\nu_1 + \nu_2)(F_i \cap A)\| + \epsilon \\
&\leq \sum_{i=1}^{n_1} \mathcal{P} \circ (\nu_1 + \nu_2)(E_i \cap B) + \sum_{i=1}^{n_2} \mathcal{P} \circ (\nu_1 + \nu_2)(F_i \cap A) + \epsilon \\
&= \mathcal{P} \circ (\nu_1 + \nu_2)(E \cap B) + \mathcal{P} \circ (\nu_1 + \nu_2)(E \cap A) + \epsilon \\
&= \mathcal{P} \circ (\nu_1 + \nu_2)(E) + \epsilon
\end{aligned}$$

where the second equality follows from the fact that $\forall \bar{A} \in \mathcal{B}$ with $\bar{A} \subseteq A$, we have $\mathcal{P} \circ \nu_1(\bar{A}) = 0$ and $\nu_1(\bar{A}) = \vartheta_{\mathbf{y}}$, and the fact that $\forall \bar{B} \in \mathcal{B}$ with $\bar{B} \subseteq B$, we have $\mathcal{P} \circ \nu_2(\bar{B}) = 0$ and $\nu_2(\bar{B}) = \vartheta_{\mathbf{y}}$; the third equality follows from Proposition 11.138; and the second inequality and the last two equalities follow from Definition 11.108. By the arbitrariness of ϵ , we have $\mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) \leq \mathcal{P} \circ (\nu_1 + \nu_2)(E)$. This case is proved.

Case 2: $\mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) = \infty$. Without loss of generality, assume $\mathcal{P} \circ \nu_1(E) = \infty$. Since ν_1 and ν_2 are σ -finite, then $\exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mathcal{P} \circ \nu_i(X_n) < \infty$, $\forall n \in \mathbb{N}$, $i = 1, 2$. Without loss of generality, we may assume that $X_n \subseteq X_{n+1}$, $\forall n \in \mathbb{N}$. Let $E_n := E \cap X_n$, $\forall n \in \mathbb{N}$. Then, $E = \bigcup_{n=1}^{\infty} E_n$ and $E_n \subseteq E_{n+1}$, $\forall n \in \mathbb{N}$. By Proposition 11.7, we have $\lim_{n \in \mathbb{N}} \mathcal{P} \circ \nu_1(E_n) = \mathcal{P} \circ \nu_1(E) = \infty$. $\forall M \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\mathcal{P} \circ \nu_1(E_{n_0}) > M + 1$. Clearly, $\mathcal{P} \circ \nu_i(E_{n_0}) \leq \mathcal{P} \circ \nu_i(X_{n_0}) < \infty$, $i = 1, 2$. Then, $\exists n_1 \in \mathbb{Z}_+$, $\exists$ pairwise disjoint $(\bar{E}_i)_{i=1}^{n_1} \subseteq \mathcal{B}$ with $E_{n_0} = \bigcup_{i=1}^{n_1} \bar{E}_i$ such that $\mathcal{P} \circ \nu_1(E_{n_0}) < \sum_{i=1}^{n_1} \|\nu_1(\bar{E}_i)\| + 1$. Then,

$$M < \sum_{i=1}^{n_1} \|\nu_1(\bar{E}_i)\| = \sum_{i=1}^{n_1} \|\nu_1(\bar{E}_i \cap A) + \nu_1(\bar{E}_i \cap B)\|$$

$$\begin{aligned}
&= \sum_{i=1}^{n_1} \|\nu_2(\bar{E}_i \cap B) + \nu_1(\bar{E}_i \cap B)\| \\
&= \sum_{i=1}^{n_1} \|(\nu_1 + \nu_2)(\bar{E}_i \cap B)\| \leq \sum_{i=1}^{n_1} \mathcal{P} \circ (\nu_1 + \nu_2)(\bar{E}_i \cap B) \\
&= \mathcal{P} \circ (\nu_1 + \nu_2)(E_{n_0} \cap B) \leq \mathcal{P} \circ (\nu_1 + \nu_2)(E)
\end{aligned}$$

where the first equality follows from Definition 11.108; the second equality follows from the fact that $\forall \bar{A} \in \mathcal{B}$ with $\bar{A} \subseteq A$, we have $\mathcal{P} \circ \nu_1(\bar{A}) = 0$ and $\nu_1(\bar{A}) = \vartheta_{\mathcal{Y}}$, and the fact that $\forall \bar{B} \in \mathcal{B}$ with $\bar{B} \subseteq B$, we have $\mathcal{P} \circ \nu_2(\bar{B}) = 0$ and $\nu_2(\bar{B}) = \vartheta_{\mathcal{Y}}$; the third equality follows from Proposition 11.138; and the second inequality follows from Definition 11.108. By the arbitrariness of M , we have $\mathcal{P} \circ (\nu_1 + \nu_2)(E) = \infty$. Then, $\mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) = \mathcal{P} \circ (\nu_1 + \nu_2)(E) = \infty$. This case is proved.

In both cases, we have $\mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) \leq \mathcal{P} \circ (\nu_1 + \nu_2)(E)$. By the arbitrariness of E , we have $\mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 \leq \mathcal{P} \circ (\nu_1 + \nu_2)$. Hence, $\mathcal{P} \circ (\nu_1 + \nu_2) = \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2$.

(viii) By Proposition 11.138, $\alpha\nu_1, \beta\nu_2, \alpha\nu_1 + \beta\nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$. By (v), we have $\mathcal{P} \circ (\alpha\nu_1) \perp \mathcal{P} \circ (\beta\nu_2)$. By (vii) and Proposition 11.138, we have $\mathcal{P} \circ (\alpha\nu_1 + \beta\nu_2) = \mathcal{P} \circ (\alpha\nu_1) + \mathcal{P} \circ (\beta\nu_2) = |\alpha| \mathcal{P} \circ \nu_1 + |\beta| \mathcal{P} \circ \nu_2$.

(ix) By Proposition 11.138, $\alpha\nu_1 + \beta\nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$. $\forall E \in \mathcal{B}$ with $\mu(E) = 0$, by $\mathcal{P} \circ \nu_i \ll \mu$, we have $\mathcal{P} \circ \nu_i(E) = 0$, $i = 1, 2$. Then, $0 \leq \mathcal{P} \circ (\alpha\nu_1 + \beta\nu_2)(E) \leq |\alpha| \mathcal{P} \circ \nu_1(E) + |\beta| \mathcal{P} \circ \nu_2(E) = 0$, where the second inequality follows from Proposition 11.138. Hence, $\mathcal{P} \circ (\alpha\nu_1 + \beta\nu_2) \ll \mu$.

(x) By Proposition 11.138, $\alpha\nu_1 + \beta\nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$. By $\mathcal{P} \circ \nu_i \perp \mu$, $\exists A_i, B_i \in \mathcal{B}$ with $A_i = X \setminus B_i$ such that $\mathcal{P} \circ \nu_i(A_i) = \mu(B_i) = 0$, $i = 1, 2$. Let $A := A_1 \cap A_2 \in \mathcal{B}$ and $B := B_1 \cup B_2 \in \mathcal{B}$. Then, $A = X \setminus B$ and $0 \leq \mathcal{P} \circ (\alpha\nu_1 + \beta\nu_2)(A) \leq (|\alpha| \mathcal{P} \circ \nu_1 + |\beta| \mathcal{P} \circ \nu_2)(A) \leq |\alpha| \mathcal{P} \circ \nu_1(A_1) + |\beta| \mathcal{P} \circ \nu_2(A_2) = 0 \leq \mu(B) \leq \mu(B_1) + \mu(B_2) = 0$, where the second inequality follows from Proposition 11.138. Hence, $\mathcal{P} \circ (\alpha\nu_1 + \beta\nu_2) \perp \mu$.

This completes the proof of the proposition. $\square$

Proposition 11.157 *Let $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space, $A \in \mathcal{B}$, and $B := X \setminus A$. Define μ_A to be a function from $\mathcal{B}$ to $\mathcal{Y}$ by: $\forall E \in \mathcal{B}$ with $E \cap A \in \text{dom}(\mu)$, $\mu_A(E) := \mu(E \cap A) \in \mathcal{Y}$; $\forall E \in \mathcal{B}$ with $E \cap A \in \mathcal{B} \setminus \text{dom}(\mu)$, $\mu_A(E)$ is undefined. Define μ_B to be a function from $\mathcal{B}$ to $\mathcal{Y}$ similarly as μ_A with A replaced with B . Then, the following statements hold.*

- (i) μ_A and μ_B are $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$ with $\mathcal{P} \circ \mu_A(E) = \mathcal{P} \circ \mu(E \cap A)$ and $\mathcal{P} \circ \mu_B(E) = \mathcal{P} \circ \mu(E \cap B)$, $\forall E \in \mathcal{B}$.
- (ii) $\mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B$ and $\mathcal{P} \circ \mu = \mathcal{P} \circ \mu_A + \mathcal{P} \circ \mu_B$.
- (iii) if μ is σ -finite and $\mathcal{Y}$ is a Banach space, then μ_A and μ_B are σ -finite and $\mu = \mu_A + \mu_B$.

Proof (i) Define $\nu_A : \mathcal{B} \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $\nu_A(E) = \mathcal{P} \circ \mu(E \cap A)$, $\forall E \in \mathcal{B}$. It is straightforward to check that ν_A is a measure on $(X, \mathcal{B})$. $\forall E \in \mathcal{B}$ with $\nu_A(E) = \infty$, then $\mathcal{P} \circ \mu(E \cap A) = \infty$, $E \cap A \in \mathcal{B} \setminus \text{dom}(\mu)$, and $\mu_A(E)$ is undefined. $\forall E \in \mathcal{B}$ with $\nu_A(E) < \infty$, then $\mathcal{P} \circ \mu(E \cap A) < \infty$, $E \cap A \in \text{dom}(\mu)$, and $\mu_A(E) = \mu(E \cap A) \in \mathcal{Y}$. $\forall$ pairwise disjoint $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$ with $E = \bigcup_{n=1}^\infty E_n$, clearly, $0 \leq \mathcal{P} \circ \mu(E_n \cap A) \leq \mathcal{P} \circ \mu(E \cap A) < \infty$, $E_n \cap A \in \text{dom}(\mu)$, and $\mu_A(E_n) = \mu(E_n \cap A) \in \mathcal{Y}$, $\forall n \in \mathbb{N}$. By Definition 11.108, $\mu_A(E) = \mu(E \cap A) = \sum_{n=1}^\infty \mu(E_n \cap A) = \sum_{n=1}^\infty \mu_A(E_n)$ and $\sum_{n=1}^\infty \|\mu_A(E_n)\| = \sum_{n=1}^\infty \|\mu(E_n \cap A)\| < \infty$. $\forall E \in \mathcal{B}$ with $\nu_A(E) < \infty$,

$$\begin{aligned}
\nu_A(E) &= \mathcal{P} \circ \mu(E \cap A) \\
&= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E \cap A = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(E_i)\| \\
&= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E \cap A = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu_A(E_i)\| \\
&= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu_A(E_i)\|
\end{aligned}$$

Hence, μ_A is a $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \mu_A = \nu_A$. By symmetry, μ_B is a $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \mu_B(E) = \mathcal{P} \circ \mu(E \cap B)$, $\forall E \in \mathcal{B}$.

(ii) Clearly, $\mathcal{P} \circ \mu_A(B) = 0 = \mathcal{P} \circ \mu_B(A)$. Then, $\mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B$. $\forall E \in \mathcal{B}$, $\mathcal{P} \circ \mu_A(E) + \mathcal{P} \circ \mu_B(E) = \mathcal{P} \circ \mu(E \cap A) + \mathcal{P} \circ \mu(E \cap B) = \mathcal{P} \circ \mu(E)$. Hence, $\mathcal{P} \circ \mu = \mathcal{P} \circ \mu_A + \mathcal{P} \circ \mu_B$.

(iii) If μ is σ -finite and $\mathcal{Y}$ is a Banach space, then $\mathcal{P} \circ \mu$ is σ -finite and $\mathcal{P} \circ \mu_A$ and $\mathcal{P} \circ \mu_B$ are σ -finite. Therefore, μ_A and μ_B are σ -finite. Then, $\mu_A + \mu_B$ is well-defined as in Proposition 11.138. $\forall E \in \text{dom}(\mu_A) \cap \text{dom}(\mu_B)$, we have $E \cap A \in \text{dom}(\mu)$ and $E \cap B \in \text{dom}(\mu)$. Then, $\mathcal{P} \circ \mu(E) = \mathcal{P} \circ \mu(E \cap A) + \mathcal{P} \circ \mu(E \cap B) < \infty$ and $E \in \text{dom}(\mu)$. Hence, $\mu(E) = \mu(E \cap A) + \mu(E \cap B) = \mu_A(E) + \mu_B(E)$. By Proposition 11.138, $\mu = \mu_A + \mu_B$.

This completes the proof of the proposition. $\square$

11.8 The Radon-Nikodym Theorem

Definition 11.158 Let $(X, \mathcal{B}, \mu)$ be an $\mathbb{R}$ -valued pre-measure space (or a σ -finite $\mathbb{R}$ -valued measure space). $A \in \mathcal{B}$ is said to be a positive set with respect to μ if $\forall E \in \text{dom}(\mu)$ with $E \subseteq A$, we have $\mu(E) \in [0, \infty) \subset \mathbb{R}$. $B \in \mathcal{B}$ is said to be a negative set with respect to μ if $\forall E \in \text{dom}(\mu)$ with $E \subseteq B$, we have $\mu(E) \in (-\infty, 0] \subset \mathbb{R}$.

Lemma 11.159 Let $(X, \mathcal{B}, \mu)$ be an $\mathbb{R}$ -valued pre-measure space (or a σ -finite $\mathbb{R}$ -valued measure space). Then, every measurable subset of a positive

set is a positive set; the union of countably many positive sets is a positive set. Similar statements hold for negative sets.

Proof The first statement follows directly from Definition 11.158. For the second statement, let $(A_n)_{n=1}^\infty \subseteq \mathcal{B}$ be a sequence of positive sets and $A := \bigcup_{n=1}^\infty A_n$. $\forall E \in \text{dom}(\mu)$ with $E \subseteq A$, let $E_n := E \cap A_n \setminus (\bigcup_{i=1}^{n-1} A_i) \in \mathcal{B}$, $\forall n \in \mathbb{N}$. Then, $E = \bigcup_{n=1}^\infty E_n$, $(E_n)_{n=1}^\infty$ is pairwise disjoint, $E_n \in \text{dom}(\mu)$, $E_n \subseteq A_n$, and $\mu(E_n) \in [0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$. This leads to $\mu(E) = \sum_{i=1}^\infty \mu(E_n) \in [0, \infty) \subset \mathbb{R}$. Hence, A is a positive set. For the case of finitely many positive sets, we may form a sequence of positive sets by appending $\emptyset$, which is clearly a positive set, to the collection. This completes the proof of the lemma. $\square$

Lemma 11.160 *Let $(X, \mathcal{B}, \mu)$ be an $\mathbb{R}$ -valued pre-measure space (or a σ -finite $\mathbb{R}$ -valued measure space) and $E \in \mathcal{B}$. Assume that $E \in \text{dom}(\mu)$ and $\mu(E) \in (0, +\infty) \subset \mathbb{R}$. Then, there exists a positive set $A \subseteq E$ with $\mu(A) \in (0, +\infty) \subset \mathbb{R}$. Similar statement holds for negative sets.*

Proof Note that, $\forall \bar{E} \in \mathcal{B}$ with $\bar{E} \subseteq E$, we have $\bar{E} \in \text{dom}(\mu)$ and $\mu(\bar{E}) \in \mathbb{R}$. Either E is a positive set, in which case the result holds, or E contains a subset of negative measure. In the latter case, let $n_1 \in \mathbb{N}$ be the smallest natural number such that there is a measurable set $E_1 \subseteq E$ with $\mu(E_1) < -1/n_1$. Then, $\mu(E \setminus E_1) = \mu(E) - \mu(E_1) \in (0, +\infty) \subset \mathbb{R}$. Inductively, assume that we have obtained $n_j \in \mathbb{N}$, $E_j \in \mathcal{B}$, $\forall j \in \{1, \dots, k\} \subseteq \mathbb{N}$, such that $E_j \subseteq E \setminus (\bigcup_{l=1}^{j-1} E_l)$, $\mu(E_j) < -1/n_j$, and $\mu(E \setminus (\bigcup_{l=1}^j E_l)) \in (0, +\infty) \subset \mathbb{R}$, $\forall j \in \{1, \dots, k\}$. Either $E \setminus (\bigcup_{j=1}^k E_j)$ is a positive set, in which case the result holds, or it contains a subset of negative measure. In the latter case, let $n_{k+1} \in \mathbb{N}$ be the smallest natural number such that there is a measurable set $E_{k+1} \subseteq E \setminus (\bigcup_{j=1}^k E_j)$ with $\mu(E_{k+1}) < -1/n_{k+1}$. Then, $\mu(E \setminus (\bigcup_{j=1}^{k+1} E_j)) = \mu(E \setminus (\bigcup_{j=1}^k E_j)) - \mu(E_{k+1}) \in (0, +\infty) \subset \mathbb{R}$. Thus, either the result holds or we have $(n_k)_{k=1}^\infty \subseteq \mathbb{N}$ and $(E_k)_{k=1}^\infty \subseteq \mathcal{B}$. Let $A := E \setminus (\bigcup_{k=1}^\infty E_k) \in \mathcal{B}$. Then, $E = A \cup (\bigcup_{k=1}^\infty E_k)$, where the sets in the union are pairwise disjoint. Then, $\mu(E) = \mu(A) + \sum_{k=1}^\infty \mu(E_k)$ and $|\mu(A)| - \sum_{k=1}^\infty \mu(E_k) < +\infty$. Then, $\sum_{k=1}^\infty 1/n_k < +\infty$, $\lim_{k \in \mathbb{N}} n_k = \infty$, and $\mu(A) \in (0, +\infty) \subset \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists k_0 \in \mathbb{N}$ such that $(n_{k_0} - 1)^{-1} < \epsilon$. Since $A \subseteq E \setminus (\bigcup_{j=1}^{k_0-1} E_j)$, then A can not contain any measurable subset with measure less than $-1/(n_{k_0} - 1) > -\epsilon$. Thus, A contains no measurable subset with measure less than $-\epsilon$. By the arbitrariness of ϵ , A is a positive set. This completes the proof of the lemma. $\square$

Theorem 11.161 (Hahn Decomposition Theorem) *Let $(X, \mathcal{B}, \mu) =: \mathcal{X}$ be an $\mathbb{R}$ -valued pre-measure space (or a σ -finite $\mathbb{R}$ -valued measure space). Then, there is a positive set $A \in \mathcal{B}$ and a negative set $B \in \mathcal{B}$ such that $X = A \cup B$ and $A \cap B = \emptyset$.*

Proof We will distinguish two exhaustive and mutually exclusives cases: Case 1: $\mathcal{X}$ is an $\mathbb{R}$ -valued pre-measure space; Case 2: $\mathcal{X}$ is a σ -finite $\mathbb{R}$ -valued measure space.

Case 1: $\mathcal{X}$ is an $\mathbb{R}$ -valued pre-measure space. Let $\lambda := \sup \{ \mu(A) \in \mathbb{R} \mid A \text{ is a positive set} \} \in \mathbb{R}_e$. Since $\emptyset$ is a positive set, then $\lambda \geq 0$. Let $(A_i)_{i=1}^\infty \subseteq \mathcal{B}$ be a sequence of positive sets such that $\lim_{i \in \mathbb{N}} \mu(A_i) = \lambda$. Let $A := \bigcup_{i=1}^\infty A_i \in \mathcal{B}$. By Lemma 11.159, A is a positive set. This implies that $\mu(A) \leq \lambda$. Note that $\mu(A) = \mu(A_i) + \mu(A \setminus A_i) \geq \mu(A_i)$, $\forall i \in \mathbb{N}$. Then, $\lambda = \mu(A) \in \mathbb{R}$. Let $B := X \setminus A$. Suppose that there exists $E \subseteq B$ such that E is a positive set with $\mu(E) > 0$. Then $A \cup E$ is a positive set by Lemma 11.159 and $\mu(A \cup E) = \mu(A) + \mu(E) > \lambda$. This contradicts with the definition of λ . Hence, B does not contain any subset that is a positive set with positive measure. Then, by Lemma 11.160, B does not contain any subset with positive measure. Hence, B is a negative set. This case is proved.

Case 2: $\mathcal{X}$ is a σ -finite $\mathbb{R}$ -valued measure space. Then, $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty X_n$ and $\mathcal{P} \circ \mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $(X_n)_{n=1}^\infty$ is pairwise disjoint. Let $\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite $\mathbb{R}$ -valued measure subspace of $\mathcal{X}$, $\forall n \in \mathbb{N}$. Fix any $n \in \mathbb{N}$. $\mathcal{X}_n$ is a finite $\mathbb{R}$ -valued pre-measure space. By Case 1, there exists $A_n, B_n \in \mathcal{B}_n \subseteq \mathcal{B}$ such that $A_n = X_n \setminus B_n$ and A_n is a positive set with respect to μ_n and B_n is a negative set with respect to μ_n . It is straightforward to see that A_n is a positive set with respect to μ and B_n is a negative set with respect to μ . Let $A := \bigcup_{n=1}^\infty A_n \in \mathcal{B}$ and $B := \bigcup_{n=1}^\infty B_n \in \mathcal{B}$. Then, by Lemma 11.159, A is a positive set with respect to μ and B is a negative set with respect to μ . $A \cup B = (\bigcup_{n=1}^\infty A_n) \cup (\bigcup_{n=1}^\infty B_n) = \bigcup_{n=1}^\infty (A_n \cup B_n) = \bigcup_{n=1}^\infty X_n = X$ and $A \cap B = (\bigcup_{n=1}^\infty A_n) \cap (\bigcup_{n=1}^\infty B_n) = \bigcup_{n=1}^\infty \bigcup_{m=1}^\infty (A_n \cap B_m) = \emptyset$. This case is proved.

This completes the proof of the theorem. $\square$

Theorem 11.162 (Jordan Decomposition Theorem) *Let $(X, \mathcal{B}, \mu)$ be an $\mathbb{R}$ -valued pre-measure space. Then, there exists a unique pair of mutually singular finite measures μ_+ and μ_- on $(X, \mathcal{B})$ such that $\mu = \mu_+ - \mu_-$. Furthermore, $\mathcal{P} \circ \mu = \mu_+ + \mu_-$ and $(X, \mathcal{B}, \mu)$ is a finite $\mathbb{R}$ -valued measure space.*

Proof By Hahn Decomposition Theorem 11.161, $\exists A, B \in \mathcal{B}$ with $A \cap B = \emptyset$ and $X = A \cup B$ such that A is a positive set with respect to μ and B is a negative set with respect to μ . Define $\mu_+ : \mathcal{B} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\mu_+(E) = \mu(A \cap E)$, $\forall E \in \mathcal{B}$ and $\mu_- : \mathcal{B} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\mu_-(E) = -\mu(B \cap E)$, $\forall E \in \mathcal{B}$. It is straightforward to show that μ_+ and μ_- are finite measures on $(X, \mathcal{B})$. Clearly, $\mu_+(B) = \mu_-(A) = 0$. Then, $\mu_+ \perp \mu_-$. It is straightforward to show that $\mu = \mu_+ - \mu_-$. Then, by Proposition 11.136, μ is a finite $\mathbb{R}$ -valued measure on $(X, \mathcal{B})$. By Proposition 11.156, $\mathcal{P} \circ \mu = \mathcal{P} \circ \mu_+ + \mathcal{P} \circ \mu_- = \mu_+ + \mu_-$.

Let $\hat{\mu}_+$ and $\hat{\mu}_-$ be any pair of mutually singular finite measures on $(X, \mathcal{B})$ such that $\mu = \hat{\mu}_+ - \hat{\mu}_-$. Then, $\exists \hat{A}, \hat{B} \in \mathcal{B}$ with $\hat{A} \cap \hat{B} = \emptyset$ and $\hat{A} \cup \hat{B} = X$ such that $\hat{\mu}_+(\hat{B}) = \hat{\mu}_-(\hat{A}) = 0$. Note that $\mu_+(A \cap \hat{A}) = \mu_+(A) - \mu_+(A \setminus \hat{A}) \leq \mu_+(A) = \mu(A) = \hat{\mu}_+(A) - \hat{\mu}_-(A) = \hat{\mu}_+(A \cap \hat{A}) + \hat{\mu}_+(A \setminus \hat{A}) - \hat{\mu}_-(A \cap \hat{A}) - \hat{\mu}_-(A \setminus \hat{A}) = \hat{\mu}_+(A \cap \hat{A}) - \hat{\mu}_-(A \setminus \hat{A}) \leq \hat{\mu}_+(A \cap \hat{A}) = \hat{\mu}_+(\hat{A}) - \hat{\mu}_+(\hat{A} \setminus A) \leq \hat{\mu}_+(\hat{A}) = \mu(\hat{A}) = \mu_+(\hat{A}) - \mu_-(\hat{A}) = \mu_+(A \cap \hat{A}) + \mu_+(\hat{A} \setminus A) - \mu_-(A \cap \hat{A}) - \mu_-(\hat{A} \setminus A) = \mu_+(A \cap \hat{A}) - \mu_-(\hat{A} \setminus A) \leq \mu_+(A \cap \hat{A}) \in \mathbb{R}$. This implies that $\mu_+(A \cap \hat{A}) = \mu_+(A) = \mu(A) = \hat{\mu}_+(A \cap \hat{A}) = \hat{\mu}_+(\hat{A}) = \mu(\hat{A})$ and $\mu_+(A \setminus \hat{A}) = \hat{\mu}_+(A \setminus \hat{A}) = \hat{\mu}_+(\hat{A} \setminus A) = \mu_-(\hat{A} \setminus A) = 0$. $\forall E \in \mathcal{B}$, we have $\mu_+(E) = \mu_+(E \cap A) = \mu_+((E \cap A) \setminus \hat{A}) + \mu_+(E \cap A \cap \hat{A}) = \mu_+(E \cap A \cap \hat{A}) = \mu(E \cap A \cap \hat{A}) = \hat{\mu}_+(E \cap A \cap \hat{A}) = \hat{\mu}_+(E \cap \hat{A}) - \hat{\mu}_+((E \cap \hat{A}) \setminus A) = \hat{\mu}_+(E \cap \hat{A}) = \hat{\mu}_+(E) - \hat{\mu}_+(E \setminus \hat{A}) = \hat{\mu}_+(E) \in \mathbb{R}$ and $\mu_-(E) = \mu_+(E) - \mu(E) = \hat{\mu}_+(E) - \mu(E) = \hat{\mu}_-(E) \in \mathbb{R}$. Hence, $\mu_+ = \hat{\mu}_+$ and $\mu_- = \hat{\mu}_-$. This shows that the pair μ_+ and μ_- is unique.

This completes the proof of the theorem. $\square$

Theorem 11.163 (Jordan Decomposition Theorem) *Let $(X, \mathcal{B}, \mu) =: \mathcal{X}$ be an σ -finite $\mathbb{R}$ -valued measure space. Then, there exists a unique pair of mutually singular σ -finite measures μ_+ and μ_- on $(X, \mathcal{B})$ such that $\mu = \mu_+ - \mu_-$. Furthermore, $\mathcal{P} \circ \mu = \mu_+ + \mu_-$.*

Proof By Hahn Decomposition Theorem 11.161, there exists $A, B \in \mathcal{B}$ such that $A = X \setminus B$, A is a positive set with respect to μ , and B is a negative set with respect to μ . Define μ_A to be a function from $\mathcal{B}$ to $\mathbb{R}$ by: $\mu_A(E)$ is undefined if $E \in \mathcal{B}$ and $E \cap A \in \mathcal{B} \setminus \text{dom}(\mu)$; $\mu_A(E) = \mu(E \cap A) \in [0, \infty) \subset \mathbb{R}$ if $E \in \mathcal{B}$ and $E \cap A \in \text{dom}(\mu)$. Define μ_B to be a function from $\mathcal{B}$ to $\mathbb{R}$ by: $\mu_B(E)$ is undefined if $E \in \mathcal{B}$ and $E \cap B \in \mathcal{B} \setminus \text{dom}(\mu)$; $\mu_B(E) = \mu(E \cap B) \in (-\infty, 0] \subset \mathbb{R}$ if $E \in \mathcal{B}$ and $E \cap B \in \text{dom}(\mu)$. Then, by Proposition 11.157, μ_A and μ_B are σ -finite $\mathbb{R}$ -valued measures on $(X, \mathcal{B})$ such that $\mu = \mu_A + \mu_B$, $\mathcal{P} \circ \mu = \mathcal{P} \circ \mu_A + \mathcal{P} \circ \mu_B$, and $\mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B$. Since $\mu_A(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\mu_A)$, then μ_A is identified with a σ -finite measure μ_+ on $(X, \mathcal{B})$ according to Proposition 11.135. By Proposition 11.138, $-\mu_B$ is a σ -finite $\mathbb{R}$ -valued measure on $(X, \mathcal{B})$. Since $-\mu_B(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\mu_B)$, then $-\mu_B$ is identified with a σ -finite measure μ_- on $(X, \mathcal{B})$ according to Proposition 11.135. Then, $\mathcal{P} \circ \mu = \mathcal{P} \circ \mu_A + \mathcal{P} \circ (-\mu_B) = \mu_+ + \mu_-$. Since $\mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B$, then, $\mu_+ \perp \mu_-$. $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu(E) = \mu_+(E) + \mu_-(E) < \infty$, we have $\mu(E) = (\mu_A + \mu_B)(E) = \mu_A(E) + \mu_B(E) = \mu_+(E) - \mu_-(E) = (\mu_+ - \mu_-)(E) \in \mathbb{R}$. By Propositions 11.137 and 11.138, we have $\mu = \mu_+ - \mu_-$. Hence, the pair μ_+ and μ_- is the pair of mutually singular σ -finite measures on $(X, \mathcal{B})$ that we seek.

Next, we will show that the pair is unique. Let $\hat{\mu}_+$ and $\hat{\mu}_-$ be another pair of mutually singular σ -finite measures on $(X, \mathcal{B})$ such that $\mu = \hat{\mu}_+ - \hat{\mu}_-$ and $\mathcal{P} \circ \mu = \hat{\mu}_+ + \hat{\mu}_-$. Since the four measures μ_+ , μ_- , $\hat{\mu}_+$, and $\hat{\mu}_-$ are σ -finite, then $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty X_n$ and $\mu_+(X_n) \in \mathbb{R}$, $\mu_-(X_n) \in \mathbb{R}$, $\hat{\mu}_+(X_n) \in \mathbb{R}$, and $\hat{\mu}_-(X_n) \in \mathbb{R}$, $\forall n \in \mathbb{N}$. Then, $\mathcal{P} \circ$

$\mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Fix any $n \in \mathbb{N}$. Let $\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite $\mathbb{R}$ -valued measure subspace of $\mathcal{X}$, $(X_n, \mathcal{B}_n, \mu_{n,+})$ be the finite measure subspace of $(X, \mathcal{B}, \mu_+)$, $(X_n, \mathcal{B}_n, \mu_{n,-})$ be the finite measure subspace of $(X, \mathcal{B}, \mu_-)$, $(X_n, \mathcal{B}_n, \hat{\mu}_{n,+})$ be the finite measure subspace of $(X, \mathcal{B}, \hat{\mu}_+)$, and $(X_n, \mathcal{B}_n, \hat{\mu}_{n,-})$ be the finite measure subspace of $(X, \mathcal{B}, \hat{\mu}_-)$. Then, it is easy to see that $\mu_n = \mu_{n,+} - \mu_{n,-} = \hat{\mu}_{n,+} - \hat{\mu}_{n,-}$, $\mu_{n,+} \perp \mu_{n,-}$, $\hat{\mu}_{n,+} \perp \hat{\mu}_{n,-}$, and $\mathcal{P} \circ \mu_n = \mu_{n,+} + \mu_{n,-} = \hat{\mu}_{n,+} + \hat{\mu}_{n,-}$. By Jordan Decomposition Theorem 11.162, we have $\mu_{n,+} = \hat{\mu}_{n,+}$ and $\mu_{n,-} = \hat{\mu}_{n,-}$. By Proposition 11.117, we have $\mu_+ = \hat{\mu}_+$ and $\mu_- = \hat{\mu}_-$. Hence, the pair is unique.

This completes the proof of the theorem. $\square$

Proposition 11.164 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathbb{C}$ -valued pre-measure space. Then, $\mu = \mu_r + i\mu_i$, where $\mu_r : \mathcal{B} \rightarrow \mathbb{R}$ and $\mu_i : \mathcal{B} \rightarrow \mathbb{R}$ are finite $\mathbb{R}$ -valued measures on $(X, \mathcal{B})$, and $\max\{\mathcal{P} \circ \mu_r(E), \mathcal{P} \circ \mu_i(E)\} \leq \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E)$, $\forall E \in \mathcal{B}$. Then, $\mathcal{X}$ is a finite $\mathbb{C}$ -valued measure space.*

Proof Define $\mu_r : \mathcal{B} \rightarrow \mathbb{R}$ and $\mu_i : \mathcal{B} \rightarrow \mathbb{R}$ by $\mu_r(E) = \operatorname{Re}(\mu(E))$ and $\mu_i(E) = \operatorname{Im}(\mu(E))$, $\forall E \in \mathcal{B}$. Clearly, $\mu = \mu_r + i\mu_i$, $\mu_r(\emptyset) = 0$, and $\mu_i(\emptyset) = 0$. $\forall$ pairwise disjoint sequence $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$, $\sum_{n=1}^\infty |\mu_r(E_n)| \leq \sum_{n=1}^\infty |\mu(E_n)| < +\infty$, where the last inequality follows from the fact that μ is a $\mathbb{C}$ -value pre-measure. We also have $\sum_{n=1}^\infty \mu_r(E_n) = \sum_{n=1}^\infty \operatorname{Re}(\mu(E_n)) = \operatorname{Re}(\sum_{n=1}^\infty \mu(E_n)) = \operatorname{Re}(\mu(\bigcup_{n=1}^\infty E_n)) = \mu_r(\bigcup_{n=1}^\infty E_n)$. Hence, μ_r is an $\mathbb{R}$ -valued pre-measure on $(X, \mathcal{B})$. Note that $\sum_{n=1}^\infty |\mu_i(E_n)| \leq \sum_{n=1}^\infty |\mu(E_n)| < +\infty$ and $\sum_{n=1}^\infty \mu_i(E_n) = \sum_{n=1}^\infty \operatorname{Im}(\mu(E_n)) = \operatorname{Im}(\sum_{n=1}^\infty \mu(E_n)) = \operatorname{Im}(\mu(\bigcup_{n=1}^\infty E_n)) = \mu_i(\bigcup_{n=1}^\infty E_n)$. Hence, μ_i is an $\mathbb{R}$ -valued pre-measure on $(X, \mathcal{B})$. By Jordan Decomposition Theorem 11.162, μ_r and μ_i are finite $\mathbb{R}$ -valued measures on $(X, \mathcal{B})$.

$\forall E \in \mathcal{B}$, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_k)_{k=1}^n \subseteq \mathcal{B}$ with $E = \bigcup_{k=1}^n E_k$, $\sum_{k=1}^n |\mu_r(E_k)| \leq \sum_{k=1}^n |\mu(E_k)| \leq \mathcal{P} \circ \mu(E)$. Hence, $\mathcal{P} \circ \mu_r(E) \leq \mathcal{P} \circ \mu(E)$. $\forall E \in \mathcal{B}$, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_k)_{k=1}^n \subseteq \mathcal{B}$ with $E = \bigcup_{k=1}^n E_k$, $\sum_{k=1}^n |\mu_i(E_k)| \leq \sum_{k=1}^n |\mu(E_k)| \leq \mathcal{P} \circ \mu(E)$. Hence, $\mathcal{P} \circ \mu_i(E) \leq \mathcal{P} \circ \mu(E)$. $\forall E \in \mathcal{B}$, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_k)_{k=1}^n \subseteq \mathcal{B}$ with $E = \bigcup_{k=1}^n E_k$, $\sum_{k=1}^n |\mu(E_k)| = \sum_{k=1}^n \sqrt{(\mu_r(E_k))^2 + (\mu_i(E_k))^2} \leq \sum_{k=1}^n (|\mu_r(E_k)| + |\mu_i(E_k)|) \leq \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E)$. Hence, $\mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E)$. Then, $\mathcal{P} \circ \mu(X) < +\infty$. Hence, $\mathcal{X}$ is a finite $\mathbb{C}$ -valued measure space.

This completes the proof of the proposition. $\square$

Proposition 11.165 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite $\mathbb{C}$ -valued measure space. Then, there exists a pair of σ -finite $\mathbb{R}$ -valued measures μ_r and μ_i on $(X, \mathcal{B})$ such that $\mu = \mu_r + i\mu_i$ and $0 \leq \max\{\mathcal{P} \circ \mu_r(E), \mathcal{P} \circ \mu_i(E)\} \leq \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E)$, $\forall E \in \mathcal{B}$.*

Proof Since $\mathcal{X}$ is σ -finite, then $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty X_n$ and $\mathcal{P} \circ \mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality,

we may assume that $(X_n)_{n=1}^\infty$ is pairwise disjoint. Fix any $n \in \mathbb{N}$. Let $\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite $\mathbb{C}$ -valued measure subspace of $\mathcal{X}$. Then, $\mathcal{X}_n$ is a finite $\mathbb{C}$ -valued pre-measure space. By Proposition 11.164, there exists finite $\mathbb{R}$ -valued measures $\mu_{n,r}$ and $\mu_{n,i}$ on $(X_n, \mathcal{B}_n)$ such that $\mu_n = \mu_{n,r} + i\mu_{n,i}$ and $0 \leq \max\{\mathcal{P} \circ \mu_{n,r}(E), \mathcal{P} \circ \mu_{n,i}(E)\} \leq \mathcal{P} \circ \mu_n(E) \leq \mathcal{P} \circ \mu_{n,r}(E) + \mathcal{P} \circ \mu_{n,i}(E) < \infty, \forall E \in \mathcal{B}_n$. By Proposition 11.118, the generation process on $((X_n, \mathcal{B}_n, \mu_{n,r}))_{n=1}^\infty$ yields a unique σ -finite $\mathbb{R}$ -valued measure space $(X, \mathcal{B}, \mu_r)$ on X and the generation process on $((X_n, \mathcal{B}_n, \mu_{n,i}))_{n=1}^\infty$ yields a unique σ -finite $\mathbb{R}$ -valued measure space $(X, \mathcal{B}, \mu_i)$ on X .

$\forall E \in \mathcal{B}$, let $E_n := E \cap X_n \in \mathcal{B}_n, \forall n \in \mathbb{N}$. Then, $\mathcal{P} \circ \mu(E) = \sum_{n=1}^\infty \mathcal{P} \circ \mu(E_n) = \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(E_n)$, where the first equality follows from the fact that $\mathcal{P} \circ \mu$ is a measure; and the second equality follows from Propositions 11.118 and 11.115. By Propositions 11.118 and 11.115, we have $\mathcal{P} \circ \mu(E) \geq \sum_{n=1}^\infty \mathcal{P} \circ \mu_{n,r}(E_n) = \sum_{n=1}^\infty \mathcal{P} \circ \mu_r(E_n) = \mathcal{P} \circ \mu_r(E)$. Similarly, we have $\mathcal{P} \circ \mu(E) \geq \mathcal{P} \circ \mu_i(E)$. Again, by Propositions 11.118 and 11.115, $\mathcal{P} \circ \mu(E) \leq \sum_{n=1}^\infty (\mathcal{P} \circ \mu_{n,r}(E_n) + \mathcal{P} \circ \mu_{n,i}(E_n)) = \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E)$. Hence, we have $0 \leq \max\{\mathcal{P} \circ \mu_r(E), \mathcal{P} \circ \mu_i(E)\} \leq \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E) \leq \infty, \forall E \in \mathcal{B}$.

By Proposition 11.138, $\mu_r + i\mu_i$ is a σ -finite $\mathbb{C}$ -valued measure on $(X, \mathcal{B})$. $\forall n \in \mathbb{N}, \forall E \in \mathcal{B}_n, \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu(X_n) < \infty$ and $E \in \text{dom}(\mu)$. This implies that $\max\{\mathcal{P} \circ \mu_r(E), \mathcal{P} \circ \mu_i(E)\} \leq \mathcal{P} \circ \mu(E) < \infty$ and $E \in \text{dom}(\mu_r) \cap \text{dom}(\mu_i)$. Then, $(\mu_r + i\mu_i)(E) = \mu_r(E) + i\mu_i(E) = \mu_{n,r}(E) + i\mu_{n,i}(E) = (\mu_{n,r} + i\mu_{n,i})(E) = \mu_n(E) = \mu(E)$, where the first equality follows from Proposition 11.138; the second equality follows from Proposition 11.115; the third equality follows from Proposition 11.138; and the last equality follows from Proposition 11.115. Hence, $\mathcal{X}_n$ is the finite $\mathbb{C}$ -valued measure subspace of $(X, \mathcal{B}, \mu_r + i\mu_i), \forall n \in \mathbb{N}$. By Proposition 11.117, we have $\mu = \mu_r + i\mu_i$.

This completes the proof of the proposition. $\square$

Definition 11.166 Let $(X, \mathcal{B})$ be a measurable space, λ_1 and λ_2 be measures on $(X, \mathcal{B})$, and $f : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable. We will say that f is a Radon-Nikodym derivative of λ_1 with respect to λ_2 if $\lambda_1(E) = \int_E f d\lambda_2, \forall E \in \mathcal{B}$. Let μ be a $\mathbb{K}$ -valued measure on $(X, \mathcal{B})$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, ν be a $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$, and $\bar{f} : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. We will say that $\bar{f}$ is a Radon-Nikodym derivative of ν with respect to μ if $\mathcal{P} \circ \bar{f}$ is a Radon-Nikodym derivative of $\mathcal{P} \circ \nu$ with respect to $\mathcal{P} \circ \mu$ (that is, $\mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ \bar{f} d\mathcal{P} \circ \mu \in [0, \infty) \subset \mathbb{R}_e, \forall E \in \mathcal{B}$) and $\nu(E) = \int_E \bar{f} d\mu \in \mathcal{Y}, \forall E \in \text{dom}(\nu)$.

It is easy to see that the definition above is compatible with the identification of σ -finite $\mathbb{R}$ -valued measures and σ -finite measures eluded to in Section 11.7. It is straightforward to show that if $\bar{f} : X \rightarrow \mathcal{Y}$ is a Radon-Nikodym derivative of ν with respect to μ , then $\mathcal{P} \circ \nu \ll \mathcal{P} \circ \mu$. When μ is σ -finite, $\mathcal{Y}$ is a Banach space, and $\mathcal{W}$ is a separable subspace of $\mathcal{Y}$, then Proposition 11.154 guarantees that any measurable function $\bar{f} : X \rightarrow \mathcal{W}$

is a Radon-Nikodym derivative of some σ -finite $\mathcal{Y}$ -valued measure ν with respect to μ . The next result gives the sufficient condition under which the Radon-Nikodym derivative is unique. The Radon-Nikodym Theorems then give sufficient conditions on ν and μ under which the Radon-Nikodym derivative exists.

Proposition 11.167 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite $\mathbb{K}$ -valued measure space, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, ν be a σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$, and $f : X \rightarrow \mathcal{Y}$ be a Radon-Nikodym derivative of the ν with respect to μ , and $g : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Then, g is a Radon-Nikodym derivative of ν with respect to μ if, and only if, $g = f$ a.e. in $\mathcal{X}$. In this case, we will write $\frac{d\nu}{d\mu} = f$ a.e. in $\mathcal{X}$.*

Proof “Sufficiency” Let $g = f$ a.e. in $\mathcal{X}$. Then, by Propositions 7.21, 7.23, 11.38, and 11.39, $\mathcal{P} \circ g = \mathcal{P} \circ f$ a.e. in $\mathcal{X}$. $\forall E \in \mathcal{B}$, $\mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu = \int_E \mathcal{P} \circ g d\mathcal{P} \circ \mu$, where the second equality follows from Proposition 11.83. $\forall E \in \text{dom}(\nu)$, $\infty > \mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu = \int_E \mathcal{P} \circ g d\mathcal{P} \circ \mu$. Then, $f|_E$ and $g|_E$ are absolutely integrable over $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$, which is the $\mathbb{K}$ -valued measure subspace of $\mathcal{X}$. This implies that, by Proposition 11.132, $\nu(E) = \int_E f d\mu = \int_E g d\mu \in \mathcal{Y}$. Hence, g is a Radon-Nikodym derivative of ν with respect to μ .

“Necessity” Let $g : X \rightarrow \mathcal{Y}$ be another Radon-Nikodym derivative of ν with respect to μ . Define $\lambda := \nu - \nu$, which is a σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$, by Proposition 11.138. Then, by Proposition 11.137, $\lambda(E) = \vartheta_{\mathcal{Y}}$ and $\mathcal{P} \circ \lambda(E) = 0$, $\forall E \in \mathcal{B}$. By Proposition 11.154, we may define a σ -finite $\mathcal{Y}$ -valued measure $\bar{\lambda}$ on $(X, \mathcal{B})$ such that $\mathcal{P} \circ \bar{\lambda}(E) = \int_E \mathcal{P} \circ (f - g) d\mathcal{P} \circ \mu$, $\forall E \in \mathcal{B}$, and $\bar{\lambda}(E) = \int_E (f - g) d\mu$, $\forall E \in \mathcal{B}$ with $\int_E \mathcal{P} \circ (f - g) d\mathcal{P} \circ \mu < \infty$. $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \nu(E) < \infty$, then $f|_E$ and $g|_E$ are absolutely integrable over $\mathcal{E}$. By Proposition 11.132, $(f - g)|_E$ is absolutely integrable over $\mathcal{E}$. Then, we have $\vartheta_{\mathcal{Y}} = \lambda(E) = \nu(E) - \nu(E) = \int_E f d\mu - \int_E g d\mu = \int_E (f - g) d\mu = \bar{\lambda}(E)$. By Proposition 11.137, we have $\lambda = \bar{\lambda}$. Then, $\forall E \in \mathcal{B}$, $E \in \text{dom}(\lambda)$, $0 = \mathcal{P} \circ \lambda(E) = \mathcal{P} \circ \bar{\lambda}(E) = \int_E \mathcal{P} \circ (f - g) d\mathcal{P} \circ \mu$. By Proposition 11.96, $\mathcal{P} \circ (f - g) = 0$ a.e. in $\mathcal{X}$ and $f = g$ a.e. in $\mathcal{X}$.

This completes the proof of the proposition. $\square$

Proposition 11.168 *Let $(X, \mathcal{B})$ be a measurable space, λ and λ_1 be σ -finite measures on $(X, \mathcal{B})$, $\hat{\mathcal{X}} := (X, \mathcal{B}, \lambda)$, $\hat{\mathcal{X}}_1 := (X, \mathcal{B}, \lambda_1)$, $\mu, \mu_1 \in \mathcal{M}_{\sigma}(X, \mathcal{B}, \mathbb{K})$, $\mathcal{X} := (X, \mathcal{B}, \mu)$, $\mathcal{X}_1 := (X, \mathcal{B}, \mu_1)$, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, $\nu_1, \nu_2 \in \mathcal{M}_{\sigma}(X, \mathcal{B}, \mathcal{Y})$, $\bar{\mathcal{X}}_1 := (X, \mathcal{B}, \nu_1)$, $\frac{d\nu_i}{d\mu} =: \bar{f}_i : X \rightarrow \mathcal{Y}$ a.e. in $\mathcal{X}$, $i = 1, 2$, $\frac{d\lambda}{d\lambda_1} =: \hat{f} : X \rightarrow [0, \infty) \subset \mathbb{R}$ a.e. in $\hat{\mathcal{X}}_1$, $\frac{d\mu}{d\mu_1} =: f : X \rightarrow \mathbb{K}$ a.e. in $\mathcal{X}_1$, $g : X \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable, $\bar{f} : X \rightarrow [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$ -measurable, $\hat{f} : X \rightarrow \mathbb{K}$ be $\mathcal{B}$ -measurable, $\alpha_0 \in \mathbb{K}$, $A \in B(\mathcal{Y}, \mathcal{Z})$, and $y_0 \in \mathcal{Y}$. Then, the following statements hold.*

$$(i) \int_X \bar{f} d\lambda = \int_X (\bar{f} \hat{f}) d\lambda_1 \in [0, \infty] \subset \mathbb{R}_e.$$

- (ii) g is absolutely integrable over $\bar{\mathcal{X}}_1$ if, and only if, $(\mathcal{P} \circ g)(\mathcal{P} \circ \bar{f}_1)$ is integrable over $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. In which case, $\int_X g d\nu_1 = \int_X (g\bar{f}_1) d\mu \in \mathbb{Z}$.
- (iii) $\frac{d(\nu_1 + \nu_2)}{d\mu} = \bar{f}_1 + \bar{f}_2 = \frac{d\nu_1}{d\mu} + \frac{d\nu_2}{d\mu}$ a.e. in $\mathcal{X}$.
- (iv) $A\bar{f}_1$ is a Radon-Nikodym derivative of $A\nu_1$ with respect to μ .
- (v) $\frac{d(\alpha_0\nu_1)}{d\mu} = \alpha_0\bar{f}_1 = \alpha_0\frac{d\nu_1}{d\mu}$ a.e. in $\mathcal{X}$.
- (vi) $\frac{d(\mu y_0)}{d\mu_1} = f y_0 = \left(\frac{d\mu}{d\mu_1}\right) y_0$ a.e. in $\mathcal{X}_1$.
- (vii) $\frac{d\nu_1}{d\mu_1} = \bar{f}_1 f = \left(\frac{d\nu_1}{d\mu}\right) \left(\frac{d\mu}{d\mu_1}\right)$ a.e. in $\mathcal{X}_1$.
- (viii) $\frac{d\mu_1}{d\mu} = \hat{f}$ a.e. in $\mathcal{X}$ if, and only if, $\hat{f}f = 1$ a.e. in $\mathcal{X}_1$.

Proof (i) By Definition 11.166, $\hat{f}$ is $\mathcal{B}$ -measurable and $\lambda(E) = \int_E \hat{f} d\lambda_1 \in \mathbb{R}_e$, $\forall E \in \mathcal{B}$. By Propositions 7.23, 11.38, and 11.39, $\bar{f}\hat{f}$ is $\mathcal{B}$ -measurable. By Proposition 11.66, there exists a sequence of simple functions $(\bar{\phi}_n)_{n=1}^\infty$, $\bar{\phi}_n : \tilde{\mathcal{X}} \rightarrow [0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, such that $0 \leq \bar{\phi}_n(x) \leq \bar{f}(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \bar{\phi}_n = \bar{f}$ a.e. in $\tilde{\mathcal{X}}$, where $\tilde{\mathcal{X}} := (X, \mathcal{B}, \lambda + \lambda_1)$ is a σ -finite measure space by Proposition 11.136. $\forall n \in \mathbb{N}$, define $\phi_n := \bigvee_{i=1}^n \bar{\phi}_i$. Then, ϕ_n is a simple function of $\tilde{\mathcal{X}}$ to $[0, \infty) \subset \mathbb{R}$, $0 \leq \phi_n(x) \leq \phi_{n+1}(x) \leq \bar{f}(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Proposition 11.50, $\lim_{n \in \mathbb{N}} \phi_n = \bar{f}$ a.e. in $\tilde{\mathcal{X}}$. Since $\lambda \leq \lambda + \lambda_1$, then ϕ_n is a simple function of $\tilde{\mathcal{X}}$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \phi_n = \bar{f}$ a.e. in $\tilde{\mathcal{X}}$. By Monotone Convergence Theorem 11.81, $\int_X \bar{f} d\lambda = \lim_{n \in \mathbb{N}} \int_X \phi_n d\lambda \in \mathbb{R}_e$. Fix any $n \in \mathbb{N}$. Let ϕ_n admit the canonical representation $\phi_n = \sum_{i=1}^{\bar{n}} a_i \chi_{A_i, X}$. Then, by Propositions 11.75 and 11.83, $\int_X \phi_n d\lambda = \sum_{i=1}^{\bar{n}} a_i \lambda(A_i) = \sum_{i=1}^{\bar{n}} a_i \int_{A_i} \hat{f} d\lambda_1 = \sum_{i=1}^{\bar{n}} \int_X (a_i \chi_{A_i, X} \hat{f}) d\lambda_1 = \int_X (\sum_{i=1}^{\bar{n}} a_i \chi_{A_i, X} \hat{f}) d\lambda_1 = \int_X (\phi_n \hat{f}) d\lambda_1$. Note that $0 \leq \phi_n(x) \hat{f}(x) \leq \phi_{n+1}(x) \hat{f}(x) \leq \bar{f}(x) \hat{f}(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Proposition 11.50, $\lim_{n \in \mathbb{N}} \phi_n \hat{f} = \bar{f} \hat{f}$ a.e. in $\tilde{\mathcal{X}}$. Since $\lambda_1 \leq \lambda + \lambda_1$, then $\lim_{n \in \mathbb{N}} \phi_n \hat{f} = \bar{f} \hat{f}$ a.e. in $\tilde{\mathcal{X}}_1$. By Monotone Convergence Theorem 11.81, we have $\int_X (\bar{f} \hat{f}) d\lambda_1 = \lim_{n \in \mathbb{N}} \int_X (\phi_n \hat{f}) d\lambda_1 = \lim_{n \in \mathbb{N}} \int_X \phi_n d\lambda = \int_X \bar{f} d\lambda \in \mathbb{R}_e$.

(ii) Since $\frac{d\nu_1}{d\mu} = \bar{f}_1$ a.e. in $\mathcal{X}$, then $\frac{d\mathcal{P} \circ \nu_1}{d\mathcal{P} \circ \mu} = \mathcal{P} \circ \bar{f}_1$ a.e. in $\mathcal{X}$. By (i), $\int_X \mathcal{P} \circ g d\mathcal{P} \circ \nu_1 = \int_X (\mathcal{P} \circ g)(\mathcal{P} \circ \bar{f}_1) d\mathcal{P} \circ \mu \in \mathbb{R}_e$. Hence, g is absolutely integrable over $\tilde{\mathcal{X}}_1$ if, and only if, $(\mathcal{P} \circ g)(\mathcal{P} \circ \bar{f}_1)$ is integrable over $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. Let g be absolutely integrable over $\tilde{\mathcal{X}}_1$, then $\int_X \mathcal{P} \circ g d\mathcal{P} \circ \nu_1 = \int_X (\mathcal{P} \circ g)(\mathcal{P} \circ \bar{f}_1) d\mathcal{P} \circ \mu < \infty$. By Proposition 11.66, there exists a sequence of simple functions $(\phi_n)_{n=1}^\infty$, $\phi_n : \tilde{\mathcal{X}}_1 \rightarrow \mathbb{W}$, $\forall n \in \mathbb{N}$, such that $\|\phi_n(x)\| \leq \|g(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \phi_n = g$ a.e. in $\tilde{\mathcal{X}}_1$, where $\tilde{\mathcal{X}}_1 := (X, \mathcal{B}, \mathcal{P} \circ \mu + \mathcal{P} \circ \nu_1)$ is a σ -finite measure space by Proposition 11.136. Since $\mathcal{P} \circ \nu_1 \leq \mathcal{P} \circ \mu + \mathcal{P} \circ \nu_1$, then ϕ_n is a simple function

of $\bar{\mathcal{X}}_1$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \phi_n = g$ a.e. in $\bar{\mathcal{X}}_1$. By Lebesgue Dominated Convergence Theorem 11.131, $\int_X g \, d\nu_1 = \lim_{n \in \mathbb{N}} \int_X \phi_n \, d\nu_1 \in \mathcal{Z}$. Fix any $n \in \mathbb{N}$. Let ϕ_n admit the canonical representation $\phi_n = \sum_{i=1}^{\bar{n}} w_i \chi_{A_i, X}$. Note that $\hat{\mathcal{W}} := \text{span}(\{z \in \mathcal{Z} \mid z = wy, w \in \mathcal{W}, y \in \mathcal{Y}\})$ is a separable subspace of $\mathcal{Z}$. Then, by Propositions 11.125 and 11.132, we have $\int_X \phi_n \, d\nu_1 = \sum_{i=1}^{\bar{n}} w_i \nu_1(A_i) = \sum_{i=1}^{\bar{n}} w_i \int_{A_i} \bar{f}_1 \, d\mu = \sum_{i=1}^{\bar{n}} \int_X (\chi_{A_i, X} w_i \bar{f}_1) \, d\mu = \int_X (\sum_{i=1}^{\bar{n}} \chi_{A_i, X} w_i \bar{f}_1) \, d\mu = \int_X (\phi_n \bar{f}_1) \, d\mu$. Note that $\mathcal{P} \circ (\phi_n \bar{f}_1)(x) = \|\phi_n(x) \bar{f}_1(x)\| \leq \|\phi_n(x)\| \|\bar{f}_1(x)\| \leq (\mathcal{P} \circ g(x)) (\mathcal{P} \circ \bar{f}_1(x))$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Note also that $\phi_n \bar{f}_1 : X \rightarrow \hat{\mathcal{W}} \subseteq \mathcal{Z}$ and $g \bar{f}_1 : X \rightarrow \hat{\mathcal{W}}$, $\forall n \in \mathbb{N}$. By Propositions 7.65, 11.38, and 11.39, $\phi_n \bar{f}_1$ and $g \bar{f}_1$ are $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. By Proposition 11.50, $\lim_{n \in \mathbb{N}} \phi_n \bar{f}_1 = g \bar{f}_1$ a.e. in $\bar{\mathcal{X}}_1$. Since $\mathcal{P} \circ \mu \leq \mathcal{P} \circ \mu + \mathcal{P} \circ \nu_1$, then $\lim_{n \in \mathbb{N}} \phi_n \bar{f}_1 = g \bar{f}_1$ a.e. in $\mathcal{X}$. By Lebesgue Dominated Convergence Theorem 11.131, $\int_X (g \bar{f}_1) \, d\mu = \lim_{n \in \mathbb{N}} \int_X (\phi_n \bar{f}_1) \, d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n \, d\nu_1 = \int_X g \, d\nu_1 \in \mathcal{Z}$.

(iii) By Proposition 11.154, $\bar{f}_1 + \bar{f}_2$ is a Radon-Nikodym derivative of a σ -finite $\mathcal{Y}$ -valued measure $\bar{\nu}$ on $(X, \mathcal{B})$ with respect to μ . $\forall E \in \mathcal{B}$ with $(\mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 + \mathcal{P} \circ \bar{\nu})(E) < \infty$, then $\mathcal{P} \circ \nu_i(E) = \int_E \mathcal{P} \circ f_i \, d\mathcal{P} \circ \mu < \infty$, $i = 1, 2$. This implies that $\bar{f}_i|_E$, $i = 1, 2$, are absolutely integrable over $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$, which is the $\mathbb{K}$ -valued measure subspace of $\mathcal{X}$. Then, $\bar{\nu}(E) = \int_E (\bar{f}_1 + \bar{f}_2) \, d\mu = \int_E \bar{f}_1 \, d\mu + \int_E \bar{f}_2 \, d\mu = \nu_1(E) + \nu_2(E) = (\nu_1 + \nu_2)(E) \in \mathcal{Y}$, where the second equality follows from Proposition 11.132; and the last equality follows from Proposition 11.138. Note that $\mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 + \mathcal{P} \circ \bar{\nu}$ is a σ -finite measure on $(X, \mathcal{B})$ by Proposition 11.136. By Proposition 11.137, we have $\bar{\nu} = \nu_1 + \nu_2$. Hence, by Proposition 11.167, $\frac{d(\nu_1 + \nu_2)}{d\mu} = \bar{f}_1 + \bar{f}_2$ a.e. in $\mathcal{X}$.

(iv) Note that $A\bar{f}_1 : X \rightarrow \mathcal{R}(A) \subseteq \mathcal{Z}$, where $\mathcal{R}(A)$ is a separable subspace of $\mathcal{Z}$ since $\mathcal{Y}$ is separable. By Proposition 11.154, $A\bar{f}_1$ is a Radon-Nikodym derivative of a σ -finite $\mathcal{Z}$ -valued measure $\bar{\nu}$ on $(X, \mathcal{B})$ with respect to μ . $\forall E \in \mathcal{B}$ with $(\mathcal{P} \circ \nu_1 + \mathcal{P} \circ \bar{\nu})(E) < \infty$, then $\mathcal{P} \circ \nu_1(E) = \int_E \mathcal{P} \circ \bar{f}_1 \, d\mathcal{P} \circ \mu < \infty$. This implies that $\bar{f}_1|_E$ is absolutely integrable over $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$, which is the $\mathbb{K}$ -valued measure subspace of $\mathcal{X}$. Then, $\bar{\nu}(E) = \int_E (A\bar{f}_1) \, d\mu = A \int_E \bar{f}_1 \, d\mu = A\nu_1(E) = (A\nu_1)(E) \in \mathcal{Z}$, where the second equality follows from Proposition 11.132; and the last equality follows from Proposition 11.138. Note that $\mathcal{P} \circ \nu_1 + \mathcal{P} \circ \bar{\nu}$ is a σ -finite measure on $(X, \mathcal{B})$ by Proposition 11.136. By Proposition 11.137, we have $\bar{\nu} = A\nu_1$. Hence, $A\bar{f}_1$ is a Radon-Nikodym derivative of $A\nu_1$ with respect to μ .

(v) and (vi) follows immediately from (iv) and Proposition 11.167.

(vii) By Proposition 11.154, $\bar{f}_1 f$ is a Radon-Nikodym derivative of a σ -finite $\mathcal{Y}$ -valued measure $\bar{\nu}$ on $(X, \mathcal{B})$ with respect to μ_1 . $\forall E \in \mathcal{B}$ with $(\mathcal{P} \circ \mu + \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \bar{\nu})(E) < \infty$, then $\mathcal{P} \circ \nu_1(E) = \int_E \mathcal{P} \circ \bar{f}_1 \, d\mathcal{P} \circ \mu < \infty$, $\mathcal{P} \circ \bar{\nu}(E) = \int_E \mathcal{P} \circ (\bar{f}_1 f) \, d\mathcal{P} \circ \mu_1 = \int_E (\mathcal{P} \circ \bar{f}_1) (\mathcal{P} \circ f) \, d\mathcal{P} \circ \mu_1 < \infty$, and $\bar{\nu}(E) = \int_E (\bar{f}_1 f) \, d\mu_1 = \int_E \bar{f}_1 \, d\mu = \nu_1(E) \in \mathcal{Y}$, where the second equality follows from (ii). Note that $\mathcal{P} \circ \mu + \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \bar{\nu}$ is a σ -finite measure on $(X, \mathcal{B})$ by Proposition 11.136. By Proposition 11.137, we have

$\bar{\nu} = \nu_1$. Hence, by Proposition 11.167, $\frac{d\nu_1}{d\mu_1} = \bar{f}_1 f$ a.e. in $\mathcal{X}_1$.

(viii) “Necessity” Let $\frac{d\mu_1}{d\mu} = \hat{f}$ a.e. in $\mathcal{X}$. By (vii), $\frac{d\mu_1}{d\mu_1} = \hat{f}f$ a.e. in $\mathcal{X}_1$. Clearly, the constant function 1 of X is a Radon-Nikodym derivative of μ_1 with respect to μ_1 . By Proposition 11.167, we have $\hat{f}f = 1$ a.e. in $\mathcal{X}_1$.

“Sufficiency” Let $\hat{f}$ be such that $\hat{f}f = 1$ a.e. in $\mathcal{X}_1$. By Proposition 11.154, $\hat{f}$ is a Radon-Nikodym derivative of a σ -finite $\mathbb{K}$ -valued measure $\bar{\mu}$ on $(X, \mathcal{B})$ with respect to μ . $\forall E \in \mathcal{B}$ with $(\mathcal{P} \circ \mu_1 + \mathcal{P} \circ \bar{\mu})(E) < \infty$, $\bar{\mu}(E) = \int_E \hat{f} d\mu = \int_E (\hat{f}f) d\mu_1 = \int_E 1 d\mu_1 = \mu_1(E) \in \mathbb{K}$, where the second equality follows from (ii); the third equality follows from Proposition 11.132; and the last equality follows from Proposition 11.125. By Proposition 11.137, we have $\bar{\mu} = \mu_1$. Then, by Proposition 11.167, $\frac{d\mu_1}{d\mu} = \hat{f}$ a.e. in $\mathcal{X}$.

This completes the proof of the proposition. $\square$

Theorem 11.169 (Radon-Nikodym) *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, ν be a σ -finite $\mathbb{R}$ -valued measure on $(X, \mathcal{B})$. Assume that $\mathcal{P} \circ \nu \ll \mu$. Then, there exists $f : X \rightarrow \mathbb{R}$ such that f is the Radon-Nikodym derivative of ν with respect to μ . The function f is unique in the sense that $g : X \rightarrow \mathbb{R}$ is another function with this property if, and only if, g is $\mathcal{B}$ -measurable and $f = g$ a.e. in $\mathcal{X}$. Hence, $\frac{d\nu}{d\mu} = f$ a.e. in $\mathcal{X}$. Furthermore, when ν is a σ -finite measure on $(X, \mathcal{B})$, then we may take $f : X \rightarrow [0, \infty) \subset \mathbb{R}$.*

Proof We will show the existence of f in three steps. First, consider the special case that $\mathcal{X}$ is a finite measure space and ν is a finite measure on $(X, \mathcal{B})$. Define $\mathcal{F} := \{f : X \rightarrow [0, \infty) \subset \mathbb{R} \mid f \text{ is } \mathcal{B}\text{-measurable, and } \forall A \in \mathcal{B}, \int_A f d\mu \leq \nu(A)\}$. Clearly, the identically 0 function belongs to $\mathcal{F} \neq \emptyset$. Let $\lambda := \sup_{f \in \mathcal{F}} \int_X f d\mu \in \mathbb{R}_e$. Then, $\lambda \geq 0$. Clearly, $\lambda \leq \nu(X) < +\infty$, by the definition of $\mathcal{F}$. $\forall f_1, f_2 \in \mathcal{F}$, by Propositions 11.38, 11.39, 7.23, $f_1 - f_2$ is $\mathcal{B}$ -measurable. Then, $B := \{x \in X \mid f_1(x) - f_2(x) > 0\} \in \mathcal{B}$ and $C := \{x \in X \mid f_1(x) - f_2(x) \leq 0\} \in \mathcal{B}$. Clearly, $B = X \setminus C$. By Proposition 11.40, $f_1 \vee f_2$ is $\mathcal{B}$ -measurable. $\forall A \in \mathcal{B}$, by Proposition 11.83, $\int_A (f_1 \vee f_2) d\mu = \int_{A \cap B} (f_1 \vee f_2) d\mu + \int_{A \cap C} (f_1 \vee f_2) d\mu = \int_{A \cap B} f_1 d\mu + \int_{A \cap C} f_2 d\mu \leq \nu(A \cap B) + \nu(A \cap C) = \nu(A)$. Therefore, $f_1 \vee f_2 \in \mathcal{F}$. By the definition of λ , there exists $(\tilde{f}_n)_{n=1}^\infty \subseteq \mathcal{F}$ such that $\lambda = \lim_{n \in \mathbb{N}} \int_X \tilde{f}_n d\mu$. $\forall n \in \mathbb{N}$, let $\tilde{f}_n := \bigvee_{i=1}^n \tilde{f}_i \in \mathcal{F}$. Then, $\tilde{f}_n(x) \leq \tilde{f}_{n+1}(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Proposition 11.83 and the definition of λ , we have $\int_X \tilde{f}_n d\mu \leq \int_X \tilde{f}_{n+1} d\mu \leq \lambda$. Then, $\lambda = \lim_{n \in \mathbb{N}} \int_X \tilde{f}_n d\mu \in [0, \nu(X)] \subset \mathbb{R}$. By Proposition 11.82 and Monotone Convergence Theorem 11.81, $\exists f : X \rightarrow [0, \infty) \subset \mathbb{R}$, which is $\mathcal{B}$ -measurable, such that $\lim_{n \in \mathbb{N}} \tilde{f}_n = f$ a.e. in $\mathcal{X}$ and $\int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \tilde{f}_n d\mu = \lambda$. $\forall A \in \mathcal{B}$, by Monotone Convergence Theorem 11.81, we have $\int_A f d\mu = \int_X (f \chi_{A,X}) d\mu = \lim_{n \in \mathbb{N}} \int_X (\tilde{f}_n \chi_{A,X}) d\mu = \lim_{n \in \mathbb{N}} \int_A \tilde{f}_n d\mu \leq \nu(A)$. Then, $f \in \mathcal{F}$. Clearly, f is integrable over $\mathcal{X}$.

Claim 11.169.1 $\nu(E) = \int_E f \, d\mu, \forall E \in \mathcal{B}$.

Proof of claim: Suppose that $\exists E_0 \in \mathcal{B}$ such that $\nu(E_0) > \int_{E_0} f \, d\mu \geq 0$. Then, $\exists \epsilon_0 > 0$ such that $\nu(E_0) > \epsilon_0 \mu(E_0) + \int_{E_0} f \, d\mu$. Define $\bar{\nu} : \mathcal{B} \rightarrow \mathbb{R}$ by $\bar{\nu}(E) = \nu(E) - \epsilon_0 \mu(E) - \int_E f \, d\mu$. By Propositions 11.116 and 11.136, $\bar{\nu}$ is a finite $\mathbb{R}$ -valued measure with $\bar{\nu}(E_0) \in (0, \infty) \subset \mathbb{R}$. By Lemma 11.160, $\exists A_0 \in \mathcal{B}$ with $A_0 \subseteq E_0$ such that A_0 is a positive set for $\bar{\nu}$ with $\bar{\nu}(A_0) \in (0, \infty) \subset \mathbb{R}$. Clearly, $0 < \bar{\nu}(A_0) \leq \nu(A_0)$. By $\nu \ll \mu$ and μ is finite, we have $\mu(A_0) \in (0, \infty) \subset \mathbb{R}$. $\forall E \in \mathcal{B}$, $\nu(E) = \nu(E \cap A_0) + \nu(E \setminus A_0) \geq \bar{\nu}(E \cap A_0) + \epsilon_0 \mu(E \cap A_0) + \int_{E \cap A_0} f \, d\mu + \int_{E \setminus A_0} f \, d\mu \geq \epsilon_0 \mu(E \cap A_0) + \int_E f \, d\mu$, where the last inequality follows from Proposition 11.83. Let $f_1 := \epsilon_0 \chi_{A_0, X}$. By Propositions 11.38, 11.39, and 7.23, $f + f_1$ is $\mathcal{B}$ -measurable. $\forall E \in \mathcal{B}$, $\int_E (f + f_1) \, d\mu = \int_E f \, d\mu + \int_E f_1 \, d\mu = \int_E f \, d\mu + \epsilon_0 \mu(E \cap A_0) \leq \nu(E)$, where the first equality follows from Proposition 11.83; and the second equality follows from Proposition 11.75. Hence, $f + f_1 \in \mathcal{F}$. Then, $\int_X (f + f_1) \, d\mu = \int_X f \, d\mu + \int_X f_1 \, d\mu = \lambda + \epsilon_0 \mu(A_0) > \lambda$. This contradicts with the definition of λ . Therefore, the claim holds. $\square$

Hence, f is a Radon-Nikodym derivative of ν with respect to μ . This special case is proved.

Next step, we consider the case when $\mathcal{X}$ is a σ -finite measure space and ν is a σ -finite measure on $(X, \mathcal{B})$. Then, $X = \bigcup_{n=1}^{\infty} X_n$ where $(X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$, $\nu(X_n) < +\infty$, and $\mu(X_n) < +\infty, \forall n \in \mathbb{N}$. Without loss of generality, we may assume that X_n 's are pairwise disjoint. $\forall n \in \mathbb{N}$, let $(X_n, \mathcal{B}_n, \mu_n)$ be the finite measure subspace of $\mathcal{X}$ and $(X_n, \mathcal{B}_n, \nu_n)$ be the finite measure subspace of $(X, \mathcal{B}, \nu)$ as defined in Proposition 11.13. By $\nu \ll \mu$, we have $\nu_n \ll \mu_n$. Then, by the first step, $\exists \bar{f}_n : X_n \rightarrow [0, \infty) \subset \mathbb{R}$ such that $\bar{f}_n$ is $\mathcal{B}_n$ -measurable and $\nu_n(E_n) = \int_{E_n} \bar{f}_n \, d\mu_n, \forall E_n \in \mathcal{B}_n$. Define $f : X \rightarrow [0, \infty) \subset \mathbb{R}$ by $f(x) = \bar{f}_n(x), \forall x \in X_n, \forall n \in \mathbb{N}$. By Proposition 11.41, f is $\mathcal{B}$ -measurable. $\forall E \in \mathcal{B}$, $\nu(E) = \sum_{n=1}^{\infty} \nu(E \cap X_n) = \sum_{n=1}^{\infty} \nu_n(E \cap X_n) = \sum_{n=1}^{\infty} \int_{E \cap X_n} \bar{f}_n \, d\mu_n = \sum_{n=1}^{\infty} \int_{E \cap X_n} f|_{X_n} \, d\mu_n = \sum_{n=1}^{\infty} \int_{E \cap X_n} f \, d\mu = \int_E f \, d\mu$, where the fifth and sixth equalities follow from Proposition 11.83. Hence, f is a Radon-Nikodym derivative of ν with respect to μ .

The third step, we consider the case when $\mathcal{X}$ is a σ -finite measure space and ν is an σ -finite $\mathbb{R}$ -valued measure on $(X, \mathcal{B})$. By Jordan Decomposition Theorem 11.163, there exists a pair of mutually singular σ -finite measures ν_+ and ν_- on $(X, \mathcal{B})$ such that $\nu = \nu_+ - \nu_-$ and $\mathcal{P} \circ \nu = \nu_+ + \nu_-$. By $\mathcal{P} \circ \nu \ll \mu$, we have $\nu_+ \ll \mu$ and $\nu_- \ll \mu$. By the second step, $\exists f_+ : X \rightarrow [0, \infty) \subset \mathbb{R}$ and $\exists f_- : X \rightarrow [0, \infty) \subset \mathbb{R}$ such that f_+ and f_- are Radon-Nikodym derivatives of ν_+ and ν_- with respect to μ , respectively. Let $f := f_+ - f_-$. $\forall E \in \text{dom}(\nu)$, we have $\mathcal{P} \circ \nu(E) = \nu_+(E) + \nu_-(E) = \int_E f_+ \, d\mu + \int_E f_- \, d\mu < \infty$. This leads to $\int_E f_+ \, d\mu < \infty$ and $\int_E f_- \, d\mu < \infty$. Then, $f_+|_E$ and $f_-|_E$ are (absolutely) integrable over $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$, which is the σ -finite measure subspace of $\mathcal{X}$. Then, $\nu(E) = \nu_+(E) - \nu_-(E) = \int_E f_+ \, d\mu - \int_E f_- \, d\mu = \int_E f \, d\mu \in \mathbb{R}$, where the third equality follows from

Proposition 11.92. By Proposition 11.116, f is a Radon-Nikodym derivative of a σ -finite $\mathbb{R}$ -valued measure $\bar{\nu}$ on $(X, \mathcal{B})$ with respect to μ . The above shows that $\bar{\nu}(E) = \nu(E) \in \mathbb{R}$, $\forall E \in \text{dom}(\nu)$. By Proposition 11.137, we have $\bar{\nu} = \nu$. Hence, f is a Radon-Nikodym derivative of ν with respect to μ .

This completes the three steps to show the existence of f . The uniqueness of f follows directly from Proposition 11.167. This completes the proof of the theorem. $\square$

Proposition 11.170 *Let $\mathcal{X} := (X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, and $f : X \rightarrow \mathcal{Y}$. Assume that $\mathcal{Y}^*$ is separable, $D_* \subseteq \mathcal{Y}^*$ is a countable dense set, and, $\forall y_* \in D_*$, $f_{y_*} : X \rightarrow \mathbb{K}$ defined by $f_{y_*}(x) = \langle y_*, f(x) \rangle$, $\forall x \in X$, is $\mathcal{B}$ -measurable. Then, f is $\mathcal{B}$ -measurable.*

Proof Let $D \subseteq \mathcal{Y}$ be a countable dense set. By Proposition 4.4, $\mathcal{Y}$ is second countable with a countable basis $\mathcal{E} := \{\mathcal{B}_{\mathcal{Y}}(y_0, r) \mid y_0 \in D, r \in \mathbb{Q}, r > 0\}$. By Proposition 11.34, all we need to show is that $\forall \mathcal{B}_{\mathcal{Y}}(y_0, r) \in \mathcal{E}$, $V := f_{\text{inv}}(\mathcal{B}_{\mathcal{Y}}(y_0, r)) \in \mathcal{B}$. $\exists N \in \mathbb{N}$ such that $r > 1/N$. We will show that $V = U := \bigcup_{n=N}^{\infty} \bigcap_{y_* \in D_*, y_* \neq \vartheta_{\mathcal{Y}^*}} f_{y_* \text{ inv}}(\mathcal{B}_{\mathbb{K}}(\langle y_*, y_0 \rangle, \|y_*\| (r - 1/n))) \in \mathcal{B}$. $\forall n \in \mathbb{N}$ with $n \geq N$, $\forall y_* \in D_*$ with $y_* \neq \vartheta_{\mathcal{Y}^*}$, f_{y_*} is $\mathcal{B}$ -measurable implies that $f_{y_* \text{ inv}}(\mathcal{B}_{\mathbb{K}}(\langle y_*, y_0 \rangle, \|y_*\| (r - 1/n))) \in \mathcal{B}$. Hence, $U \in \mathcal{B}$.

$\forall x \in V$, we have $\|f(x) - y_0\| < r$. Then, $\exists n \in \mathbb{N}$ with $n \geq N$ such that $\|f(x) - y_0\| < r - 1/n$. $\forall y_* \in D_*$ with $y_* \neq \vartheta_{\mathcal{Y}^*}$, we have $|f_{y_*}(x) - \langle y_*, y_0 \rangle| = |\langle y_*, f(x) - y_0 \rangle| \leq \|y_*\| \|f(x) - y_0\| < \|y_*\| (r - 1/n)$. Hence, $x \in f_{y_* \text{ inv}}(\mathcal{B}_{\mathbb{K}}(\langle y_*, y_0 \rangle, \|y_*\| (r - 1/n)))$. Then, $x \in U$. By the arbitrariness of x , we have $V \subseteq U$.

On the other hand, $\forall x \in U$, $\exists n \in \mathbb{N}$ with $n \geq N$, $\forall y_* \in D_*$ with $y_* \neq \vartheta_{\mathcal{Y}^*}$, we have $|f_{y_*}(x) - \langle y_*, y_0 \rangle| < \|y_*\| (r - 1/n)$. Then, $|\langle y_*, f(x) - y_0 \rangle| < \|y_*\| (r - 1/n)$. If $f(x) = y_0$, then $x \in V$. If $f(x) \neq y_0$, by Proposition 7.85, $\exists y_{*0} \in \mathcal{Y}^*$ with $\|y_{*0}\| = 1$ such that $\|f(x) - y_0\| = \langle y_{*0}, f(x) - y_0 \rangle$. Since D_* is dense in $\mathcal{Y}^*$, then, by Proposition 4.13, $\exists (y_{*i})_{i=1}^{\infty} \subseteq D_*$ that converges to y_{*0} . Without loss of generality, assume $y_{*i} \neq \vartheta_{\mathcal{Y}^*}$, $\forall i \in \mathbb{N}$. By Propositions 3.66, 3.67, 7.21, and 7.72, we have $\|f(x) - y_0\| = \lim_{i \in \mathbb{N}} |\langle y_{*i}, f(x) - y_0 \rangle| \leq \lim_{i \in \mathbb{N}} \|y_{*i}\| (r - 1/n) = r - 1/n < r$. Hence, $x \in V$. This implies that $U \subseteq V$.

Hence, $V = U \in \mathcal{B}$. This completes the proof of the proposition. $\square$

Theorem 11.171 (Radon-Nikodym) *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite $\mathbb{K}$ -valued measure space, $\mathcal{Y}$ be a separable reflexive Banach space over $\mathbb{K}$, ν be a σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$. Assume that $\mathcal{P} \circ \nu \ll \mathcal{P} \circ \mu$ and $\mathcal{Y}^*$ is separable. Then, there exists $f : X \rightarrow \mathcal{Y}$ such that f is the Radon-Nikodym derivative of ν with respect to μ . The function f is unique in the sense that $g : X \rightarrow \mathcal{Y}$ is another function with this property if, and only if, g is $\mathcal{B}$ -measurable and $f = g$ a.e. in $\mathcal{X}$. Hence, $\frac{d\nu}{d\mu} = f$ a.e. in $\mathcal{X}$.*

Proof First, we consider the special case when $\mathcal{X}$ is a σ -finite measure space and $\mathcal{Y} = \mathbb{C}$. By Proposition 11.165, $\nu = \nu_r + i\nu_i$, where ν_r and ν_i are σ -finite $\mathbb{R}$ -valued measures on $(X, \mathcal{B})$, and $\max\{\mathcal{P} \circ \nu_r(E), \mathcal{P} \circ \nu_i(E)\} \leq \mathcal{P} \circ \nu(E) \leq \mathcal{P} \circ \nu_r(E) + \mathcal{P} \circ \nu_i(E)$, $\forall E \in \mathcal{B}$. Then, $\mathcal{P} \circ \nu \ll \mu$ implies that $\mathcal{P} \circ \nu_r \ll \mu$ and $\mathcal{P} \circ \nu_i \ll \mu$. By Radon-Nikodym Theorem 11.169, there exist $f_r : X \rightarrow \mathbb{R}$ and $f_i : X \rightarrow \mathbb{R}$ such that $\frac{d\nu_r}{d\mu} = f_r$ a.e. in $\mathcal{X}$ and $\frac{d\nu_i}{d\mu} = f_i$ a.e. in $\mathcal{X}$. By Proposition 11.168, we have $\frac{d\nu}{d\mu} = \frac{d(\nu_r + i\nu_i)}{d\mu} = \frac{d\nu_r}{d\mu} + i\frac{d\nu_i}{d\mu} = f_r + if_i =: f : X \rightarrow \mathbb{C}$ a.e. in $\mathcal{X}$. This completes the proof of the special case.

Next, we consider the special case when $\mathcal{X}$ is a σ -finite measure space. $\forall y_* \in \mathcal{Y}^*$, define ν_{y_*} to be the σ -finite $\mathbb{K}$ -valued measure on $(X, \mathcal{B})$ such that $\nu_{y_*}(E) = \langle\langle y_*, \nu(E) \rangle\rangle$, $\forall E \in \text{dom}(\nu)$, as shown in Proposition 11.138. Then, by Proposition 11.138, $\mathcal{P} \circ \nu_{y_*} \leq \|y_*\| \mathcal{P} \circ \nu$. Then, $\mathcal{P} \circ \nu_{y_*} \ll \mu$ since $\mathcal{P} \circ \nu \ll \mu$. By Radon-Nikodym Theorem 11.169 and the special case, $\exists! f_{y_*} : X \rightarrow \mathbb{K}$ such that $\frac{d\nu_{y_*}}{d\mu} = f_{y_*}$ a.e. in $\mathcal{X}$. $\forall y_{*1}, y_{*2} \in \mathcal{Y}^*$, $\forall \alpha, \beta \in \mathbb{K}$, $\forall E \in \text{dom}(\nu)$, we have $E \in \text{dom}(\nu_{y_{*1}}) \cap \text{dom}(\nu_{y_{*2}}) \cap \text{dom}(\nu_{\alpha y_{*1} + \beta y_{*2}})$. Then, $\mathbb{K} \ni \nu_{\alpha y_{*1} + \beta y_{*2}}(E) = \int_E f_{\alpha y_{*1} + \beta y_{*2}} d\mu = \langle\langle \alpha y_{*1} + \beta y_{*2}, \nu(E) \rangle\rangle = \alpha \langle\langle y_{*1}, \nu(E) \rangle\rangle + \beta \langle\langle y_{*2}, \nu(E) \rangle\rangle = \alpha \nu_{y_{*1}}(E) + \beta \nu_{y_{*2}}(E) = \alpha \int_E f_{y_{*1}} d\mu + \beta \int_E f_{y_{*2}} d\mu$, $\int_E \mathcal{P} \circ f_{\alpha y_{*1} + \beta y_{*2}} d\mu < \infty$, $\int_E \mathcal{P} \circ f_{y_{*1}} d\mu < \infty$, and $\int_E \mathcal{P} \circ f_{y_{*2}} d\mu < \infty$. Then, $f_{y_{*1}}|_E$ and $f_{y_{*2}}|_E$ are absolutely integrable over $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$, which is the σ -finite measure subspace of $\mathcal{X}$. By Proposition 11.92, we have $\nu_{\alpha y_{*1} + \beta y_{*2}}(E) = \int_E f_{\alpha y_{*1} + \beta y_{*2}} d\mu = \int_E (\alpha f_{y_{*1}} + \beta f_{y_{*2}}) d\mu \in \mathbb{K}$, $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \nu(E) < \infty$. By Propositions 11.116 and 11.137, $\alpha f_{y_{*1}} + \beta f_{y_{*2}}$ is a Radon-Nikodym derivative of $\nu_{\alpha y_{*1} + \beta y_{*2}}$ with respect to μ . Then, by Proposition 11.167,

$$f_{\alpha y_{*1} + \beta y_{*2}} = \alpha f_{y_{*1}} + \beta f_{y_{*2}} \text{ a.e. in } \mathcal{X}; \quad \forall y_{*1}, y_{*2} \in \mathcal{Y}^*, \forall \alpha, \beta \in \mathbb{K}$$

Note that $\mathcal{P} \circ \nu$ is a σ -finite measure on $(X, \mathcal{B})$ and $\mathcal{P} \circ \nu \ll \mu$. Again, by Radon-Nikodym Theorem 11.169, $\exists! \hat{f} : X \rightarrow [0, \infty) \subset \mathbb{R}$ such that $\frac{d\mathcal{P} \circ \nu}{d\mu} = \hat{f}$ a.e. in $\mathcal{X}$. This yields $\mathcal{P} \circ \nu(E) = \int_E \hat{f} d\mu$, $\forall E \in \mathcal{B}$. $\forall y_* \in \mathcal{Y}^*$, by Definition 11.166, $\mathcal{P} \circ \nu_{y_*}(E) = \int_E (\mathcal{P} \circ f_{y_*}) d\mu \leq (\|y_*\| \mathcal{P} \circ \nu)(E)$, $\forall E \in \mathcal{B}$. When $\|y_*\| > 0$, $\mathcal{P} \circ \nu_{y_*}(E) \leq (\|y_*\| \mathcal{P} \circ \nu)(E) = \|y_*\| \mathcal{P} \circ \nu(E) = \|y_*\| \int_E \hat{f} d\mu = \int_E (\|y_*\| \hat{f}) d\mu$, where the first equality follows from Proposition 11.136; and the third equality follows from Proposition 11.83. When $\|y_*\| = 0$, $\mathcal{P} \circ \nu_{y_*}(E) \leq 0 = \int_E (\|y_*\| \hat{f}) d\mu$, where the inequality follows from Proposition 11.138; and the equality follows from Proposition 11.75. Hence, $\mathcal{P} \circ \nu_{y_*}(E) = \int_E (\mathcal{P} \circ f_{y_*}) d\mu \leq \int_E (\|y_*\| \hat{f}) d\mu$, $\forall E \in \mathcal{B}$, $\forall y_* \in \mathcal{Y}^*$. Then, by Propositions 11.96,

$$\mathcal{P} \circ f_{y_*} \leq \|y_*\| \hat{f} \text{ a.e. in } \mathcal{X}; \quad \forall y_* \in \mathcal{Y}^*$$

Let $\mathbb{K}_{\mathbb{Q}} := \mathbb{Q}$ if $\mathbb{K} = \mathbb{R}$ and $\mathbb{K}_{\mathbb{Q}} := \{\alpha + i\beta \in \mathbb{C} \mid \alpha, \beta \in \mathbb{Q}\}$ if $\mathbb{K} = \mathbb{C}$. Clearly, $\mathbb{K}_{\mathbb{Q}}$ is a countable dense set in $\mathbb{K}$.

Let $D \subseteq \mathcal{Y}^*$ be a countable dense set and $\hat{D} := \{ \sum_{i=1}^n \alpha_i y_{*i} \in \mathcal{Y}^* \mid n \in \mathbb{Z}_+, \alpha_1, \dots, \alpha_n \in \mathbb{K}_{\mathbb{Q}}, y_{*1}, \dots, y_{*n} \in D \}$. Then, $\hat{D} \subseteq \mathcal{Y}^*$ is also a countable dense set. Note that, $\forall y_{*1}, y_{*2} \in \hat{D}, \forall \alpha, \beta \in \mathbb{K}_{\mathbb{Q}}$, we have $\alpha y_{*1} + \beta y_{*2} \in \hat{D}$. Define

$$E_0 := \{ x \in X \mid \exists y_{*1}, y_{*2} \in \hat{D}, \exists \alpha, \beta \in \mathbb{K}_{\mathbb{Q}} \ni \cdot f_{\alpha y_{*1} + \beta y_{*2}}(x) \neq \alpha f_{y_{*1}}(x) + \beta f_{y_{*2}}(x) \text{ or } |f_{y_{*1}}(x)| > \|y_{*1}\| \hat{f}(x) \}$$

Then, $E_0 \in \mathcal{B}$ and $\mu(E_0) = 0$.

We have the following results.

$$\begin{aligned} f_{\alpha y_{*1} + \beta y_{*2}}(x) &= \alpha f_{y_{*1}}(x) + \beta f_{y_{*2}}(x); \\ \forall x \in X \setminus E_0, \forall y_{*1}, y_{*2} \in \hat{D}, \forall \alpha, \beta \in \mathbb{K}_{\mathbb{Q}} \end{aligned} \quad (11.4)$$

and

$$|f_{y_*}(x)| \leq \|y_*\| \hat{f}(x); \quad \forall x \in X \setminus E_0, \forall y_* \in \hat{D} \quad (11.5)$$

$\forall x \in X \setminus E_0, \forall y_* \in \mathcal{Y}^*$, by Proposition 4.13, $\exists (y_{*i})_{i=1}^{\infty} \subseteq \hat{D}$ that converges to y_* . Then, this sequence is a Cauchy sequence. By (11.4) and (11.5), $(f_{y_{*i}}(x))_{i=1}^{\infty} \subseteq \mathbb{K}$ is a Cauchy sequence and admits limit $F(x, y_*) \in \mathbb{K}$. It should be clear that $F(x, y_*)$ is independent of the choice of the sequence $(y_{*i})_{i=1}^{\infty}$, since let $(\bar{y}_{*i})_{i=1}^{\infty} \subseteq \hat{D}$ be any sequence converging to y_* , the combined sequence $(y_{*1}, \bar{y}_{*1}, y_{*2}, \bar{y}_{*2}, \dots) \subseteq \hat{D}$ converges to y_* . Therefore, the sequence $(f_{y_{*1}}(x), f_{\bar{y}_{*1}}(x), f_{y_{*2}}(x), f_{\bar{y}_{*2}}(x), \dots) \subseteq \mathbb{K}$ is a Cauchy sequence and admits a limit, which has to equal to $F(x, y_*)$. Then, $\forall x \in X \setminus E_0, \forall y_* \in \mathcal{Y}^*, F(x, y_*) \in \mathbb{K}$. $\forall x \in X \setminus E_0, \forall \bar{y}_* \in \hat{D}, F(x, \bar{y}_*) = f_{\bar{y}_*}(x)$.

$\forall x \in X \setminus E_0, \forall y_{*1}, y_{*2} \in \mathcal{Y}^*, \forall \alpha, \beta \in \mathbb{K}, \exists (\bar{y}_{*1,i})_{i=1}^{\infty} \subseteq \hat{D}$ that converges to $y_{*1}, \exists (\bar{y}_{*2,i})_{i=1}^{\infty} \subseteq \hat{D}$ that converges to $y_{*2}, \exists (\alpha_i)_{i=1}^{\infty} \subseteq \mathbb{K}_{\mathbb{Q}}$ that converges to α , and $\exists (\beta_i)_{i=1}^{\infty} \subseteq \mathbb{K}_{\mathbb{Q}}$ that converges to β . By Propositions 7.23, 3.66, and 3.67, we have $\lim_{i \in \mathbb{N}} (\alpha_i \bar{y}_{*1,i} + \beta_i \bar{y}_{*2,i}) = \alpha y_{*1} + \beta y_{*2}$, where the sequence on the left-hand-side is contained in $\hat{D}$. This implies that $F(x, \alpha y_{*1} + \beta y_{*2}) = \lim_{i \in \mathbb{N}} f_{\alpha_i \bar{y}_{*1,i} + \beta_i \bar{y}_{*2,i}}(x) = \lim_{i \in \mathbb{N}} (\alpha_i f_{\bar{y}_{*1,i}}(x) + \beta_i f_{\bar{y}_{*2,i}}(x)) = \alpha F(x, y_{*1}) + \beta F(x, y_{*2})$, where the second equality follows from (11.4); and the last equality follows from Propositions 7.23, 3.66, and 3.67. Furthermore, $|F(x, y_{*1})| = \lim_{i \in \mathbb{N}} |f_{\bar{y}_{*1,i}}(x)| \leq \lim_{i \in \mathbb{N}} \|\bar{y}_{*1,i}\| \hat{f}(x) = \|y_{*1}\| \hat{f}(x)$, where the inequality follows from (11.5); and the last equality follows from Propositions 3.66 and 7.21. Hence, $\forall x \in X \setminus E_0, F(x, \cdot) : \mathcal{Y}^* \rightarrow \mathbb{K}$ is a bounded linear functional on $\mathcal{Y}^*$. Since $\mathcal{Y}$ is reflexive, then $\exists f(x) \in \mathcal{Y}$ such that $F(x, y_*) = \langle\langle y_*, f(x) \rangle\rangle, \forall y_* \in \mathcal{Y}^*$ and $\|f(x)\| \leq \hat{f}(x)$. Therefore, we may define $f : X \rightarrow \mathcal{Y}$ by assigning $f(x) = \vartheta_{\mathcal{Y}}, \forall x \in E_0$.

$\forall y_* \in \mathcal{Y}^*$, define $F_{y_*} : X \rightarrow \mathbb{K}$ by $F_{y_*}(x) = \langle\langle y_*, f(x) \rangle\rangle, \forall x \in X$. Fix any $\bar{y}_* \in \hat{D}$. Then, $\forall x \in X \setminus E_0, F_{\bar{y}_*}(x) = \langle\langle \bar{y}_*, f(x) \rangle\rangle = F(x, \bar{y}_*) = f_{\bar{y}_*}(x)$ and, $\forall x \in E_0, F_{\bar{y}_*}(x) = \langle\langle \bar{y}_*, f(x) \rangle\rangle = 0$. By Proposition 11.41, $F_{\bar{y}_*}$ is $\mathcal{B}$ -measurable. By the arbitrariness of $\bar{y}_*$ and Proposition 11.170, f is $\mathcal{B}$ -measurable.

Claim 11.171.1 $\forall y_{*0} \in \mathcal{Y}^*, f_{y_{*0}}(x) = \langle \langle y_{*0}, f(x) \rangle \rangle$ a.e. $x \in \mathcal{X}$.

Proof of claim: Let $\bar{D} := D \cup \{y_{*0}\}$ and $\hat{\bar{D}} := \{ \sum_{i=1}^n \alpha_i y_{*i} \in \mathcal{Y}^* \mid n \in \mathbb{Z}_+, \alpha_1, \dots, \alpha_n \in \mathbb{K}_{\mathbb{Q}}, y_{*1}, \dots, y_{*n} \in \bar{D} \}$. Then, $\hat{\bar{D}} \subseteq \mathcal{Y}^*$ is also a countable dense set and $\hat{D} \subseteq \hat{\bar{D}}$. Define

$$E_1 := \{ x \in X \mid \exists y_{*1}, y_{*2} \in \hat{\bar{D}}, \exists \alpha, \beta \in \mathbb{K}_{\mathbb{Q}} \ni \cdot f_{\alpha y_{*1} + \beta y_{*2}}(x) \neq \alpha f_{y_{*1}}(x) + \beta f_{y_{*2}}(x) \text{ or } |f_{y_{*1}}(x)| > \|y_{*1}\| \hat{f}(x) \}$$

Then, $E_1 \in \mathcal{B}$, $\mu(E_1) = 0$ and $E_0 \subseteq E_1$.

$\forall x \in X \setminus E_1, \forall y_{*1}, y_{*2} \in \hat{\bar{D}}, \forall \alpha, \beta \in \mathbb{K}_{\mathbb{Q}}$, we have $f_{\alpha y_{*1} + \beta y_{*2}}(x) = \alpha f_{y_{*1}}(x) + \beta f_{y_{*2}}(x)$ and $|f_{y_{*1}}(x)| \leq \|y_{*1}\| \hat{f}(x)$. By Proposition 4.13, $\exists (y_{*i})_{i=1}^{\infty} \subseteq \hat{\bar{D}}$ that converges to y_{*0} . Then, this sequence is a Cauchy sequence. By the above, $(f_{y_{*i}}(x))_{i=1}^{\infty} \subseteq \mathbb{K}$ is a Cauchy sequence and admits limit $\bar{F}(x, y_{*0}) \in \mathbb{K}$. It should be clear that $\bar{F}(x, y_{*0})$ is independent of the choice of the sequence $(y_{*i})_{i=1}^{\infty}$. Two particular sequences $(\bar{y}_{*i})_{i=1}^{\infty} \subseteq \hat{D}$ converging to y_{*0} and $(y_{*0}, y_{*0}, \dots) \subseteq \hat{D}$ leads to $f_{y_{*0}}(x) = \bar{F}(x, y_{*0}) = F(x, y_{*0})$. Therefore, $f_{y_{*0}}(x) = F(x, y_{*0}) = \langle \langle y_{*0}, f(x) \rangle \rangle = F_{y_{*0}}(x)$, $\forall x \in X \setminus E_1$. Note that $f_{y_{*0}}$ is $\mathcal{B}$ -measurable and, by Propositions 7.72 and 11.38, $F_{y_{*0}}$ is $\mathcal{B}$ -measurable. Then, $E_1 \supseteq \{ x \in X \mid f_{y_{*0}}(x) \neq F_{y_{*0}}(x) \} = \{ x \in X \mid |f_{y_{*0}}(x) - F_{y_{*0}}(x)| > 0 \} \in \mathcal{B}$. Hence, $f_{y_{*0}} = F_{y_{*0}}$ a.e. in $\mathcal{X}$. Therefore, the claim holds. $\square$

Note that $\mathcal{P} \circ f(x) \leq \hat{f}(x)$, $\forall x \in X$, and $\mathcal{P} \circ f$ is $\mathcal{B}$ -measurable, by Propositions 7.21 and 11.38. $\forall E \in \text{dom}(\nu)$, we have $\mathcal{P} \circ \nu(E) = \int_E \hat{f} d\mu < \infty$. This implies that $\int_E \mathcal{P} \circ f d\mu < \infty$. By Proposition 11.92, we have $\int_E f d\mu \in \mathcal{Y}$. $\forall y_* \in \mathcal{Y}^*, \mathcal{P} \circ \nu_{y_*}(E) \leq (\|y_*\| \mathcal{P} \circ \nu)(E) < \infty$, $E \in \text{dom}(\nu_{y_*})$, and $\langle \langle y_*, \nu(E) \rangle \rangle = \nu_{y_*}(E) = \int_E f_{y_*} d\mu = \int_E F_{y_*} d\mu = \int_E \langle \langle y_*, f(x) \rangle \rangle d\mu(x) = \langle \langle y_*, \int_E f d\mu \rangle \rangle$, where the last three equalities follow from Proposition 11.92. Hence, by Proposition 7.85, $\nu(E) = \int_E f d\mu \in \mathcal{Y}, \forall E \in \text{dom}(\nu)$. Hence, by Propositions 11.116 and 11.137, f is a Radon-Nikodym derivative of ν with respect to μ . By Proposition 11.167, f is unique as desired and $\frac{d\nu}{d\mu} = f$ a.e. in $\mathcal{X}$.

Finally, we consider the general case. Let $\bar{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$, which is a σ -finite measure space. By the second special case, $\exists! f_1 : X \rightarrow \mathcal{Y}$ such that $\frac{d\nu}{d\mathcal{P} \circ \mu} = f_1$ a.e. in $\bar{\mathcal{X}}$. Again by the second special case, $\exists! f_2 : X \rightarrow \mathbb{K}$ such that $\frac{d\mu}{d\mathcal{P} \circ \mu} = f_2$ a.e. in $\bar{\mathcal{X}}$. By Definition 11.166, $\frac{d\mathcal{P} \circ \mu}{d\mathcal{P} \circ \mu} = \mathcal{P} \circ f_2$ a.e. in $\bar{\mathcal{X}}$. Clearly, the constant function 1 is the Radon-Nikodym derivative of $\mathcal{P} \circ \mu$ with respect to $\mathcal{P} \circ \mu$. Then, $\mathcal{P} \circ f_2 = 1$ a.e. in $\bar{\mathcal{X}}$ and $\exists f_3 : X \rightarrow \mathbb{K}$ such that $f_3 f_2 = 1$ a.e. in $\bar{\mathcal{X}}$. By Proposition 11.168, $\frac{d\mathcal{P} \circ \mu}{d\mu} = f_3$ a.e. in $\mathcal{X}$ and $\frac{d\nu}{d\mu} = \left(\frac{d\nu}{d\mathcal{P} \circ \mu} \right) \left(\frac{d\mathcal{P} \circ \mu}{d\mu} \right) = f_1 f_3 =: f : X \rightarrow \mathcal{Y}$ a.e. in $\mathcal{X}$.

This completes the proof of the theorem. $\square$

Theorem 11.172 (Lebesgue Decomposition) *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a Banach space, and ν be a σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$. Then, there exists a unique pair of σ -finite $\mathcal{Y}$ -valued measures ν_0 and ν_1 on $(X, \mathcal{B})$ such that $\mathcal{P} \circ \nu_0 \perp \mu$, $\mathcal{P} \circ \nu_1 \ll \mu$, $\nu = \nu_0 + \nu_1$, and $\mathcal{P} \circ \nu = \mathcal{P} \circ \nu_0 + \mathcal{P} \circ \nu_1$. Furthermore, the following statements hold.*

- (i) *if ν is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$, then ν_0 and ν_1 are finite $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$.*
- (ii) *if ν is a σ -finite measure on $(X, \mathcal{B})$, then ν_0 and ν_1 are σ -finite measures on $(X, \mathcal{B})$.*
- (iii) *if ν is a finite measure on $(X, \mathcal{B})$, then ν_0 and ν_1 are finite measures on $(X, \mathcal{B})$.*

Proof Let $\lambda := \mu + \mathcal{P} \circ \nu$. By Proposition 11.136, λ is a σ -finite measure on $(X, \mathcal{B})$. Let $\bar{\mathcal{X}} := (X, \mathcal{B}, \lambda)$. Clearly, we have $\mu \ll \lambda$ and $\mathcal{P} \circ \nu \ll \lambda$. By Radon-Nikodym Theorem 11.169, $\exists f : X \rightarrow [0, \infty) \subset \mathbb{R}$ such that f is the Radon-Nikodym derivative of μ with respect to λ . Then, $\mu(E) = \int_E f d\lambda$, $\forall E \in \mathcal{B}$. Let $A := \{x \in X \mid f(x) > 0\} \in \mathcal{B}$ and $B := \{x \in X \mid f(x) = 0\} \in \mathcal{B}$. Then, $A \cap B = \emptyset$ and $A \cup B = X$. By Proposition 11.75, $\mu(B) = \int_B f d\lambda = 0$. Define ν_0 to be a function from $\mathcal{B}$ to $\mathcal{Y}$ by: $\nu_0(E)$ is undefined, $\forall E \in \mathcal{B}$ with $E \cap B \in \mathcal{B} \setminus \text{dom}(\nu)$; and $\nu_0(E) = \nu(E \cap B) \in \mathcal{Y}$, $\forall E \in \mathcal{B}$ with $E \cap B \in \text{dom}(\nu)$. Define ν_1 to be a function from $\mathcal{B}$ to $\mathcal{Y}$ by: $\nu_1(E)$ is undefined, $\forall E \in \mathcal{B}$ with $E \cap A \in \mathcal{B} \setminus \text{dom}(\nu)$; and $\nu_1(E) = \nu(E \cap A) \in \mathcal{Y}$, $\forall E \in \mathcal{B}$ with $E \cap A \in \text{dom}(\nu)$. By Proposition 11.157, ν_0 and ν_1 are σ -finite $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$ such that $\nu = \nu_0 + \nu_1$, $\mathcal{P} \circ \nu = \mathcal{P} \circ \nu_0 + \mathcal{P} \circ \nu_1$, $\mathcal{P} \circ \nu_0(E) = \mathcal{P} \circ \nu(E \cap B)$, and $\mathcal{P} \circ \nu_1(E) = \mathcal{P} \circ \nu(E \cap A)$, $\forall E \in \mathcal{B}$. This implies that $\mathcal{P} \circ \nu_0(A) = \mathcal{P} \circ \nu(A \cap B) = 0$. Then, $\mathcal{P} \circ \nu_0 \perp \mu$.

$\forall E \in \mathcal{B}$ with $\mu(E) = 0$, we have $\int_E f d\lambda = 0$. Let $\bar{\mathcal{E}} := (E, \mathcal{B}_E, \lambda_E)$ be the σ -finite measure subspace of $\bar{\mathcal{X}}$ as defined in Proposition 11.13. Then, by Proposition 11.96, we have $f = 0$ a.e. in $\bar{\mathcal{E}}$. Then, $\lambda(E \cap A) = 0$ and $\mathcal{P} \circ \nu_1(E) = \mathcal{P} \circ \nu(E \cap A) = 0$, since $\mathcal{P} \circ \nu \ll \lambda$. Hence, $\mathcal{P} \circ \nu_1 \ll \mu$. This shows that ν_0 and ν_1 is the pair of σ -finite $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$ that we seek.

Next, we show the uniqueness of the pair ν_0 and ν_1 . Let $\hat{\nu}_0$ and $\hat{\nu}_1$ be another pair of σ -finite $\mathcal{Y}$ -valued measures on $(X, \mathcal{B})$ such that $\mathcal{P} \circ \hat{\nu}_0 \perp \mu$, $\mathcal{P} \circ \hat{\nu}_1 \ll \mu$, $\nu = \hat{\nu}_0 + \hat{\nu}_1$, and $\mathcal{P} \circ \nu = \mathcal{P} \circ \hat{\nu}_0 + \mathcal{P} \circ \hat{\nu}_1$. Then, $\exists \hat{A}, \hat{B} \in \mathcal{B}$ with $\hat{A} = X \setminus \hat{B}$ such that $\mathcal{P} \circ \hat{\nu}_0(\hat{A}) = \mu(\hat{B}) = 0$. $\forall E \in \mathcal{B}$ with $\lambda(E) < \infty$, we have $\mu(E) < \infty$, $\mathcal{P} \circ \nu(E) = \mathcal{P} \circ \nu_0(E) + \mathcal{P} \circ \nu_1(E) = \mathcal{P} \circ \hat{\nu}_0(E) + \mathcal{P} \circ \hat{\nu}_1(E) < \infty$. Then, $\hat{\nu}_0(E) = \hat{\nu}_0(E \cap \hat{A}) + \hat{\nu}_0(E \cap \hat{B}) = \hat{\nu}_0(E \cap \hat{B}) = \hat{\nu}_0(E \cap \hat{B}) + \hat{\nu}_1(E \cap \hat{B}) = \nu(E \cap \hat{B}) = \nu_0(E \cap \hat{B}) + \nu_1(E \cap \hat{B}) + \hat{\nu}_0(E \cap \hat{A} \cap B) + \hat{\nu}_1(E \cap \hat{A} \cap B) = \nu_0(E \cap \hat{B}) + \nu(E \cap \hat{A} \cap B) = \nu_0(E \cap \hat{B}) + \nu_0(E \cap \hat{A}) = \nu_0(E) \in \mathcal{Y}$, where the second equality follows from the fact that $\mathcal{P} \circ \hat{\nu}_0(\hat{A}) = 0$; the third equality follows from $\mu(\hat{B}) = 0$ and $\mathcal{P} \circ \hat{\nu}_1 \ll \mu$; the fifth equality follows from the fact that $\mathcal{P} \circ \hat{\nu}_0(\hat{A}) = 0$, $\mu(B) = 0$ and $\mathcal{P} \circ \hat{\nu}_1 \ll \mu$; and the sixth equality

follows from the fact that $\mu(\hat{B}) = 0$ and $\mathcal{P} \circ \nu_1 \ll \mu$. We also have $\hat{\nu}_1(E) = \hat{\nu}_1(E \cap \hat{A}) + \hat{\nu}_1(E \cap \hat{B}) = \hat{\nu}_1(E \cap \hat{A}) = \hat{\nu}_1(E \cap \hat{A}) + \hat{\nu}_0(E \cap \hat{A}) = \nu(E \cap \hat{A}) = \nu_0(E \cap \hat{A}) + \nu_1(E \cap \hat{A}) = \nu(E \cap \hat{A} \cap B) + \nu_1(E \cap \hat{A}) = \hat{\nu}_0(E \cap \hat{A} \cap B) + \hat{\nu}_1(E \cap \hat{A} \cap B) + \nu_1(E \cap \hat{A}) = \nu_1(E \cap \hat{A}) = \nu_1(E \cap \hat{A}) + \nu_1(E \cap \hat{B}) = \nu_1(E) \in \mathcal{Y}$ where the second equality follows from the fact that $\mu(\hat{B}) = 0$ and $\mathcal{P} \circ \hat{\nu}_1 \ll \mu$; the third equality follows from $\mathcal{P} \circ \hat{\nu}_0(\hat{A}) = 0$; the eighth equality follows from the fact that $\mathcal{P} \circ \hat{\nu}_0(\hat{A}) = 0$, $\mu(B) = 0$, and $\mathcal{P} \circ \hat{\nu}_1 \ll \mu$; and the ninth equality follows from the fact that $\mu(\hat{B}) = 0$ and $\mathcal{P} \circ \nu_1 \ll \mu$. By Proposition 11.137, we have $\hat{\nu}_0 = \nu_0$ and $\hat{\nu}_1 = \nu_1$. Hence, the pair ν_0 and ν_1 is unique.

(i) If ν is finite, then $\mathcal{P} \circ \nu(X) = \mathcal{P} \circ \nu_0(X) + \mathcal{P} \circ \nu_1(X) < \infty$. Then, ν_0 and ν_1 are finite.

(ii) If ν is a σ -finite measure on $(X, \mathcal{B})$, by Proposition 11.135, it is identified with a σ -finite $\mathbb{R}$ -valued measure $\bar{\nu}$ on $(X, \mathcal{B})$ such that $\bar{\nu}(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\bar{\nu})$. By the general case, there exists a pair of σ -finite $\mathbb{R}$ -valued measures $\bar{\nu}_0$ and $\bar{\nu}_1$ on $(X, \mathcal{B})$ that satisfies the desired properties. By the proof of the general case, $\bar{\nu}_0(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\bar{\nu}_0)$, and $\bar{\nu}_1(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\bar{\nu}_1)$. By Proposition 11.135, $\bar{\nu}_0$ and $\bar{\nu}_1$ are identified with σ -finite measures ν_0 and ν_1 , respectively. Then, $\nu_0 = \mathcal{P} \circ \bar{\nu}_0 \perp \mu$, $\nu_1 = \mathcal{P} \circ \bar{\nu}_1 \ll \mu$, and $\nu = \mathcal{P} \circ \bar{\nu} = \mathcal{P} \circ \bar{\nu}_0 + \mathcal{P} \circ \bar{\nu}_1 = \nu_0 + \nu_1$. Then, ν_0 and ν_1 are σ -finite measures on $(X, \mathcal{B})$.

(iii) If ν is a finite measure on $(X, \mathcal{B})$, then, by (i) and (ii), ν_0 and ν_1 are finite measures on $(X, \mathcal{B})$.

This completes the proof of the theorem. $\square$

11.9 L_p Spaces

Example 11.173 Let $p \in [1, \infty) \subset \mathbb{R}$, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, and $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})$ be the vector space of functions of $\mathcal{X}$ to $\mathcal{Y}$ as defined in Example 6.20 with the usual vector addition, scalar multiplication, and the null vector ϑ . Define $l : [0, \infty) \subset \mathbb{R} \rightarrow [0, \infty) \subset \mathbb{R}$ by $l(t) = t^p$, $\forall t \in [0, \infty) \subset \mathbb{R}$. We will introduce the notation $\mathcal{P}_p \circ f$ to denote $l \circ \mathcal{P} \circ f$. Let $Z_p := \{f \in \mathcal{M}(\mathcal{X}, \mathcal{Y}) \mid f \text{ is } \mathcal{B}\text{-measurable and } \mathcal{P}_p \circ f \text{ is integrable over } \mathcal{X}\}$. Define $\|\cdot\|_p : Z_p \rightarrow [0, \infty) \subset \mathbb{R}$ by $\|f\|_p = (\int_X (\mathcal{P}_p \circ f) d\mu)^{1/p}$, $\forall f \in Z_p$. We will next show that Z_p is a subspace of $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})$ and $\|\cdot\|_p$ defines a pseudo-norm on Z_p .

Clearly, $\vartheta \in Z_p \neq \emptyset$. $\forall f, g \in Z_p$, $\forall \alpha \in \mathbb{K}$, by Propositions 7.23, 11.38, and 11.39, $f + g$ and αf are $\mathcal{B}$ -measurable. By Minkowski's Inequality 11.174, $f + g \in Z_p$ and $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. $\|\alpha f\|_p = (\int_X (\mathcal{P}_p \circ (\alpha f)) d\mu)^{1/p} = (|\alpha|^p \int_X (\mathcal{P}_p \circ f) d\mu)^{1/p} = |\alpha| (\int_X (\mathcal{P}_p \circ f) d\mu)^{1/p} = |\alpha| \|f\|_p \in \mathbb{R}$, where the second equality follows from Proposition 11.92. Hence, $\alpha f \in Z_p$.

The above shows that Z_p is a subspace of $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})$. Hence, $(Z_p, \mathbb{K})$ is a vector space. Clearly, $\|\vartheta\|_p = 0$. Then, combined with the above, we have $\|\cdot\|_p$ defines a pseudo-norm on $(Z_p, \mathbb{K})$. By Proposition 7.47, the quotient space of $(Z_p, \mathbb{K})$ modulo $\|\cdot\|_p$ is a normed linear space, to be denoted $L_p(\mathcal{X}, \mathcal{Y})$. We will denote the vector space $(Z_p, \mathbb{K})$ with the pseudo-norm $\|\cdot\|_p$ by $\bar{L}_p(\mathcal{X}, \mathcal{Y})$.

$\forall f \in Z_p$ with $\|f\|_p = 0$, we have $\int_X \mathcal{P}_p \circ f \, d\mu = 0$. By Proposition 11.96, we have $\mathcal{P}_p \circ f = 0$ a.e. in $\mathcal{X}$ and $f = \vartheta_{\mathcal{Y}}$ a.e. in $\mathcal{X}$. On the other hand, $\forall f \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ with $f = \vartheta_{\mathcal{Y}}$ a.e. in $\mathcal{X}$ and f being $\mathcal{B}$ -measurable, then, by Proposition 11.83, $\|f\|_p = 0$ and $f \in Z_p$. Hence, $\forall f \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$, $f \in Z_p$ and $\|f\|_p = 0$ if, and only if, $f = \vartheta_{\mathcal{Y}}$ a.e. in $\mathcal{X}$ and f is $\mathcal{B}$ -measurable. We will denote the norm in $L_p(\mathcal{X}, \mathcal{Y})$ by $\|\cdot\|_p$ and elements in $L_p(\mathcal{X}, \mathcal{Y})$ by $[f]$, where $f \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$. $\diamond$

Theorem 11.174 (Minkowski's Inequality) *Let $p \in [1, \infty) \subset \mathbb{R}$, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, and $\bar{L}_p(\mathcal{X}, \mathcal{Y})$ be the vector space over $\mathbb{K}$ with the pseudo-norm $\|\cdot\|_p$ as defined in Example 11.173. $\forall f, g \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$, then, $f + g \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$ and*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

When $1 < p < \infty$, equality holds if, and only if, $\exists \alpha, \beta \in [0, \infty) \subset \mathbb{R}$, which are not both zeros, such that $\alpha \mathcal{P} \circ f = \beta \mathcal{P} \circ g$ a.e. in $\mathcal{X}$ and $\mathcal{P} \circ (f + g) = \mathcal{P} \circ f + \mathcal{P} \circ g$ a.e. in $\mathcal{X}$.

Proof Clearly, f and g are $\mathcal{B}$ -measurable. By Propositions 7.23, 11.38, and 11.39, $f + g$ is $\mathcal{B}$ -measurable. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\|f\|_p \|g\|_p = 0$; Case 2: $\|f\|_p \|g\|_p > 0$. Case 1: $\|f\|_p \|g\|_p = 0$. Without loss of generality, assume that $\|g\|_p = 0$. Then, $\int_X (\mathcal{P}_p \circ g) \, d\mu = 0$. By Proposition 11.96, we have $\mathcal{P}_p \circ g = 0$ a.e. in $\mathcal{X}$. Thus, $g = \vartheta_{\mathcal{Y}}$ a.e. in $\mathcal{X}$ and $f = f + g$ a.e. in $\mathcal{X}$. By Proposition 11.83, the integral $\int_X (\mathcal{P}_p \circ (f + g)) \, d\mu = \int_X (\mathcal{P}_p \circ f) \, d\mu \in \mathbb{R}$. Hence, $\|f + g\|_p = \|f\|_p = \|f\|_p + \|g\|_p \in \mathbb{R}$ and $f + g \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$. Clearly, equality holds implies $\alpha = 0, \beta = 1, \alpha \mathcal{P} \circ f = \beta \mathcal{P} \circ g$ a.e. in $\mathcal{X}$ and $\mathcal{P} \circ (f + g) = \mathcal{P} \circ f + \mathcal{P} \circ g$ a.e. in $\mathcal{X}$.

Case 2: $\|f\|_p \|g\|_p > 0$. Let $\lambda := \frac{\|f\|_p}{\|f\|_p + \|g\|_p} \in (0, 1) \subset \mathbb{R}$. $\forall x \in X$, we have

$$\begin{aligned} \frac{\|f(x) + g(x)\|^p}{(\|f\|_p + \|g\|_p)^p} &\leq \left(\frac{\|f(x)\| + \|g(x)\|}{\|f\|_p + \|g\|_p} \right)^p \\ &= \left(\lambda \frac{\|f(x)\|}{\|f\|_p} + (1 - \lambda) \frac{\|g(x)\|}{\|g\|_p} \right)^p \leq \lambda \frac{\|f(x)\|^p}{\|f\|_p^p} + (1 - \lambda) \frac{\|g(x)\|^p}{\|g\|_p^p} \end{aligned}$$

where the second inequality follows from the convexity of the function t^p on $[0, \infty) \subset \mathbb{R}$. In the above, when $p > 1$, equality holds if, and only if, $\|f(x) + g(x)\| = \|f(x)\| + \|g(x)\|$ and $\frac{\|f(x)\|}{\|f\|_p} = \frac{\|g(x)\|}{\|g\|_p}$.

By Proposition 11.83, we have $\int_X \frac{(\mathcal{P}_p \circ (f+g))}{(\|f\|_p + \|g\|_p)^p} d\mu \leq \int_X \left(\lambda \frac{(\mathcal{P}_p \circ f)}{\|f\|_p^p} + (1 - \lambda) \frac{(\mathcal{P}_p \circ g)}{\|g\|_p^p} \right) d\mu = \frac{\lambda}{\|f\|_p^p} \int_X (\mathcal{P}_p \circ f) d\mu + \frac{1-\lambda}{\|g\|_p^p} \int_X (\mathcal{P}_p \circ g) d\mu = 1$. This implies that $\int_X (\mathcal{P}_p \circ (f+g)) d\mu \leq (\|f\|_p + \|g\|_p)^p < +\infty$ and $\|f+g\|_p \leq \|f\|_p + \|g\|_p$. Therefore, $f+g \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$.

Fix any $p \in (1, +\infty) \subset \mathbb{R}$. By Proposition 11.97, equality holds if, and only if, $\mathcal{P} \circ (f+g) = \mathcal{P} \circ f + \mathcal{P} \circ g$ a.e. in $\mathcal{X}$ and $\frac{\mathcal{P} \circ f}{\|f\|_p^p} = \frac{\mathcal{P} \circ g}{\|g\|_p^p}$ a.e. in $\mathcal{X}$. Equality implies that $\alpha = 1/\|f\|_p$ and $\beta = 1/\|g\|_p$ such that $\alpha \mathcal{P} \circ f = \beta \mathcal{P} \circ g$ a.e. in $\mathcal{X}$ and $\mathcal{P} \circ (f+g) = \mathcal{P} \circ f + \mathcal{P} \circ g$ a.e. in $\mathcal{X}$. On the other hand, if $\exists \alpha, \beta \in [0, \infty) \subset \mathbb{R}$, which are not both zeros, such that $\alpha \mathcal{P} \circ f = \beta \mathcal{P} \circ g$ a.e. in $\mathcal{X}$ and $\mathcal{P} \circ (f+g) = \mathcal{P} \circ f + \mathcal{P} \circ g$ a.e. in $\mathcal{X}$. Without loss of generality, assume that $\beta \neq 0$. Then, $\mathcal{P} \circ g = \alpha_1 \mathcal{P} \circ f$ a.e. in $\mathcal{X}$ with $\alpha_1 := \alpha/\beta$. It is easy to show that $\|g\|_p = \alpha_1 \|f\|_p$. This implies that $\alpha_1 = \|g\|_p / \|f\|_p$. Then, equality holds. Hence, equality holds, if, and only if, $\exists \alpha, \beta \in [0, \infty) \subset \mathbb{R}$, which are not both zeros, such that $\alpha \mathcal{P} \circ f = \beta \mathcal{P} \circ g$ a.e. in $\mathcal{X}$ and $\mathcal{P} \circ (f+g) = \mathcal{P} \circ f + \mathcal{P} \circ g$ a.e. in $\mathcal{X}$.

Hence, the result holds in both cases. This completes the proof of the theorem. $\square$

Definition 11.175 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space and $f : X \rightarrow \mathbb{R}$ be $\mathcal{B}$ -measurable. The essential supremum of f is

$$\operatorname{ess\,sup}_{x \in \mathcal{X}} f(x) := \inf \{ M \in \mathbb{R} \mid \mu(\{x \in X \mid f(x) > M\}) = 0 \} \in \mathbb{R}_e$$

Proposition 11.176 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space and $f : X \rightarrow \mathbb{R}$ and $g : X \rightarrow \mathbb{R}$ be $\mathcal{B}$ -measurable. Then,

- (i) $\operatorname{ess\,sup}_{x \in \mathcal{X}} (f(x) + g(x)) \leq \operatorname{ess\,sup}_{x \in \mathcal{X}} f(x) + \operatorname{ess\,sup}_{x \in \mathcal{X}} g(x)$;
- (ii) if $f \leq g$ a.e. in $\mathcal{X}$, then $\operatorname{ess\,sup}_{x \in \mathcal{X}} f(x) \leq \operatorname{ess\,sup}_{x \in \mathcal{X}} g(x)$;
- (iii) $\forall \alpha \in (0, \infty) \subset \mathbb{R}$, $\operatorname{ess\,sup}_{x \in \mathcal{X}} (\alpha f(x)) = \alpha \operatorname{ess\,sup}_{x \in \mathcal{X}} f(x)$; and $\forall \alpha \in [0, \infty) \subset \mathbb{R}$, $\operatorname{ess\,sup}_{x \in \mathcal{X}} (\alpha f(x)) = \alpha \operatorname{ess\,sup}_{x \in \mathcal{X}} f(x)$ if $\operatorname{ess\,sup}_{x \in \mathcal{X}} f(x) \in \mathbb{R}$.
- (iv) $\lambda := \operatorname{ess\,sup}_{x \in \mathcal{X}} f(x) \in \mathbb{R}_e$. Then, $f(x) \leq \lambda$ a.e. $x \in \mathcal{X}$.

Proof (i) We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu(X) = 0$; Case 2: $\mu(X) > 0$.

Case 1: $\mu(X) = 0$. Then, $\forall M \in \mathbb{R}$, $0 \leq \mu(\{x \in X \mid f(x) > M\}) \leq \mu(X) = 0$. Then, $\operatorname{ess\,sup}_{x \in \mathcal{X}} f(x) = -\infty$. Similarly, $\operatorname{ess\,sup}_{x \in \mathcal{X}} g(x) = \operatorname{ess\,sup}_{x \in \mathcal{X}} (f(x) + g(x)) = -\infty$. Then, the (i) holds.

Case 2: $\mu(X) > 0$. Note that $\bigcup_{n=1}^{\infty} A_n := \bigcup_{n=1}^{\infty} \{x \in X \mid f(x) > -n\} = X$. By Proposition 11.7, $\exists n \in \mathbb{N}$ such that $\mu(A_n) > 0$. Then, $\operatorname{ess\,sup}_{x \in \mathcal{X}} f(x) \geq -n > -\infty$. Similarly, $\operatorname{ess\,sup}_{x \in \mathcal{X}} g(x) > -\infty$ and $\operatorname{ess\,sup}_{x \in \mathcal{X}} (f(x) + g(x)) > -\infty$. Then, $\lambda := \operatorname{ess\,sup}_{x \in \mathcal{X}} f(x) +$

$\text{ess sup}_{x \in \mathcal{X}} g(x) \in (-\infty, +\infty] \subset \mathbb{R}_e$. If $\lambda = +\infty$, then (i) holds. On the other hand, if $\lambda \in \mathbb{R}$, $\forall M \in \mathbb{R}$ with $\lambda < M$, $\exists M_1, M_2 \in \mathbb{R}$ with $\text{ess sup}_{x \in \mathcal{X}} f(x) < M_1$ and $\text{ess sup}_{x \in \mathcal{X}} g(x) < M_2$ and $M_1 + M_2 = M$. Then, $\mu(\{x \in X \mid f(x) > M_1\}) = 0 = \mu(\{x \in X \mid g(x) > M_2\})$. By Propositions 7.23, 11.38, and 11.39, $f + g$ is $\mathcal{B}$ -measurable. Hence, $\mu(\{x \in X \mid f(x) + g(x) > M\}) = 0$ and $\text{ess sup}_{x \in \mathcal{X}} (f(x) + g(x)) \leq M$. By the arbitrariness of M , (i) holds.

(ii) Let $\lambda := \text{ess sup}_{x \in \mathcal{X}} g(x) \in \mathbb{R}_e$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\lambda = +\infty$; Case 2: $\lambda < +\infty$. Case 1: $\lambda = +\infty$. The result holds. Case 2: $\lambda < +\infty$. $\forall M \in \mathbb{R}$ with $\lambda < M$, $\mu(\{x \in X \mid g(x) > M\}) = 0$. Then, $\{x \in X \mid f(x) > M\} \subseteq \{x \in X \mid g(x) > M\} \cup \{x \in X \mid f(x) - g(x) > 0\}$. By Propositions 7.23, 11.38, and 11.39, $f - g$ is $\mathcal{B}$ -measurable. Then, $0 \leq \mu(\{x \in X \mid f(x) > M\}) \leq \mu(\{x \in X \mid g(x) > M\}) + \mu(\{x \in X \mid f(x) - g(x) > 0\}) = 0$. Hence, $\text{ess sup}_{x \in \mathcal{X}} f(x) \leq M$. By the arbitrariness of M , the result holds.

(iii) Let $\alpha \in (0, +\infty) \subset \mathbb{R}$. Then, $\text{ess sup}_{x \in \mathcal{X}} (\alpha f(x)) = \inf \{M \in \mathbb{R} \mid \mu(\{x \in X \mid \alpha f(x) > M\}) = 0\} = \inf \{\alpha M \in \mathbb{R} \mid \mu(\{x \in X \mid \alpha f(x) > \alpha M\}) = 0\} = \alpha \inf \{M \in \mathbb{R} \mid \mu(\{x \in X \mid f(x) > M\}) = 0\} = \alpha \text{ess sup}_{x \in \mathcal{X}} f(x)$, where the third equality follows from Proposition 3.81.

Let $\alpha = 0$ and $\text{ess sup}_{x \in \mathcal{X}} f(x) \in \mathbb{R}$. Then, $\mu(X) > 0$. Then, $\text{ess sup}_{x \in \mathcal{X}} (\alpha f(x)) = \text{ess sup}_{x \in \mathcal{X}} 0 = 0 = \alpha \text{ess sup}_{x \in \mathcal{X}} f(x)$.

(iv) We will distinguish three exhaustive and mutually exclusive cases: Case 1: $\mu(X) = 0$; Case 2: $\mu(X) > 0$ and $\lambda < +\infty$; Case 3: $\mu(X) > 0$ and $\lambda = +\infty$. Case 1: $\mu(X) = 0$. Then, $\lambda = -\infty$. Clearly, the result holds. Case 2: $\mu(X) > 0$ and $\lambda < +\infty$. Then, $\lambda \in \mathbb{R}$. $\forall n \in \mathbb{N}$, $\mu(E_n) := \mu(\{x \in X \mid f(x) > \lambda + 1/n\}) = 0$. By Proposition 11.7, $\mu(E) := \mu(\{x \in X \mid f(x) > \lambda\}) = \mu(\bigcup_{n=1}^{\infty} E_n) = \lim_{n \in \mathbb{N}} \mu(E_n) = 0$. Hence, $f(x) \leq \lambda$ a.e. $x \in \mathcal{X}$. Case 3: $\mu(X) > 0$ and $\lambda = +\infty$. Clearly, the result holds. Hence, (iv) holds in all three cases.

This completes the proof of the proposition. $\square$

Example 11.177 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, and $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})$ be the vector space of functions of $\mathcal{X}$ to $\mathcal{Y}$ as defined in Example 6.20 with the usual vector addition, scalar multiplication, and the null vector ϑ . Let $Z_{\infty} := \{f \in \mathcal{M}(\mathcal{X}, \mathcal{Y}) \mid f \text{ is } \mathcal{B}\text{-measurable and } \text{ess sup}_{x \in \mathcal{X}} \|f(x)\| < +\infty\}$. Define $\|\cdot\|_{\infty} : Z_{\infty} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\|f\|_{\infty} = \max\{\text{ess sup}_{x \in \mathcal{X}} \|f(x)\|, 0\}$, $\forall f \in Z_{\infty}$. We will next show that Z_{∞} is a subspace of $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})$ and $\|\cdot\|_{\infty}$ defines a pseudo-norm on Z_{∞} . Clearly, $\vartheta \in Z_{\infty} \neq \emptyset$. $\forall f, g \in Z_{\infty}$, $\forall \alpha \in \mathbb{K}$, by Propositions 7.23, 11.38, and 11.39, $f + g$ and αf are $\mathcal{B}$ measurable. By Proposition 11.176,

$$\begin{aligned} \|f + g\|_{\infty} &= \max\{\text{ess sup}_{x \in \mathcal{X}} \|f(x) + g(x)\|, 0\} \\ &\leq \max\{\text{ess sup}_{x \in \mathcal{X}} (\|f(x)\| + \|g(x)\|), 0\} \end{aligned}$$

$$\begin{aligned}
&\leq \max\{\operatorname{ess\,sup}_{x \in \mathcal{X}} \|f(x)\| + \operatorname{ess\,sup}_{x \in \mathcal{X}} \|g(x)\|, 0\} \\
&= \begin{cases} \operatorname{ess\,sup}_{x \in \mathcal{X}} \|f(x)\| + \operatorname{ess\,sup}_{x \in \mathcal{X}} \|g(x)\| & \text{if } \mu(X) > 0 \\ 0 & \text{if } \mu(X) = 0 \end{cases} \\
&\leq \|f\|_\infty + \|g\|_\infty < +\infty
\end{aligned}$$

and

$$\begin{aligned}
\|\alpha f\|_\infty &= \max\{\operatorname{ess\,sup}_{x \in \mathcal{X}} \|\alpha f(x)\|, 0\} = \max\{\operatorname{ess\,sup}_{x \in \mathcal{X}} |\alpha| \|f(x)\|, 0\} \\
&= \begin{cases} |\alpha| \operatorname{ess\,sup}_{x \in \mathcal{X}} \|f(x)\| & \text{if } \mu(X) > 0 \\ 0 & \text{if } \mu(X) = 0 \end{cases} = |\alpha| \|f\|_\infty < +\infty
\end{aligned}$$

Then, $f + g, \alpha f \in Z_\infty$. Hence, Z_∞ is a subspace of $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})$. This implies that $(Z_\infty, \mathbb{K})$ is a vector space. Clearly, $\|\vartheta\|_\infty = 0$. Therefore, $\|\cdot\|_\infty$ defines a pseudo-norm on $(Z_\infty, \mathbb{K})$. By Proposition 7.47, the quotient space of $(Z_\infty, \mathbb{K})$ modulo $\|\cdot\|_\infty$ is a normed linear space, to be denoted $L_\infty(\mathcal{X}, \mathcal{Y})$. We will denote the vector space $(Z_\infty, \mathbb{K})$ with the pseudo-norm $\|\cdot\|_\infty$ by $\bar{L}_\infty(\mathcal{X}, \mathcal{Y})$.

$\forall f \in Z_\infty$ with $\|f\|_\infty = 0$, we have $\operatorname{ess\,sup}_{x \in \mathcal{X}} \|f(x)\| \leq 0$. Then, by Proposition 11.176, $\mathcal{P} \circ f = 0$ a.e. in $\mathcal{X}$. Hence, $f = \vartheta_{\mathcal{Y}}$ a.e. in $\mathcal{X}$. On the other hand, $\forall f \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ with $f = \vartheta_{\mathcal{Y}}$ a.e. in $\mathcal{X}$ and f being $\mathcal{B}$ -measurable, we have $\|f\|_\infty = 0$ and $f \in Z_\infty$. Hence, $\forall f \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$, $f \in Z_\infty$ and $\|f\|_\infty = 0$ if, and only if, $f = \vartheta_{\mathcal{Y}}$ a.e. in $\mathcal{X}$ and f is $\mathcal{B}$ -measurable. We will denote the norm in $L_\infty(\mathcal{X}, \mathcal{Y})$ by $\|\cdot\|_\infty$ and elements in $L_\infty(\mathcal{X}, \mathcal{Y})$ by $[f]$, where $f \in \bar{L}_\infty(\mathcal{X}, \mathcal{Y})$. $\diamond$

In the following, we will write $\lim_{n \in \mathbb{N}} z_n \doteq z$ in $\bar{L}_p(\mathcal{X}, \mathcal{Y})$, when the sequence $(z_n)_{n=1}^\infty \subseteq \bar{L}_p(\mathcal{X}, \mathcal{Y})$ converges to $z \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$ in $\bar{L}_p(\mathcal{X}, \mathcal{Y})$ pseudo-norm. We will simply write $\lim_{n \in \mathbb{N}} z_n \doteq z$ if there is no confusion in which pseudo-norm convergence occurs. For $z \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$, we will denote the corresponding equivalence class in $L_p(\mathcal{X}, \mathcal{Y})$ by $[z]$. Then, for $(z_n)_{n=1}^\infty \subseteq \bar{L}_p(\mathcal{X}, \mathcal{Y})$, $\lim_{n \in \mathbb{N}} z_n \doteq z$ in $\bar{L}_p(\mathcal{X}, \mathcal{Y})$ if, and only if, $\lim_{n \in \mathbb{N}} [z_n] = [z]$ in $L_p(\mathcal{X}, \mathcal{Y})$ (or simply $\lim_{n \in \mathbb{N}} [z_n] = [z]$ when there is no confusion in which norm convergence occurs.)

Theorem 11.178 (Hölder's Inequality) *Let $p \in [1, +\infty) \subset \mathbb{R}$ and $q \in (1, +\infty] \subset \mathbb{R}_e$ with $1/p + 1/q = 1$, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, and $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable. Then, $\forall f \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$, $\forall g \in \bar{L}_q(\mathcal{X}, \mathcal{Y}^*)$, the function $r : X \rightarrow \mathbb{K}$, defined by $r(x) = \langle g(x), f(x) \rangle$, $\forall x \in X$, is absolutely integrable over $\mathcal{X}$ and*

$$\int_X |\langle g(x), f(x) \rangle| d\mu(x) \leq \|f\|_p \|g\|_q$$

When $q < \infty$, equality holds if, and only if, $|\langle g(x), f(x) \rangle| = \|f(x)\| \cdot \|g(x)\|$ a.e. $x \in \mathcal{X}$ and $\exists \alpha, \beta \in \mathbb{R}$, which are not both zeros, such that $\alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g$ a.e. in $\mathcal{X}$.

Proof By Propositions 7.65, 11.38, 11.39, and 7.21, r and $\mathcal{P} \circ r$ are $\mathcal{B}$ -measurable. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $q = \infty$; Case 2: $1 < q < +\infty$. Case 1: $q = \infty$. Then, $p = 1$. By Proposition 11.176, $\|g(x)\| \leq \|g\|_\infty$ a.e. $x \in \mathcal{X}$. Note that $|\langle\langle g(x), f(x) \rangle\rangle| \leq \|g(x)\| \|f(x)\| \leq \|g\|_\infty \|f(x)\|$ a.e. $x \in \mathcal{X}$. Then, by Propositions 11.83 and 11.92,

$$\int_X |\langle\langle g(x), f(x) \rangle\rangle| d\mu(x) \leq \|g\|_\infty \int_X \mathcal{P} \circ f d\mu = \|f\|_1 \|g\|_\infty$$

Case 2: $1 < q < +\infty$. Then, $1 < p < +\infty$. We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: $\|f\|_p \|g\|_q = 0$; Case 2b: $\|f\|_p \|g\|_q > 0$. Case 2a: $\|f\|_p \|g\|_q = 0$. Without loss of generality, assume $\|g\|_q = 0$. Then, $g = 0$ a.e. in $\mathcal{X}$. This implies that $|\langle\langle g(x), f(x) \rangle\rangle| = 0$ a.e. $x \in \mathcal{X}$ and, by Propositions 11.83 and 11.75,

$$\int_X |\langle\langle g(x), f(x) \rangle\rangle| d\mu(x) = 0 = \|f\|_p \|g\|_q$$

Equality holds $\Rightarrow \alpha = 0$, $\beta = 1$, $\alpha \mathcal{P}_p \circ f = 0 = \beta \mathcal{P}_q \circ g$ a.e. in $\mathcal{X}$, and $|\langle\langle g(x), f(x) \rangle\rangle| = 0 = \|f(x)\| \|g(x)\|$ a.e. $x \in \mathcal{X}$. This subcase is proved. Case 2b: $\|f\|_p \|g\|_q > 0$. Then, $\|f\|_p > 0$ and $\|g\|_q > 0$. $\forall x \in X$, by Lemma 7.7 with $a = \left(\frac{\|f(x)\|}{\|f\|_p}\right)^p$, $b = \left(\frac{\|g(x)\|}{\|g\|_q}\right)^q$, and $\lambda = 1/p$, we have

$$\frac{|\langle\langle g(x), f(x) \rangle\rangle|}{\|f\|_p \|g\|_q} \leq \frac{\|g(x)\| \|f(x)\|}{\|g\|_q \|f\|_p} \leq \frac{1}{p} \frac{\|f(x)\|^p}{\|f\|_p^p} + \frac{1}{q} \frac{\|g(x)\|^q}{\|g\|_q^q}$$

with equality holding if, and only if, $|\langle\langle g(x), f(x) \rangle\rangle| = \|g(x)\| \|f(x)\|$ and $\frac{\|f(x)\|^p}{\|f\|_p^p} = \frac{\|g(x)\|^q}{\|g\|_q^q}$. Integrating the above inequality over $\mathcal{X}$, we have, by Propositions 11.83 and 11.97,

$$\frac{\int_X |\langle\langle g(x), f(x) \rangle\rangle| d\mu(x)}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

with equality holding if, and only if, $|\langle\langle g(x), f(x) \rangle\rangle| = \|g(x)\| \cdot \|f(x)\|$ a.e. $x \in \mathcal{X}$ and $\frac{\mathcal{P}_p \circ f}{\|f\|_p^p} = \frac{\mathcal{P}_q \circ g}{\|g\|_q^q}$ a.e. in $\mathcal{X}$. Equality $\Rightarrow \alpha = 1/\|f\|_p^p$, $\beta = 1/\|g\|_q^q$, $\alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g$ a.e. in $\mathcal{X}$ and $|\langle\langle g(x), f(x) \rangle\rangle| = \|g(x)\| \|f(x)\|$ a.e. $x \in \mathcal{X}$. On the other hand, if $|\langle\langle g(x), f(x) \rangle\rangle| = \|g(x)\| \|f(x)\|$ a.e. $x \in \mathcal{X}$ and $\exists \alpha, \beta \in \mathbb{R}$, which are not both zeros, such that $\alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g$ a.e. in $\mathcal{X}$, then, without loss of generality, assume $\beta \neq 0$. Let $\alpha_1 = \alpha/\beta$. Then, $\alpha_1 \mathcal{P}_p \circ f = \mathcal{P}_q \circ g$ a.e. in $\mathcal{X}$. Hence, $\alpha_1 \|f\|_p^p = \|g\|_q^q$, which further implies that $\alpha_1 = \|g\|_q^q / \|f\|_p^p$. Hence, $\frac{\mathcal{P}_p \circ f}{\|f\|_p^p} = \frac{\mathcal{P}_q \circ g}{\|g\|_q^q}$ a.e. in $\mathcal{X}$. This implies equality. Therefore, equality holds

if, and only if, $|\langle g(x), f(x) \rangle| = \|g(x)\| \|f(x)\|$ a.e. $x \in \mathcal{X}$ and $\exists \alpha, \beta \in \mathbb{R}$, which are not both zeros, such that $\alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g$ a.e. in $\mathcal{X}$. This sub-case is proved.

This completes the proof of the theorem. $\square$

When $p = 2 = q$, the Hölder's inequality becomes the well-known *Cauchy-Schwarz Inequality*:

$$\int_X |\langle g(x), f(x) \rangle| d\mu(x) \leq \left(\int_X \mathcal{P}_2 \circ f d\mu \right)^{1/2} \left(\int_X \mathcal{P}_2 \circ g d\mu \right)^{1/2}$$

Example 11.179 Let $p \in [1, \infty) \subset \mathbb{R}$, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, and $L_p(\mathcal{X}, \mathcal{Y})$ be the normed linear space over $\mathbb{K}$ as defined in Example 11.173. We will show that $L_p(\mathcal{X}, \mathcal{Y})$ is a Banach space by Proposition 7.27. Define $l : [0, \infty) \subset \mathbb{R} \rightarrow [0, \infty) \subset \mathbb{R}$ by $l(t) = t^p$, $\forall t \in [0, \infty) \subset \mathbb{R}$. Fix any $([f_n])_{n=1}^\infty \subseteq L_p(\mathcal{X}, \mathcal{Y})$ with $f_n \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$, $\forall n \in \mathbb{N}$, and $\sum_{n=1}^\infty \|f_n\|_p =: M \in [0, \infty) \subset \mathbb{R}$. $\forall n \in \mathbb{N}$, define $g_n : X \rightarrow [0, \infty) \subset \mathbb{R}$ by $g_n(x) = \sum_{i=1}^n \|f_i(x)\|$, $\forall x \in X$. By Propositions 7.23, 7.21, 11.38, and 11.39, g_n is $\mathcal{B}$ -measurable. Note that $\mathcal{P} \circ f_n \in \bar{L}_p(\mathcal{X}, \mathbb{R})$. Then, $g_n \in \bar{L}_p(\mathcal{X}, \mathbb{R})$ and $\|g_n\|_p \leq \sum_{i=1}^n \|\mathcal{P} \circ f_i\|_p = \sum_{i=1}^n \|f_i\|_p = \sum_{i=1}^n \|f_i\|_p \leq M$. This implies that $\int_X (l \circ g_n) d\mu \leq M^p$. Clearly, $(g_n(x))^p \leq (g_{n+1}(x))^p$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Proposition 11.82 and Monotone Convergence Theorem 11.81, $\exists g : X \rightarrow [0, \infty) \subset \mathbb{R}$, which is $\mathcal{B}$ -measurable, such that $\lim_{n \in \mathbb{N}} l \circ g_n = l \circ g$ a.e. in $\mathcal{X}$ and $\int_X (l \circ g) d\mu = \lim_{n \in \mathbb{N}} \int_X (l \circ g_n) d\mu \leq M^p$.

Let $E := \{x \in X \mid (g_n(x))_{n=1}^\infty \text{ does not converge to } g(x)\}$. Then, $E \in \mathcal{B}$ and $\mu(E) = 0$. $\forall x \in X \setminus E$, $\lim_{n \in \mathbb{N}} g_n(x) = g(x) \in \mathbb{R}$ and $\lim_{n \in \mathbb{N}} \sum_{i=1}^n \|f_i(x)\| = g(x) \in \mathbb{R}$. By Proposition 7.27 and the completeness of $\mathcal{Y}$, we have $\lim_{n \in \mathbb{N}} \sum_{i=1}^n f_i(x) =: \lim_{n \in \mathbb{N}} s_n(x) =: s(x) \in \mathcal{Y}$. Define $s(x) = s_n(x) = \vartheta_{\mathcal{Y}}$, $\forall x \in E$ and $\forall n \in \mathbb{N}$. Then, by Proposition 11.41, s_n is $\mathcal{B}$ -measurable, $\forall n \in \mathbb{N}$. Clearly, $\lim_{n \in \mathbb{N}} s_n(x) = s(x)$, $\forall x \in X$. By Proposition 11.48, s is $\mathcal{B}$ -measurable. By Lemma 11.43, $s_n = \sum_{i=1}^n f_i$ a.e. in $\mathcal{X}$. This yields that $[s_n] = [\sum_{i=1}^n f_i] = \sum_{i=1}^n [f_i]$ and $s_n \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$. Then, by Proposition 11.50, $\lim_{n \in \mathbb{N}} \sum_{i=1}^n f_i = s$ a.e. in $\mathcal{X}$. Note that $\lim_{n \in \mathbb{N}} \mathcal{P}_p \circ (s_n(x) - s(x)) = 0$, $\forall x \in X$. Note also $\|s_n(x) - s(x)\|^p \leq (\|s_n(x)\| + \|s(x)\|)^p \leq (g_n(x) + g(x))^p \leq 2^p (g(x))^p$, $\forall x \in X \setminus E$, $\forall n \in \mathbb{N}$, and $\|s_n(x) - s(x)\|^p = 0 \leq 2^p (g(x))^p$, $\forall x \in E$, $\forall n \in \mathbb{N}$. Hence, $\|s_n(x) - s(x)\|^p \leq 2^p (g(x))^p$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91, $\lim_{n \in \mathbb{N}} \int_X (\mathcal{P}_p \circ (s_n - s)) d\mu = 0$. This implies that $\lim_{n \in \mathbb{N}} \|s_n - s\|_p = 0$. Hence, $s \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$ and $\lim_{n \in \mathbb{N}} \sum_{i=1}^n f_i = \lim_{n \in \mathbb{N}} s_n = s$. Hence, $([f_n])_{n=1}^\infty$ is summable in $L_p(\mathcal{X}, \mathcal{Y})$. By Proposition 7.27 and the arbitrariness of $([f_n])_{n=1}^\infty$, we have $L_p(\mathcal{X}, \mathcal{Y})$ is complete. This shows that $L_p(\mathcal{X}, \mathcal{Y})$ is a Banach space when $\mathcal{X}$ is a σ -finite measure space and $\mathcal{Y}$ is a separable Banach space. $\diamond$

Example 11.180 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, and $L_\infty(\mathcal{X}, \mathcal{Y})$ be the normed linear space

over $\mathbb{K}$ as defined in Example 11.177. We will show that $L_\infty(\mathcal{X}, \mathcal{Y})$ is a Banach space. Fix any Cauchy sequence $([f_n])_{n=1}^\infty \subseteq L_\infty(\mathcal{X}, \mathcal{Y})$ with $f_n \in \bar{L}_\infty(\mathcal{X}, \mathcal{Y})$, $\forall n \in \mathbb{N}$. $\forall k \in \mathbb{N}$, $\exists N_k \in \mathbb{N}$, $\forall n, m \in \mathbb{N}$ with $n \geq N_k$ and $m \geq N_k$, we have $\|f_n - f_m\|_\infty < 1/k$. Then, by Proposition 11.176, $A_{n,m,k} := \{x \in X \mid \|f_n(x) - f_m(x)\| \geq 1/k\} \in \mathcal{B}$ and $\mu(A_{n,m,k}) = 0$. Let $A := \bigcup_{k=1}^\infty \bigcup_{n=N_k}^\infty \bigcup_{m=N_k}^\infty A_{n,m,k} \in \mathcal{B}$. Clearly, $\mu(A) = 0$. $\forall x \in X \setminus A$, $(f_n(x))_{n=1}^\infty \subseteq \mathcal{Y}$ is a Cauchy sequence, which converges to $f(x) \in \mathcal{Y}$ by the completeness of $\mathcal{Y}$. Define $f(x) = \vartheta_y$, $\forall x \in A$. Then, $f : X \rightarrow \mathcal{Y}$ is well defined and $\lim_{n \in \mathbb{N}} f_n(x) = f(x)$, $\forall x \in X \setminus A$. By Propositions 11.48 and 11.41, f is $\mathcal{B}$ -measurable. $\forall k \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq N_k$, $\forall x \in X \setminus A$, by Propositions 3.66, 3.67, 7.21, and 7.23, we have $\|f_n(x) - f(x)\| = \lim_{m \in \mathbb{N}} \|f_n(x) - f_m(x)\| \leq 1/k$. Then, $0 \leq \mu(\{x \in X \mid \|f(x) - f_n(x)\| > 1/k\}) \leq \mu(A) = 0$. This shows that $\|f_n - f\|_\infty \leq 1/k$. Then, $\lim_{n \in \mathbb{N}} \|f_n - f\|_\infty = 0$, $\lim_{n \in \mathbb{N}} f_n \doteq f$, and $f \in \bar{L}_\infty(\mathcal{X}, \mathcal{Y})$. Hence, $\lim_{n \in \mathbb{N}} [f_n] = [f]$ in $L_\infty(\mathcal{X}, \mathcal{Y})$. Therefore, $L_\infty(\mathcal{X}, \mathcal{Y})$ is a Banach space when $\mathcal{X}$ is a measure space and $\mathcal{Y}$ is a separable Banach space. $\diamond$

Proposition 11.181 *Let $p \in [1, \infty) \subset \mathbb{R}$, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, $L_p(\mathcal{X}, \mathcal{Y})$ be the normed linear space over $\mathbb{K}$ as defined in Example 11.173, and $f \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$. Then, there exists a sequence of simple functions $(\phi_i)_{i=1}^\infty$, $\phi_i : \mathcal{X} \rightarrow \mathcal{Y}$, $\forall i \in \mathbb{N}$, such that $\lim_{i \in \mathbb{N}} \phi_i = f$ a.e. in $\mathcal{X}$, $\|\phi_i(x)\| \leq \|f(x)\|$, $\forall x \in X$, $\forall i \in \mathbb{N}$, $\lim_{i \in \mathbb{N}} \int_X \mathcal{P}_p \circ (\phi_i - f) d\mu = 0$, and $\lim_{i \in \mathbb{N}} \phi_i \doteq f$ in $\bar{L}_p(\mathcal{X}, \mathcal{Y})$.*

Proof Since $f \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$, then f is $\mathcal{B}$ -measurable and $\mathcal{P}_p \circ f$ is integrable over $\mathcal{X}$. By Proposition 11.66, there exists a sequence of simple functions $(\phi_i)_{i=1}^\infty$, $\phi_i : \mathcal{X} \rightarrow \mathcal{Y}$, $\forall i \in \mathbb{N}$, such that $\|\phi_i(x)\| \leq \|f(x)\|$, $\forall x \in X$, $\forall i \in \mathbb{N}$, and $\lim_{i \in \mathbb{N}} \phi_i = f$ a.e. in $\mathcal{X}$. By Propositions 7.23, 7.21, 11.38, and 11.39, $\mathcal{P}_p \circ (\phi_i - f)$ is $\mathcal{B}$ -measurable, $\forall i \in \mathbb{N}$. Note that, by Propositions 7.23, 7.21, 11.52, and 11.53, $\lim_{i \in \mathbb{N}} \mathcal{P}_p \circ (\phi_i - f) = 0$ a.e. in $\mathcal{X}$ and $\mathcal{P}_p \circ (\phi_i - f)(x) \leq 2^p \mathcal{P}_p \circ f(x)$, $\forall x \in X$, $\forall i \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91, we have $\lim_{i \in \mathbb{N}} \int_X \mathcal{P}_p \circ (\phi_i - f) d\mu = 0$. Hence, $\lim_{i \in \mathbb{N}} \phi_i \doteq f$. This completes the proof of the proposition. $\square$

Proposition 11.182 *Let $p \in [1, \infty) \subset \mathbb{R}$, $\mathbb{X} := (\mathcal{X}, \mathcal{B}, \mu)$ be a σ -finite normal topological measure space, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, and $f \in \bar{L}_p(\mathbb{X}, \mathcal{Y})$. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists$ a continuous function $g : \mathbb{X} \rightarrow \mathcal{Y}$ such that $g \in \bar{L}_p(\mathbb{X}, \mathcal{Y})$ and $\|g - f\|_p < \epsilon$.*

Proof Fix $\epsilon \in (0, \infty) \subset \mathbb{R}$. By Proposition 11.181, there exists a simple function $\phi : \mathbb{X} \rightarrow \mathcal{Y}$ such that $\|\phi(x)\| \leq \|f(x)\|$, $\forall x \in \mathbb{X}$, and $\|\phi - f\|_p < \epsilon/2$. Let ϕ admit the canonical representation $\phi = \sum_{i=1}^n y_i \chi_{A_i, \mathbb{X}}$, where $n \in \mathbb{Z}_+$, $y_1, \dots, y_n \in \mathcal{Y}$ are distinct and none

equals to $\vartheta_{\mathcal{Y}}$, $A_1, \dots, A_n \in \mathcal{B}$ are pairwise disjoint, nonempty, and of finite measure. $\forall i \in \{1, \dots, n\}$, by $\mathbb{X}$ being a topological measure space and Proposition 11.27, $\exists U_i, \mathbb{X} \setminus F_i \in \mathcal{O}_{\mathcal{X}}$ such that $F_i \subseteq A_i \subseteq U_i$ and $\mu(U_i \setminus F_i) = \mu(U_i \setminus A_i) + \mu(A_i \setminus F_i) < \frac{\epsilon^p}{2^p(n+1)^p \|y_i\|^p}$. By $\mathcal{X}$ being a normal topological space and Urysohn's Lemma 3.55, there exists a continuous function $g_i : \mathbb{X} \rightarrow [0, 1] \subset \mathbb{R}$ such that $g_i(x) = 1$, $\forall x \in F_i$ and $g_i(x) = 0$, $\forall x \in \mathbb{X} \setminus U_i$. By Proposition 11.37, g_i is $\mathcal{B}$ -measurable. By Definition 11.79, $\|g_i - \chi_{A_i, \mathbb{X}}\|_p \leq \frac{\epsilon}{2(n+1)\|y_i\|}$.

Define $g : \mathbb{X} \rightarrow \mathcal{Y}$ by $g(x) = \sum_{i=1}^n y_i g_i(x)$, $\forall x \in \mathbb{X}$. By Propositions 7.23, 3.12, and 3.32, g is continuous. By Proposition 11.37, g is $\mathcal{B}$ -measurable. Then, $\|g - f\|_p \leq \|g - \phi\|_p + \|\phi - f\|_p \leq \sum_{i=1}^n \|y_i(g_i - \chi_{A_i, \mathbb{X}})\|_p + \epsilon/2 = \sum_{i=1}^n \|y_i\| \|g_i - \chi_{A_i, \mathbb{X}}\|_p + \epsilon/2 \leq \sum_{i=1}^n \frac{\|y_i\| \epsilon}{2(n+1)\|y_i\|} + \epsilon/2 < \epsilon$. Hence, $g \in \bar{L}_p(\mathbb{X}, \mathcal{Y})$.

This completes the proof of the proposition. $\square$

Proposition 11.183 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a separable Banach space, $f : X \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$, $f|_E$ be absolutely integrable over $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$, which is the finite measure subspace of $\mathcal{X}$, and $M \in [0, \infty) \subset \mathbb{R}$. Assume that, $\forall E \in \mathcal{B}$ with $0 < \mu(E) < \infty$, we have $\left\| \frac{1}{\mu(E)} \int_E f \, d\mu \right\| \leq M$. Then, $\mathcal{P} \circ f \leq M$ a.e. in $\mathcal{X}$.*

Proof Consider the open set $O := \{y \in \mathcal{Y} \mid \|y\| > M\} \subseteq \mathcal{Y}$. Since $\mathcal{Y}$ is separable, by Propositions 4.38 and 4.4, O is second countable and separable. Let $D \subseteq O$ be a countable dense set in O (the relative closure of D with respect to O equals to O). It is easy to show that $\mathcal{M} := \{\mathcal{B}_{\mathcal{Y}}(y, r) \subseteq O \mid y \in D, r \in \mathbb{Q}, r > 0\}$ is a countable basis for O . Let $E := f_{\text{inv}}(O) = f_{\text{inv}}(\bigcup_{\mathcal{B}_{\mathcal{Y}}(y, r) \in \mathcal{M}} \mathcal{B}_{\mathcal{Y}}(y, r)) = \bigcup_{\mathcal{B}_{\mathcal{Y}}(y, r) \in \mathcal{M}} f_{\text{inv}}(\mathcal{B}_{\mathcal{Y}}(y, r))$.

We will show that $\mu(f_{\text{inv}}(\mathcal{B}_{\mathcal{Y}}(y, r))) = 0$, $\forall \mathcal{B}_{\mathcal{Y}}(y, r) \in \mathcal{M}$, by an argument of contradiction. Suppose that $\exists y \in D$, $\exists r \in \mathbb{Q}$ with $r > 0$ such that $\mathcal{B}_{\mathcal{Y}}(y, r) \subseteq O$ and $\mu(\bar{E}) := \mu(f_{\text{inv}}(\mathcal{B}_{\mathcal{Y}}(y, r))) > 0$. Since $\mathcal{X}$ is σ -finite, then $\exists \bar{E} \in \mathcal{B}$ with $\bar{E} \subseteq \bar{E}$ such that $0 < \mu(\bar{E}) < +\infty$. $\forall x \in \bar{E}$, we have $x \in \bar{E}$ and $\|f(x) - y\| < r$. Then, $\left\| \frac{1}{\mu(\bar{E})} \int_{\bar{E}} f \, d\mu - y \right\| = \left\| \frac{1}{\mu(\bar{E})} \int_{\bar{E}} (f - y) \, d\mu \right\| = \frac{1}{\mu(\bar{E})} \left\| \int_{\bar{E}} (f(x) - y) \, d\mu(x) \right\| \leq \frac{1}{\mu(\bar{E})} \int_{\bar{E}} \|f(x) - y\| \, d\mu(x) < r$, where the first equality follows from Propositions 11.92 and 11.75; the first inequality follows from Proposition 11.92; and the second inequality follows from Proposition 11.97. This implies that $\frac{1}{\mu(\bar{E})} \int_{\bar{E}} f \, d\mu \in \mathcal{B}_{\mathcal{Y}}(y, r) \subseteq O$

and $\left\| \frac{1}{\mu(\bar{E})} \int_{\bar{E}} f \, d\mu \right\| > M$. This contradicts the assumption. Hence, $\mu(f_{\text{inv}}(\mathcal{B}_{\mathcal{Y}}(y, r))) = 0$, $\forall \mathcal{B}_{\mathcal{Y}}(y, r) \in \mathcal{M}$.

Then, $0 \leq \mu(E) \leq \sum_{\mathcal{B}_{\mathcal{Y}}(y, r) \in \mathcal{M}} \mu(f_{\text{inv}}(\mathcal{B}_{\mathcal{Y}}(y, r))) = 0$. Hence, $\mathcal{P} \circ f \leq M$ a.e. in $\mathcal{X}$. $\square$

Lemma 11.184 *Let $p, q \in (1, \infty) \subset \mathbb{R}$ with $1/p + 1/q = 1$, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable, $\mathcal{Z} := L_p(\mathcal{X}, \mathcal{Y})$ be the normed linear space over $\mathbb{K}$ as defined in Example 11.173, and $g : X \rightarrow \mathcal{Y}^*$ be $\mathcal{B}$ -measurable. Assume that*

- (i) $\forall E \in \mathcal{B}$ with $\mu(E) < +\infty$, g is absolutely integrable over $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$, which is the finite measure subspace of $\mathcal{X}$;
- (ii) $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\forall$ simple function $\phi : \mathcal{X} \rightarrow \mathcal{Y}$ ($\phi \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$), the function $\langle\langle g(\cdot), \phi(\cdot) \rangle\rangle : X \rightarrow \mathbb{K}$ is absolutely integrable over $\mathcal{X}$ and $|\int_X \langle\langle g(x), \phi(x) \rangle\rangle d\mu(x)| \leq M \|\phi\|_p$.

Then, $g \in \bar{L}_q(\mathcal{X}, \mathcal{Y}^*)$ and $\|g\|_q \leq M$.

Proof By Proposition 11.116, define a σ -finite $\mathcal{Y}^*$ -valued measure ν on $(X, \mathcal{B})$ by: $\nu(E)$ is undefined, $\forall E \in \mathcal{B}$ with $\int_E \mathcal{P} \circ g d\mu = \infty$; $\nu(E) = \int_E g d\mu \in \mathcal{Y}^*$, $\forall E \in \mathcal{B}$ with $\int_E \mathcal{P} \circ g d\mu < \infty$. Then, $\mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ g d\mu$, $\forall E \in \mathcal{B}$. By (i), $\mathcal{P} \circ \nu(E) < \infty$, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$. By Proposition 11.92, $\nu(E) = \vartheta_{\mathcal{Y}^*}$, $\forall E \in \mathcal{B}$ with $\mu(E) = 0$. Fix any $E \in \mathcal{B}$ with $\mu(E) < \infty$. Define $\mathcal{P}_q \circ \nu(E) \in [0, \infty] \subset \mathbb{R}_e$ by

$$\mathcal{P}_q \circ \nu(E) := \sup_{\substack{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \cup_{i=1}^n E_i = E \\ E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{\substack{i=1 \\ \mu(E_i) > 0}}^n \frac{\|\nu(E_i)\|^q}{(\mu(E_i))^q} \mu(E_i)$$

$\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_i)_{i=1}^n \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^n E_i$, $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\forall i \in \{1, \dots, n\}$ with $\mu(E_i) > 0$, by Lemma 7.75, $\exists y_i \in \mathcal{Y}$ such that $\|y_i\| = \|\nu(E_i)\|^{q-1} / (\mu(E_i))^{q-1}$ and $\langle\langle \nu(E_i), y_i \rangle\rangle \geq (1 - \epsilon) \|\nu(E_i)\| \|y_i\| = (1 - \epsilon) \|\nu(E_i)\|^q / (\mu(E_i))^{q-1}$. Define a simple function $\phi = \sum_{i=1}^n \nu(E_i) y_i \chi_{E_i, X}$. Then,

$$\begin{aligned} \|\phi\|_p &= \left(\sum_{\substack{i=1 \\ \mu(E_i) > 0}}^n \|\nu(E_i)\|^{(q-1)p} (\mu(E_i))^{(1-q)p} \mu(E_i) \right)^{1/p} \\ &= \left(\sum_{\substack{i=1 \\ \mu(E_i) > 0}}^n \|\nu(E_i)\|^q (\mu(E_i))^{-q} \mu(E_i) \right)^{1/p} \end{aligned}$$

By (ii), we have

$$\begin{aligned} \sum_{\substack{i=1 \\ \mu(E_i) > 0}}^n (1 - \epsilon) \|\nu(E_i)\|^q / (\mu(E_i))^{q-1} &\leq \left| \sum_{\substack{i=1 \\ \mu(E_i) > 0}}^n \langle\langle \nu(E_i), y_i \rangle\rangle \right| \\ &= \left| \sum_{\substack{i=1 \\ \mu(E_i) > 0}}^n \langle\langle \int_{E_i} g d\mu, y_i \rangle\rangle \right| = \left| \sum_{\substack{i=1 \\ \mu(E_i) > 0}}^n \int_{E_i} \langle\langle g(x), y_i \rangle\rangle d\mu(x) \right| \end{aligned}$$

$$= \left| \int_X \langle g(x), \phi(x) \rangle d\mu(x) \right| \leq M \left(\sum_{\substack{i=1 \\ \mu(E_i) > 0}}^n \|\nu(E_i)\|^q (\mu(E_i))^{1-q} \right)^{1/p}$$

where the second and third equalities follow from Proposition 11.92. Hence, we have $\sum_{i=1}^n \mu(E_i) > 0 \|\nu(E_i)\|^q (\mu(E_i))^{-q} \mu(E_i) \leq M^q / (1 - \epsilon)^q$. By the arbitrariness of ϵ , we have $\sum_{i=1}^n \mu(E_i) > 0 \|\nu(E_i)\|^q (\mu(E_i))^{-q} \mu(E_i) \leq M^q$. Then, $\mathcal{P}_q \circ \nu(E) \leq M^q$.

Since $\mathcal{X}$ is σ -finite, then $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty X_n$ and $\mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $X_n \subseteq X_{n+1}$, $\forall n \in \mathbb{N}$. By Proposition 11.66, there exists a sequence of simple functions $(\psi_i)_{i=1}^\infty$, $\psi_i : \mathcal{X} \rightarrow \mathcal{Y}^*$, $\forall i \in \mathbb{N}$, such that $\|\psi_i(x)\| \leq \|g(x)\|$, $\forall x \in X$, $\forall i \in \mathbb{N}$, and $\lim_{i \in \mathbb{N}} \psi_i = g$ a.e. in $\mathcal{X}$. Fix any $n \in \mathbb{N}$, let $E_n := \{x \in X_n \mid \|g(x)\| \leq n\} \in \mathcal{B}$ and $\mathcal{E}_n := (E_n, \mathcal{B}_{E_n}, \mu_{E_n})$ be the finite measure subspace of $\mathcal{X}$. Then, $(\mathcal{P} \circ g)|_{E_n}$ and $(\mathcal{P}_q \circ g)|_{E_n}$ are integrable over $\mathcal{E}_n$. By Propositions 7.21, 7.23, 11.52, and 11.53, $\lim_{i \in \mathbb{N}} (\mathcal{P} \circ (\psi_i - g))|_{E_n} = 0$ a.e. in $\mathcal{E}_n$. Note that $(\mathcal{P} \circ (\psi_i - g))|_{E_n}(x) \leq 2(\mathcal{P} \circ g)|_{E_n}(x)$, $\forall x \in E_n$, $\forall i \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91, we have $\lim_{i \in \mathbb{N}} \int_{E_n} \mathcal{P} \circ (\psi_i - g) d\mu = 0$. Note also that $(\mathcal{P}_q \circ \psi_i)|_{E_n}(x) \leq (\mathcal{P}_q \circ g)|_{E_n}(x)$, $\forall x \in E_n$, $\forall i \in \mathbb{N}$. By Propositions 7.21 and 11.52, $\lim_{i \in \mathbb{N}} (\mathcal{P}_q \circ \psi_i)|_{E_n} = (\mathcal{P}_q \circ g)|_{E_n}$ a.e. in $\mathcal{E}_n$. Again, by Lebesgue Dominated Convergence Theorem 11.91, we have $\lim_{i \in \mathbb{N}} \int_{E_n} \mathcal{P}_q \circ \psi_i d\mu = \int_{E_n} \mathcal{P}_q \circ g d\mu$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists i_0 \in \mathbb{N}$ such that $0 \leq \int_{E_n} \mathcal{P} \circ (\psi_{i_0} - g) d\mu < \frac{\epsilon}{2qn^q - 1}$ and $\left| \int_{E_n} \mathcal{P}_q \circ \psi_{i_0} d\mu - \int_{E_n} \mathcal{P}_q \circ g d\mu \right| < \epsilon/2$. Let ψ_{i_0} admit the canonical representation $\psi_{i_0} = \sum_{j=1}^{\bar{n}} y_{*j} \chi_{A_j, X}$. Then,

$$\begin{aligned} \int_{E_n} \mathcal{P}_q \circ g d\mu &< \int_{E_n} \mathcal{P}_q \circ \psi_{i_0} d\mu + \epsilon/2 \\ &= \sum_{\substack{j=1 \\ \mu(A_j \cap E_n) > 0}}^{\bar{n}} \|\nu(A_j \cap E_n)\|^q (\mu(A_j \cap E_n))^{1-q} + \epsilon/2 \\ &\quad - \sum_{\substack{j=1 \\ \mu(A_j \cap E_n) > 0}}^{\bar{n}} \|\nu(A_j \cap E_n)\|^q (\mu(A_j \cap E_n))^{1-q} + \sum_{j=1}^{\bar{n}} \|y_{*j}\|^q \mu(A_j \cap E_n) \end{aligned}$$

Let $\bar{A}_j = A_j \cap E_n$, $j = 1, \dots, \bar{n}$. Then,

$$\begin{aligned} \int_{E_n} \mathcal{P}_q \circ g d\mu &< \sum_{\substack{j=1 \\ \mu(\bar{A}_j) > 0}}^{\bar{n}} \|\nu(\bar{A}_j)\|^q (\mu(\bar{A}_j))^{1-q} \\ &\quad - \sum_{\substack{j=1 \\ \mu(\bar{A}_j) > 0}}^{\bar{n}} \|\nu(\bar{A}_j)\|^q (\mu(\bar{A}_j))^{1-q} + \sum_{\substack{j=1 \\ \mu(\bar{A}_j) > 0}}^{\bar{n}} \|y_{*j}\|^q \mu(\bar{A}_j) + \epsilon/2 \end{aligned}$$

$$\begin{aligned}
&\leq \mathcal{P}_q \circ \nu(E_n) - \sum_{\substack{j=1 \\ \mu(\bar{A}_j) > 0}}^{\bar{n}} \|\nu(\bar{A}_j)\|^q (\mu(\bar{A}_j))^{1-q} \\
&\quad + \sum_{\substack{j=1 \\ \mu(\bar{A}_j) > 0}}^{\bar{n}} \left\| \int_{\bar{A}_j} \psi_{i_0} d\mu \right\|^q (\mu(\bar{A}_j))^{1-q} + \epsilon/2 \\
&\leq M^q + \sum_{\substack{j=1 \\ \mu(\bar{A}_j) > 0}}^{\bar{n}} (\mu(\bar{A}_j))^{1-q} \left(\left\| \int_{\bar{A}_j} \psi_{i_0} d\mu \right\|^q - \left\| \int_{\bar{A}_j} g d\mu \right\|^q \right) + \epsilon/2 \\
&= M^q + \sum_{\substack{j=1 \\ \mu(\bar{A}_j) > 0}}^{\bar{n}} (\mu(\bar{A}_j))^{1-q} \left(\left\| \int_{\bar{A}_j} \psi_{i_0} d\mu \right\| - \left\| \int_{\bar{A}_j} g d\mu \right\| \right) q \\
&\quad \cdot \left(t_j \left\| \int_{\bar{A}_j} \psi_{i_0} d\mu \right\| + (1-t_j) \left\| \int_{\bar{A}_j} g d\mu \right\| \right)^{q-1} + \epsilon/2 \\
&\leq M^q + \sum_{\substack{j=1 \\ \mu(\bar{A}_j) > 0}}^{\bar{n}} (\mu(\bar{A}_j))^{1-q} \left\| \int_{\bar{A}_j} \psi_{i_0} d\mu - \int_{\bar{A}_j} g d\mu \right\| q \\
&\quad \cdot (t_j n \mu(\bar{A}_j) + (1-t_j) n \mu(\bar{A}_j))^{q-1} + \epsilon/2 \\
&= M^q + \sum_{\substack{j=1 \\ \mu(\bar{A}_j) > 0}}^{\bar{n}} q n^{q-1} \left\| \int_{\bar{A}_j} (\psi_{i_0} - g) d\mu \right\| + \epsilon/2 \\
&\leq M^q + \sum_{\substack{j=1 \\ \mu(\bar{A}_j) > 0}}^{\bar{n}} q n^{q-1} \int_{\bar{A}_j} \mathcal{P} \circ (\psi_{i_0} - g) d\mu + \epsilon/2 \\
&\leq M^q + q n^{q-1} \int_{E_n} \mathcal{P} \circ (\psi_{i_0} - g) d\mu + \epsilon/2 < M^q + \epsilon
\end{aligned}$$

where the first equality follows from the Mean Value Theorem 9.20 and $t_j \in (0, 1) \subset \mathbb{R}$, $j = 1, \dots, \bar{n}$; the fourth inequality follows from Proposition 11.92 and the fact that $\|\psi_{i_0}(x)\| \leq \|g(x)\| \leq n$, $\forall x \in E_n$; the second equality and the fifth inequality follow from Proposition 11.92; and the sixth inequality follows from Proposition 11.83. By arbitrariness of ϵ , we have $\int_{E_n} \mathcal{P}_q \circ g d\mu \leq M^q$. Clearly, we have $E_n \subseteq E_{n+1}$, $\forall n \in \mathbb{N}$, and $\bigcup_{n=1}^{\infty} E_n = X$. Then, by Monotone Convergence Theorem 11.81, we have $\int_X \mathcal{P}_q \circ g d\mu = \lim_{n \in \mathbb{N}} \int_X (\mathcal{P}_q \circ g) \chi_{E_n, X} d\mu = \lim_{n \in \mathbb{N}} \int_{E_n} \mathcal{P}_q \circ g d\mu \leq M^q$. Hence, $\|g\|_q \leq M$ and $g \in \bar{L}_q(\mathcal{X}, \mathcal{Y}^*)$.

This completes the proof of the lemma. $\square$

Lemma 11.185 *Let $p \in [1, \infty) \subset \mathbb{R}$, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a separable reflexive Banach space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable, and $\mathcal{Z} := L_p(\mathcal{X}, \mathcal{Y})$ be the Banach space over $\mathbb{K}$ as defined in*

Example 11.179. Then, $\forall f \in \mathcal{Z}^*$, $\exists g : X \rightarrow \mathcal{Y}^*$, which is $\mathcal{B}$ -measurable, such that

(i) $\forall E \in \mathcal{B}$ with $\mu(E) < +\infty$, g is absolutely integrable over $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$, which is the finite measure subspace of $\mathcal{X}$;

(ii) $\forall$ simple function $\phi : \mathcal{X} \rightarrow \mathcal{Y}$ ($\phi \in \bar{\mathcal{L}} := \bar{L}_p(\mathcal{X}, \mathcal{Y})$), the function $\langle\langle g(\cdot), \phi(\cdot) \rangle\rangle : X \rightarrow \mathbb{K}$ is absolutely integrable over $\mathcal{X}$, $f([\phi]) = \int_X \langle\langle g(x), \phi(x) \rangle\rangle d\mu(x)$ and $|\int_X \langle\langle g(x), \phi(x) \rangle\rangle d\mu(x)| \leq \|f\| \|\phi\|_p$.

Furthermore, g is unique in the sense that $\tilde{g} : X \rightarrow \mathcal{Y}^*$ is another function with the above properties if, and only if, $\tilde{g}$ is $\mathcal{B}$ -measurable and $g = \tilde{g}$ a.e. in $\mathcal{X}$.

Proof Fix any $f \in \mathcal{Z}^*$. $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$, $\forall y \in \mathcal{Y}$, let $z_{E,y} := y\chi_{E,X} \in \bar{\mathcal{L}}$. Define $f_E : \mathcal{Y} \rightarrow \mathbb{K}$ by $f_E(y) = f([z_{E,y}])$, $\forall y \in \mathcal{Y}$. Since $f \in \mathcal{Z}^*$, then f_E is linear and continuous and $f_E = y_{*E} \in \mathcal{Y}^*$. Note that $[z_{\emptyset,y}] = \emptyset$, $\forall y \in \mathcal{Y}$, and $f([z_{\emptyset,y}]) = 0$. Then, $f_\emptyset = \emptyset_{\mathcal{Y}^*}$ and $y_{*\emptyset} = \emptyset_{\mathcal{Y}^*}$.

Claim 11.185.1 $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$, $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^\infty E_i$. Then, $\sum_{i=1}^\infty \|y_{*E_i}\| \leq \|f\| (\mu(E))^{1/p} < \infty$.

Proof of claim: $\sum_{i=1}^\infty \|y_{*E_i}\| = \sum_{i=1}^\infty \sup_{y \in \mathcal{Y}, \|y\| \leq 1} |f_{E_i}(y)|$. $\forall i \in \mathbb{N}$, by Propositions 7.85 and 7.90, $\exists y_i \in \mathcal{Y}$ with $\|y_i\| \leq 1$ such that $\|y_{*E_i}\| = f_{E_i}(y_i) = f([z_{E_i,y_i}])$. Then, $\forall n \in \mathbb{N}$, $\sum_{i=1}^n \|y_{*E_i}\| = \sum_{i=1}^n f([z_{E_i,y_i}]) = f([\sum_{i=1}^n z_{E_i,y_i}]) \leq \|f\| \|[\sum_{i=1}^n z_{E_i,y_i}]\|_p = \|f\| \|\sum_{i=1}^n z_{E_i,y_i}\|_p \leq \|f\| \cdot (\mu(E))^{1/p} < \infty$, where the first inequality follows from Proposition 7.72; and the second inequality follows from Propositions 11.83 and 11.75. By the arbitrariness of n , we have $\sum_{n=1}^\infty \|y_{*E_n}\| \leq \|f\| (\mu(E))^{1/p}$. This completes the proof of the claim. $\square$

Claim 11.185.2 $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$, $\forall$ pairwise disjoint $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$ with $E = \bigcup_{n=1}^\infty E_n$, we have $y_{*E} = \sum_{n=1}^\infty y_{*E_n} \in \mathcal{Y}^*$.

Proof of claim: By Claim 11.185.1 and Propositions 7.27 and 7.72, $\sum_{n=1}^\infty y_{*E_n} \in \mathcal{Y}^*$. $\forall y \in \mathcal{Y}$,

$$\begin{aligned} \langle\langle \sum_{n=1}^\infty y_{*E_n}, y \rangle\rangle &= \lim_{n \in \mathbb{N}} \langle\langle \sum_{i=1}^n y_{*E_i}, y \rangle\rangle = \lim_{n \in \mathbb{N}} \sum_{i=1}^n \langle\langle y_{*E_i}, y \rangle\rangle \\ &= \lim_{n \in \mathbb{N}} \sum_{i=1}^n f([z_{E_i,y}]) = \lim_{n \in \mathbb{N}} f([\sum_{i=1}^n z_{E_i,y}]) \end{aligned}$$

where the first equality follows from Propositions 7.72 and 3.66; and the last equality follows from the linearity of f . Note that $\lim_{n \in \mathbb{N}} \sum_{i=1}^n z_{E_i,y}(x) = z_{E,y}(x)$, $\forall x \in X$, and $\|\sum_{i=1}^n z_{E_i,y}(x) - z_{E,y}(x)\|^p \leq \|y\|^p \chi_{E,X}(x)$, $\forall x \in X$,

$\forall n \in \mathbb{N}$. Then, by Lebesgue Dominated Convergence Theorem 11.91, $\lim_{n \in \mathbb{N}} \int_X \mathcal{P}_p \circ (\sum_{i=1}^n z_{E_i, y} - z_{E, y}) d\mu = 0$ and $\lim_{n \in \mathbb{N}} \sum_{i=1}^n z_{E_i, y} \doteq z_{E, y}$ in $\mathcal{Z}$. By Propositions 7.72 and 3.66, $\lim_{n \in \mathbb{N}} f([\sum_{i=1}^n z_{E_i, y}]) = f([z_{E, y}]) = \langle\langle y_{*E}, y \rangle\rangle$. Hence, $\langle\langle y_{*E}, y \rangle\rangle = \langle\langle \sum_{n=1}^\infty y_{*E_n}, y \rangle\rangle$, $\forall y \in \mathcal{Y}$. This implies that, by Proposition 7.85, $y_{*E} = \sum_{n=1}^\infty y_{*E_n}$. $\square$

Since $\mathcal{X}$ is σ -finite, then $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^\infty X_n$ and $\mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $(X_n)_{n=1}^\infty$ is pairwise disjoint. Fix any $n \in \mathbb{N}$, let $\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite measure subspace of $\mathcal{X}$. We may define a function $\nu_n : \mathcal{B}_n \rightarrow \mathcal{Y}^*$ by $\nu_n(E) = f_E = y_{*E}$, $\forall E \in \mathcal{B}_n$. Clearly, $\nu_n(\emptyset) = y_{*\emptyset} = \vartheta_{\mathcal{Y}^*}$. $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}_n$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}_n$. By Claim 11.185.1, $\sum_{i=1}^\infty \|\nu_n(E_i)\| \leq \|f\| (\mu(E))^{1/p} \leq \|f\| (\mu(X_n))^{1/p} < \infty$. By Claim 11.185.2, $\nu_n(E) = \sum_{i=1}^\infty \nu_n(E_i) \in \mathcal{Y}^*$. This shows that ν_n is a $\mathcal{Y}^*$ -valued pre-measure on $(X_n, \mathcal{B}_n)$. By Claim 11.185.1, $\mathcal{P} \circ \nu_n(X_n) \leq \|f\| (\mu(X_n))^{1/p} < +\infty$. Then, ν_n is finite. Hence, $(X_n, \mathcal{B}_n, \nu_n)$ is a finite $\mathcal{Y}^*$ -valued measure space. By Proposition 11.118, the generation process on $((X_n, \mathcal{B}_n, \nu_n))_{n=1}^\infty$ yields a unique σ -finite $\mathcal{Y}^*$ -valued measure space $(X, \mathcal{B}, \nu)$ on X .

Next, we will show that $\mathcal{P} \circ \nu(E) \leq \|f\| (\mu(E))^{1/p} < \infty$ and $\nu(E) = y_{*E}$, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$. Fix any $E \in \mathcal{B}$ with $\mu(E) < \infty$. Let $E_n := X_n \cap E \in \mathcal{B}_n$, $\forall n \in \mathbb{N}$. By Proposition 11.118, $\mathcal{P} \circ \nu(E) = \sum_{n=1}^\infty \mathcal{P} \circ \nu_n(E_n)$. $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\exists m_n \in \mathbb{Z}_+$, $\exists$ pairwise disjoint $(E_{n,i})_{i=1}^{m_n} \subseteq \mathcal{B}_n$ with $E_n = \bigcup_{i=1}^{m_n} E_{n,i}$, such that $\mathcal{P} \circ \nu_n(E_n) < \sum_{i=1}^{m_n} \|\nu_n(E_{n,i})\| + 2^{-n}\epsilon < \infty$. Then,

$$\begin{aligned} \mathcal{P} \circ \nu(E) &< \sum_{n=1}^\infty \left(\sum_{i=1}^{m_n} \|\nu_n(E_{n,i})\| + 2^{-n}\epsilon \right) \\ &\leq \sum_{n=1}^\infty \sum_{i=1}^{m_n} \|y_{*E_{n,i}}\| + \epsilon \leq \|f\| (\mu(E))^{1/p} + \epsilon < \infty \end{aligned}$$

where the third inequality follows from Claim 11.185.1. By the arbitrariness of ϵ , we have $\mathcal{P} \circ \nu(E) \leq \|f\| (\mu(E))^{1/p} < \infty$. Then, $E \in \text{dom}(\nu)$ and $\nu(E) = \sum_{n=1}^\infty \nu_n(E_n) \in \mathcal{Y}^*$. By Claim 11.185.2, we have $\nu(E) = \sum_{n=1}^\infty y_{*E_n} = y_{*E}$. Thus, by Proposition 11.137, ν is uniquely defined independent of the choice of $(X_n)_{n=1}^\infty$.

$\forall E \in \mathcal{B}$ with $\mu(E) = 0$, $\mathcal{P} \circ \nu(E) \leq 0$. Thus, $\mathcal{P} \circ \nu \ll \mu$. By Radon-Nikodym Theorem 11.171 and Proposition 7.90, $\exists! g : \mathcal{X} \rightarrow \mathcal{Y}^*$ such that g is the Radon-Nikodym derivative of ν with respect to μ . $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$, we have $\mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ g d\mu < \infty$. Thus, (i) holds. $\forall$ simple function $\phi : \mathcal{X} \rightarrow \mathcal{Y}$, $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|\phi(x)\| \leq M$, $\forall x \in X$, and $\mu(\hat{E}) := \mu(\{x \in X \mid \phi(x) \neq \vartheta_{\mathcal{Y}}\}) < \infty$. Then, by Proposition 7.72, $|\langle\langle g(x), \phi(x) \rangle\rangle| \leq M \|g(x)\| \chi_{\hat{E}, X}(x) = M \mathcal{P} \circ g(x) \chi_{\hat{E}, X}(x)$, $\forall x \in X$. By Propositions 7.72, 11.38, and 11.39, the function $\langle\langle g(\cdot), \phi(\cdot) \rangle\rangle : X \rightarrow \mathbb{K}$ is $\mathcal{B}$ -measurable. Hence, the function $\langle\langle g(\cdot), \phi(\cdot) \rangle\rangle : X \rightarrow \mathbb{K}$ is absolutely integrable over $\mathcal{X}$. Let ϕ admit the canonical representation $\phi = \sum_{i=1}^n y_i \chi_{E_i, X}$.

Then,

$$\begin{aligned}
 f([\phi]) &= f\left(\left[\sum_{i=1}^n y_i \chi_{E_i, X}\right]\right) = \sum_{i=1}^n f([y_i \chi_{E_i, X}]) = \sum_{i=1}^n f([z_{E_i, y_i}]) \\
 &= \sum_{i=1}^n \langle\langle y_{*E_i}, y_i \rangle\rangle = \sum_{i=1}^n \langle\langle \nu(E_i), y_i \rangle\rangle = \sum_{i=1}^n \left\langle\left\langle \int_{E_i} g \, d\mu, y_i \right\rangle\right\rangle \\
 &= \sum_{i=1}^n \int_{E_i} \langle\langle g(x), y_i \rangle\rangle \, d\mu(x) = \int_X \langle\langle g(x), \phi(x) \rangle\rangle \, d\mu(x) \in \mathbb{K}
 \end{aligned}$$

where the last two equalities follow from Proposition 11.92. Then, $\left|\int_X \langle\langle g(x), \phi(x) \rangle\rangle \, d\mu(x)\right| = |f([\phi])| \leq \|f\| \cdot \|\phi\|_p = \|f\| \|\phi\|_p$. Thus, (ii) holds. Hence, g is the function we seek. Since ν is uniquely defined and g is also uniquely defined, then g is unique as desired.

This completes the proof of the lemma. $\square$

Theorem 11.186 (Riesz Representation Theorem) *Let $p \in [1, \infty) \subset \mathbb{R}$, $q \in (1, \infty] \subset \mathbb{R}_e$ with $1/p + 1/q = 1$, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a separable reflexive Banach space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable, and $\mathcal{Z} := L_p(\mathcal{X}, \mathcal{Y})$ be the Banach space over $\mathbb{K}$ as defined in Example 11.179. Then, $\mathcal{Z}^* = L_q(\mathcal{X}, \mathcal{Y}^*)$ isometrically isomorphically. In particular, $\forall f \in \mathcal{Z}^*, \exists! [g] \in L_q(\mathcal{X}, \mathcal{Y}^*)$ with $g \in \bar{L}_q(\mathcal{X}, \mathcal{Y}^*)$ such that $f([z]) = \langle\langle [g], [z] \rangle\rangle := \int_X \langle\langle g(x), z(x) \rangle\rangle \, d\mu(x)$ and the function $\langle\langle g(\cdot), z(\cdot) \rangle\rangle : X \rightarrow \mathbb{K}$ is absolutely integrable over $\mathcal{X}$, $\forall z \in \bar{\mathcal{Z}} := \bar{L}_p(\mathcal{X}, \mathcal{Y})$, and $\|f\| = \|g\|_q$.*

Proof $\forall f \in \mathcal{Z}^*$, by Lemma 11.185, $\exists g : X \rightarrow \mathcal{Y}^*$, which is $\mathcal{B}$ -measurable, such that

- (i) $\forall E \in \mathcal{B}$ with $\mu(E) < +\infty$, g is absolutely integrable over $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$, which is the finite measure subspace of $\mathcal{X}$;
- (ii) $\forall$ simple function $\phi : \mathcal{X} \rightarrow \mathcal{Y}$ ($\phi \in \bar{L}_p(\mathcal{X}, \mathcal{Y})$), the function $\langle\langle g(\cdot), \phi(\cdot) \rangle\rangle : X \rightarrow \mathbb{K}$ is absolutely integrable over $\mathcal{X}$, $f([\phi]) = \int_X \langle\langle g(x), \phi(x) \rangle\rangle \, d\mu(x)$ and $\left|\int_X \langle\langle g(x), \phi(x) \rangle\rangle \, d\mu(x)\right| \leq \|f\| \|\phi\|_p$.

Furthermore, g is unique in the sense that $\tilde{g} : X \rightarrow \mathcal{Y}^*$ is another function with the above properties if, and only if, $\tilde{g}$ is $\mathcal{B}$ -measurable and $g = \tilde{g}$ a.e. in $\mathcal{X}$.

We will show that $g \in \bar{L}_q(\mathcal{X}, \mathcal{Y}^*)$ and $\|g\|_q \leq \|f\| < +\infty$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $1 < p < +\infty$; Case 2: $p = 1$. Case 1: $1 < p < +\infty$. By Lemma 11.184, we have $g \in \bar{L}_q(\mathcal{X}, \mathcal{Y}^*)$ and $\|g\|_q \leq \|f\| < +\infty$. Case 2: $p = 1$. $\forall E \in \mathcal{B}$ with $0 < \mu(E) < \infty$, by (i) and Proposition 11.92, $y_{*E} := \int_E g \, d\mu \in \mathcal{Y}^*$ and, $\forall y \in \mathcal{Y}$, $|\langle\langle y_{*E}, y \rangle\rangle| = \left|\int_E \langle\langle g(x), y \rangle\rangle \, d\mu(x)\right| = \left|\int_X \langle\langle g(x), y \chi_{E, X}(x) \rangle\rangle \, d\mu(x)\right| \leq \|f\| \|y \chi_{E, X}\|_1 = \|f\| \|y\| \mu(E)$, where

the first two equalities follow from Proposition 11.92; and the inequality follows from (ii). Then, $\|y_{*E}\|/\mu(E) \leq \|f\|$. By Proposition 11.183, we have $\mathcal{P} \circ g \leq \|f\|$ a.e. in $\mathcal{X}$. Hence $\|g\|_\infty \leq \|f\|$ and $g \in \bar{L}_\infty(\mathcal{X}, \mathcal{Y}^*)$.

Hence, in both cases, we have $g \in \bar{L}_q(\mathcal{X}, \mathcal{Y}^*)$ and $\|g\|_q \leq \|f\| < +\infty$. $\forall z \in \tilde{\mathcal{Z}}$, by Proposition 11.181, there exists a sequence of simple functions $(\varphi_i)_{i=1}^\infty$, $\varphi_i : \mathcal{X} \rightarrow \mathcal{Y}$, $\forall i \in \mathbb{N}$, such that $\lim_{i \in \mathbb{N}} \varphi_i = z$ a.e. in $\mathcal{X}$, $\|\varphi_i(x)\| \leq \|z(x)\|$, $\forall x \in X$, $\forall i \in \mathbb{N}$ and $\lim_{i \in \mathbb{N}} \varphi_i = z$ in $\tilde{\mathcal{Z}}$. By Propositions 7.72 and 3.66, $f([z]) = \lim_{i \in \mathbb{N}} f([\varphi_i]) = \lim_{i \in \mathbb{N}} \int_X \langle g(x), \varphi_i(x) \rangle d\mu(x)$. Note that, by Propositions 7.72, 11.52, and 11.53, $\lim_{i \in \mathbb{N}} \langle g(x), \varphi_i(x) \rangle = \langle g(x), z(x) \rangle$ a.e. $x \in \mathcal{X}$ and $|\langle g(x), \varphi_i(x) \rangle| \leq \|g(x)\| \|\varphi_i(x)\| \leq \|g(x)\| \|z(x)\|$, $\forall x \in X$, $\forall i \in \mathbb{N}$. By Hölder's Inequality 11.178, $\int_X \|g(x)\| \cdot \|z(x)\| d\mu(x) \leq \|g\|_q \|z\|_p < \infty$. By Propositions 7.72, 11.38, and 11.39, $\langle g(\cdot), \varphi_i(\cdot) \rangle : X \rightarrow \mathbb{K}$ and $\langle g(\cdot), z(\cdot) \rangle : X \rightarrow \mathbb{K}$ are $\mathcal{B}$ -measurable, $\forall i \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91, $f([z]) = \lim_{i \in \mathbb{N}} \int_X \langle g(x), \varphi_i(x) \rangle d\mu(x) = \int_X \langle g(x), z(x) \rangle d\mu(x)$ and the function $\langle g(\cdot), z(\cdot) \rangle : X \rightarrow \mathbb{K}$ is absolutely integrable over $\mathcal{X}$. Note that

$$\begin{aligned} \|f\| &= \sup_{[z] \in \mathcal{Z}, \|z\|_p \leq 1} |f([z])| = \sup_{z \in \tilde{\mathcal{Z}}, \|z\|_p \leq 1} \left| \int_X \langle g(x), z(x) \rangle d\mu(x) \right| \\ &\leq \sup_{z \in \tilde{\mathcal{Z}}, \|z\|_p \leq 1} \int_X |\langle g(x), z(x) \rangle| d\mu(x) \\ &\leq \sup_{z \in \tilde{\mathcal{Z}}, \|z\|_p \leq 1} \|g\|_q \|z\|_p \leq \|g\|_q \end{aligned}$$

where the first inequality follows from Proposition 11.92; and the second inequality follows from Hölder's Inequality 11.178. Hence, we have $\|f\| = \|g\|_q$.

Thus, we may define a function $\Phi : \mathcal{Z}^* \rightarrow L_q(\mathcal{X}, \mathcal{Y}^*)$ by $\Phi(f) = [g]$, $\forall f \in \mathcal{Z}^*$. Clearly, Φ is well defined and $\|\Phi(f)\|_q = \|f\|$, $\forall f \in \mathcal{Z}^*$. $\forall f_1, f_2 \in \mathcal{Z}^*$, $\forall \alpha, \beta \in \mathbb{K}$, let $[g_i] = \Phi(f_i)$, $i = 1, 2$. $\forall [z] \in \mathcal{Z}$,

$$\begin{aligned} (\alpha f_1 + \beta f_2)([z]) &= \alpha f_1([z]) + \beta f_2([z]) \\ &= \alpha \int_X \langle g_1(x), z(x) \rangle d\mu(x) + \beta \int_X \langle g_2(x), z(x) \rangle d\mu(x) \\ &= \int_X \langle \alpha g_1(x) + \beta g_2(x), z(x) \rangle d\mu(x) \end{aligned}$$

where the last equality follows from Proposition 11.92. By the uniqueness of g , we have $\Phi(\alpha f_1 + \beta f_2) = [\alpha g_1 + \beta g_2] = \alpha [g_1] + \beta [g_2] = \alpha \Phi(f_1) + \beta \Phi(f_2)$. Hence, Φ is linear. $\forall [g] \in L_q(\mathcal{X}, \mathcal{Y}^*)$, we may define $f : \mathcal{Z} \rightarrow \mathbb{K}$ by $f([z]) = \int_X \langle g(x), z(x) \rangle d\mu(x)$, $\forall z \in \tilde{\mathcal{Z}}$. By Hölder's Inequality 11.178 and Proposition 11.92, we have $f \in \mathcal{Z}^*$. Then, $\Phi(f) = [g]$ and Φ is surjective. $\forall f_1, f_2 \in \mathcal{Z}^*$ with $\Phi(f_1) = \Phi(f_2)$, we have $0 = \|\Phi(f_1) - \Phi(f_2)\|_q = \|\Phi(f_1 - f_2)\|_q = \|f_1 - f_2\|$ and $f_1 = f_2$. Hence, Φ is injective. Therefore, $\Phi : \mathcal{Z}^* \rightarrow L_q(\mathcal{X}, \mathcal{Y}^*)$ is a isometrical isomorphism. This completes the proof of the theorem. $\square$

11.10 Dual of $\mathcal{C}(\mathcal{X}, \mathcal{Y})$

Proposition 11.187 *Let $\mathcal{X} := (X, \rho)$ be a metric space and $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ be a metric measure space. Then, the following statements hold.*

- (i) $\forall$ compact subset $K \subseteq \mathcal{X}$, we have $K \in \mathcal{B}_B(\mathcal{X})$.
- (ii) If $\mathbb{X}$ is locally finite and σ -compact, then it is σ -finite.
- (iii) If $\mathbb{X}$ is finite, then it is locally finite.

Proof (i) Fix any compact subset $K \subseteq \mathcal{X}$. We will show that $K = \bigcap_{n=1}^{\infty} \bigcup_{x \in X_n} \mathcal{B}_X(x, 1/n) =: E$, where X_n 's are finite subsets of K . Clearly, we have $E \in \mathcal{B}_B(\mathcal{X})$.

$\forall n \in \mathbb{N}$, $K \subseteq \bigcup_{x \in K} \mathcal{B}_X(x, 1/n)$. Since K is compact, then there exists a finite set $X_n \subseteq K$ such that $K \subseteq \bigcup_{x \in X_n} \mathcal{B}_X(x, 1/n)$. Then, $K \subseteq E$. On the other hand, $\forall \bar{x} \in \tilde{K}$, by Proposition 5.5, K is closed in $\mathcal{X}$. By Proposition 4.10, $\text{dist}(\bar{x}, K) > 0$. Then, $\exists n_0 \in \mathbb{N}$ such that $\text{dist}(\bar{x}, K) > 1/n_0$. Then, $\bar{x} \notin \mathcal{B}_X(x, 1/n_0)$, $\forall x \in K$, which implies that $\bar{x} \notin \bigcup_{x \in X_{n_0}} \mathcal{B}_X(x, 1/n_0)$, which further implies that $\bar{x} \in \tilde{E}$. Hence, $\tilde{K} \subseteq \tilde{E}$ and $E \subseteq K$. Thus, we have $K = E \in \mathcal{B}_B(\mathcal{X})$.

(ii) By $\mathbb{X}$ being σ -compact, there exists $(K_n)_{n=1}^{\infty} \subseteq \mathcal{X}$ such that K_n is compact, $\forall n \in \mathbb{N}$. By $\mathbb{X}$ being locally finite, we have $K_n \in \mathcal{B}_B(\mathcal{X})$ and $\mu(K_n) < \infty$. Then, $\mathbb{X}$ is σ -finite.

(iii) $\forall$ compact subset $K \subseteq \mathcal{X}$, by (i), $K \in \mathcal{B}_B(\mathcal{X})$. Since $\mathbb{X}$ is finite, then $\mu(K) \leq \mu(X) < \infty$. Then, $\mathbb{X}$ is locally finite.

This completes the proof of the proposition. $\square$

Proposition 11.188 *Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $(X_n)_{n=1}^{\infty} \subseteq \mathcal{O}$, $\mathcal{X}_n := (X_n, \mathcal{O}_n)$ be a topological subspace of $\mathcal{X}$, and $(\mathcal{X}_n, \mathcal{B}_B(\mathcal{X}_n), \mu_n) =: \mathbb{X}_n$ be a finite topological measure space, $\forall n \in \mathbb{N}$. Assume that $(\mathbb{X}_n)_{n=1}^{\infty}$ satisfies the assumptions of Proposition 11.118. By Proposition 11.118, the generation process on $(\mathbb{X}_n)_{n=1}^{\infty}$ yields a unique σ -finite measure space $\bar{\mathcal{X}} := (X, \mathcal{B}, \mu)$ such that $\mathbb{X}_n$ is the finite measure subspace of $\bar{\mathcal{X}}$, $\forall n \in \mathbb{N}$. Then, $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ is the unique σ -finite topological measure space on $\mathcal{X}$ such that $\mathbb{X}_n$ is the finite topological measure subspace of $\mathbb{X}$, $\forall n \in \mathbb{N}$.*

Proof $\forall E \in \mathcal{B}_B(\mathcal{X})$, by Proposition 11.25, we have $E \cap X_n \in \mathcal{B}_B(\mathcal{X}_n)$, $\forall n \in \mathbb{N}$. Define $E_{\langle 1 \rangle} := E \cap X_1$ and $E_{\langle n+1 \rangle} := (E \cap X_{n+1}) \setminus (\bigcup_{i=1}^n E_{\langle i \rangle})$, $\forall n \in \mathbb{N}$. Then, $(E_{\langle n \rangle})_{n=1}^{\infty}$ is pairwise disjoint, $\bigcup_{n=1}^{\infty} E_{\langle n \rangle} = E$, and $E_{\langle n \rangle} \in \mathcal{B}_B(\mathcal{X}_n) \subseteq \mathcal{B}_B(\mathcal{X})$, $\forall n \in \mathbb{N}$. Define $\bar{\mu} : \mathcal{B}_B(\mathcal{X}) \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $\bar{\mu}(E) = \sum_{n=1}^{\infty} \mu_n(E_{\langle n \rangle}) \in [0, \infty] \subset \mathbb{R}_e$, $\forall E \in \mathcal{B}_B(\mathcal{X})$.

We will show that $(X, \mathcal{B}_B(\mathcal{X}), \bar{\mu})$ is a σ -finite measure space and $\mathbb{X}_n$ is the finite measure subspace of it, $\forall n \in \mathbb{N}$. It is easy to show that $\bar{\mu}(\emptyset) = 0$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}_B(\mathcal{X})$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}_B(\mathcal{X})$. Note

that $E_{\langle 1 \rangle} = E \cap X_1 = \bigcup_{i=1}^{\infty} (E_i \cap X_1) = \bigcup_{i=1}^{\infty} E_{i, \langle 1 \rangle}$. Inductively, $E_{\langle n+1 \rangle} = (E \cap X_{n+1}) \setminus (\bigcup_{i=1}^n E_{\langle i \rangle}) = (\bigcup_{j=1}^{\infty} (E_j \cap X_{n+1})) \setminus (\bigcup_{i=1}^n \bigcup_{j=1}^{\infty} E_{j, \langle i \rangle}) = \bigcup_{j=1}^{\infty} ((E_j \cap X_{n+1}) \setminus (\bigcup_{i=1}^n E_{j, \langle i \rangle})) = \bigcup_{j=1}^{\infty} E_{j, \langle n+1 \rangle}$, $\forall n \in \mathbb{N}$, where the second equality follows from the inductive assumption; and the third equality follows from the fact that $(E_j)_{j=1}^{\infty}$ is pairwise disjoint and $E_{j, \langle i \rangle} \subseteq E_j$, $\forall i \in \mathbb{N}$, $\forall j \in \mathbb{N}$. Then,

$$\begin{aligned} \bar{\mu}(E) &= \sum_{n=1}^{\infty} \mu_n(E_{\langle n \rangle}) = \sum_{n=1}^{\infty} \mu_n\left(\bigcup_{j=1}^{\infty} E_{j, \langle n \rangle}\right) = \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} \mu_n(E_{j, \langle n \rangle}) \\ &= \sum_{j=1}^{\infty} \sum_{n=1}^{\infty} \mu_n(E_{j, \langle n \rangle}) = \sum_{j=1}^{\infty} \bar{\mu}(E_j) \end{aligned}$$

where the fourth equality follows from the fact that all summands are non-negative real numbers. Hence, $(X, \mathcal{B}_B(\mathcal{X}), \bar{\mu})$ is a measure space. It is σ -finite since $\bar{\mu}(X_n) = \sum_{i=1}^{\infty} \mu_i(X_{n, \langle i \rangle}) = \sum_{i=1}^{\infty} \mu_n(X_{n, \langle i \rangle}) = \mu_n(X_n) < \infty$, $\forall n \in \mathbb{N}$, where the second equality follows from the consistency assumption of Proposition 11.118 on $(\mathbb{X}_n)_{n=1}^{\infty}$.

Fix any $n \in \mathbb{N}$. $\forall E \in \mathcal{B}_B(\mathcal{X}_n) \subseteq \mathcal{B}_B(\mathcal{X})$, $\bar{\mu}(E) = \sum_{i=1}^{\infty} \mu_i(E_{\langle i \rangle}) = \sum_{i=1}^{\infty} \mu_n(E_{\langle i \rangle}) = \mu_n(\bigcup_{i=1}^{\infty} E_{\langle i \rangle}) = \mu_n(E)$, where the second equality follows from the consistency assumption of Proposition 11.118 on $(\mathbb{X}_n)_{n=1}^{\infty}$. By Propositions 11.25 and 11.13, we have $\mathbb{X}_n$ is the finite measure subspace of $(X, \mathcal{B}_B(\mathcal{X}), \bar{\mu})$, $\forall n \in \mathbb{N}$. By Proposition 11.118, we have $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$ and $\mu = \bar{\mu}$. Then, $\mathbb{X} = (X, \mathcal{B}_B(\mathcal{X}), \bar{\mu})$.

Finally, we will show that $\mathbb{X}$ is a topological measure space. Fix any $E \in \mathcal{B}_B(\mathcal{X})$. Then, $E = \bigcup_{n=1}^{\infty} E_{\langle n \rangle}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, since $\mathbb{X}_n$ is a topological measure space and $E_{\langle n \rangle} \in \mathcal{B}_B(\mathcal{X}_n)$, $\exists O_n \in \mathcal{O}_n$ with $E_{\langle n \rangle} \subseteq O_n$ such that $\mu_n(O_n \setminus E_{\langle n \rangle}) < 2^{-n}\epsilon$. Since $\mathcal{X}_n$ is a topological subspace of $\mathcal{X}$ and $X_n \in \mathcal{O}$, then $O_n \in \mathcal{O}_n \subseteq \mathcal{O} \subseteq \mathcal{B}_B(\mathcal{X})$. Let $O := \bigcup_{n=1}^{\infty} O_n \in \mathcal{O}$. Then, $E = \bigcup_{n=1}^{\infty} E_{\langle n \rangle} \subseteq \bigcup_{n=1}^{\infty} O_n = O$ and

$$\begin{aligned} \mu(O \setminus E) &= \bar{\mu}(O \setminus E) = \bar{\mu}\left(\left(\bigcup_{n=1}^{\infty} O_n\right) \setminus \left(\bigcup_{n=1}^{\infty} E_{\langle n \rangle}\right)\right) \\ &= \bar{\mu}\left(\bigcup_{n=1}^{\infty} (O_n \cap (\bigcap_{j=1}^{\infty} \widetilde{E_{\langle j \rangle}}))\right) \leq \bar{\mu}\left(\bigcup_{n=1}^{\infty} (O_n \cap \widetilde{E_{\langle n \rangle}})\right) \\ &= \bar{\mu}\left(\bigcup_{n=1}^{\infty} (O_n \setminus E_{\langle n \rangle})\right) \leq \sum_{n=1}^{\infty} \bar{\mu}(O_n \setminus E_{\langle n \rangle}) < \epsilon \end{aligned}$$

Hence, $\mathbb{X}$ is a topological measure space. By Proposition 11.29, $\mathbb{X}_n$ is the finite topological measure subspace of $\mathbb{X}$, $\forall n \in \mathbb{N}$. This completes the proof of the proposition. $\square$

Proposition 11.189 *Let $\mathcal{X} := (X, \rho)$ be a metric space with the natural topology $\mathcal{O}$, $(X_n)_{n=1}^{\infty} \subseteq \mathcal{O}$, $\mathcal{X}_n := (X_n, \rho)$ be a metric subspace of $\mathcal{X}$,*

and $\mathbb{X}_n := (\mathcal{X}_n, \mathcal{B}_B(\mathcal{X}_n), \mu_n)$ be a finite metric measure space, $\forall n \in \mathbb{N}$. Assume that $(\mathbb{X}_n)_{n=1}^\infty$ satisfies the assumptions of Proposition 11.118. By Proposition 11.188, the generation process on $(\mathbb{X}_n)_{n=1}^\infty$ yields a unique σ -finite metric measure space $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ such that $\mathbb{X}_n$ is the finite metric measure subspace of $\mathbb{X}$, $\forall n \in \mathbb{N}$. Then, $\mathbb{X}$ is locally finite.

Proof Fix any compact subset $K \subseteq \mathcal{X}$. By Proposition 11.187, $K \in \mathcal{B}_B(\mathcal{X})$. Clearly, $K \subseteq \bigcup_{n=1}^\infty X_n = X$. By the compactness of K , $\exists N \in \mathbb{N}$ such that $K \subseteq \bigcup_{n=1}^N X_n$. Then, $\mu(K) \leq \sum_{n=1}^N \mu(X_n) = \sum_{n=1}^N \mu_n(X_n) < \infty$, where the first inequality follows from Proposition 11.6; and the equality follows from Proposition 11.13. Hence, $\mathbb{X}$ is locally finite. $\square$

Proposition 11.190 Let $p \in [1, \infty) \subset \mathbb{R}$, $\mathcal{X}_1$ be a finite-dimensional Banach space, $\mathcal{X}_2$ be a σ -compact metric space, $\mathcal{Y}_1$ be a finite-dimensional Banach space over $\mathbb{K}$, $\mathcal{Y}_2$ be a separable normed linear space over $\mathbb{K}$, $X \subseteq \mathcal{X}_1$ be a subset with subset topology $\mathcal{O}$, $\mathcal{X} := (X, \mathcal{O})$, and $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ be a metric measure space. Then, the following statements hold.

- (i) $\mathcal{Y}_1$ is separable.
- (ii) $l_p(\mathcal{Y}_2) =: \mathcal{Z}_2$ is a separable normed linear space.
- (iii) If $\mathcal{X}$ is compact, then $\mathcal{C}(\mathcal{X}, \mathcal{Y}_1)$ is a separable Banach space.
- (iv) If $\mathbb{X}$ is a locally finite σ -compact metric measure space, then $L_p(\mathbb{X}, \mathcal{Y}_1)$ is a separable Banach space.
- (v) $\mathcal{X}_2$ is separable.

Proof Let $\mathcal{Y}_1$ be m -dimensional, where $m \in \mathbb{Z}_+$, and $\{e_1, \dots, e_m\} \subseteq \mathcal{Y}_1$ be a set of basis vectors. Then, $\mathcal{Y}_1$ is isomorphic to $\mathbb{K}^m$. The norm $\|\cdot\|_{\mathcal{Y}_1}$ induces a norm $\|\cdot\|_1$ on $\mathbb{K}^m$ as defined by, $\forall a := (\alpha_1, \dots, \alpha_m) \in \mathbb{K}^m$, $\|a\|_1 := \|\sum_{i=1}^m \alpha_i e_i\|_{\mathcal{Y}_1} \in \mathbb{R}$. Then, $\mathcal{Y}_1$ is isometrically isomorphic to the normed linear space $(\mathbb{K}^m, \mathbb{K}, \|\cdot\|_1)$, which is a Banach space by Theorem 7.36. By Theorem 7.38, $\|\cdot\|_1$ is equivalent to the Euclidean norm $|\cdot|$ on $\mathbb{K}^m$.

(i) Let $\mathbb{K}_{\mathbb{Q}} := \mathbb{Q}$ if $\mathbb{K} = \mathbb{R}$; and $\mathbb{K}_{\mathbb{Q}} := \{a + ib \in \mathbb{C} \mid a, b \in \mathbb{Q}\}$, if $\mathbb{K} = \mathbb{C}$. Clearly, $\mathbb{K}_{\mathbb{Q}}$ is a countable dense subset of $\mathbb{K}$. Then, $\mathbb{K}_{\mathbb{Q}}^m$ is a countable dense subset of $(\mathbb{K}^m, \mathbb{K}, |\cdot|)$. Since $\|\cdot\|_1$ is equivalent to $|\cdot|$, then $\mathbb{K}_{\mathbb{Q}}^m$ is a countable dense subset of $(\mathbb{K}^m, \mathbb{K}, \|\cdot\|_1)$. Hence, $\mathcal{Y}_1$ is separable, since $(\mathbb{K}^m, \mathbb{K}, \|\cdot\|_1)$ is separable.

(ii) Let $D \subseteq \mathcal{Y}_2$ be a countable denset subset and, $\forall i \in \mathbb{N}$, $D_i := \{(y_1, \dots, y_i, \vartheta_{y_2}, \vartheta_{y_2}, \dots) \in \mathcal{Z}_2 \mid y_j \in D, 1 \leq j \leq i\}$. Then, D_i is countable, $\forall i \in \mathbb{N}$. Let $\bar{D} := \bigcup_{i=1}^\infty D_i \subseteq \mathcal{Z}_2$, which is also countable. We will show that $\bar{D}$ is dense in $\mathcal{Z}_2 = l_p(\mathcal{Y}_2)$. $\forall z := (y_1, y_2, \dots) \in \mathcal{Z}_2$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, we have $\sum_{i=1}^\infty \|y_i\|_{\mathcal{Y}_2}^p < \infty$. Then, $\exists N \in \mathbb{N}$ such that $\sum_{i=N+1}^\infty \|y_i\|_{\mathcal{Y}_2}^p < \epsilon^p/2$. $\forall i \in \{1, \dots, N\}$, $\exists \bar{y}_i \in D$ such that $\|y_i - \bar{y}_i\|_{\mathcal{Y}_2} <$

$\frac{\epsilon}{2^{1/p}N^{1/p}}$. Let $\bar{z} := (\bar{y}_1, \dots, \bar{y}_N, \vartheta_{y_2}, \vartheta_{y_2}, \dots) \in \bar{D}$. Then, $\|z - \bar{z}\|_p = \left(\sum_{i=1}^N \|y_i - \bar{y}_i\|_{y_2}^p + \sum_{i=N+1}^\infty \|y_i\|_{y_2}^p \right)^{1/p} < \left(\sum_{i=1}^N \frac{\epsilon^p}{2^p} + \epsilon^p/2 \right)^{1/p} = \epsilon$. Hence, $\bar{D}$ is dense in Z_2 . Therefore, Z_2 is separable.

(iii) This result is standard, which follows directly from Corollary 7.57 of Stone-Weierstrass Theorem, and therefore the proof is omitted.

(iv) First, consider the special case where $\mathbb{X}$ is a finite compact metric measure space. By (iii), let $D \subseteq \mathcal{C}(\mathcal{X}, \mathcal{Y}_1)$ be a countable dense subset. By Proposition 4.11, $\mathcal{X}$ is a normal topological space. Then, by Proposition 11.182 and (i), $\forall f \in \bar{L}_p(\mathbb{X}, \mathcal{Y}_1)$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists g \in \bar{L}_p(\mathbb{X}, \mathcal{Y}_1)$ such that g is continuous and $\|g - f\|_p < \epsilon/2$. Clearly, $g \in \mathcal{C}(\mathcal{X}, \mathcal{Y}_1)$. Then, $\exists h \in D$ such that $\|h - g\|_{\mathcal{C}(\mathcal{X}, \mathcal{Y}_1)} < \frac{\epsilon}{2(\mu(X))^{1/p+1}}$. Since X is compact and $h \in \mathcal{C}(\mathcal{X}, \mathcal{Y}_1)$, then $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\mathcal{P} \circ h(x) \leq M < \infty$, $\forall x \in X$. By Proposition 11.37, h is $\mathcal{B}$ -measurable. Hence, $h \in \bar{L}_p(\mathbb{X}, \mathcal{Y}_1)$. Then, $\|f - h\|_p \leq \|f - g\|_p + \|g - h\|_p < \epsilon/2 + \left(\int_X \mathcal{P}_p \circ (g - h) d\mu \right)^{1/p} \leq \epsilon/2 + \left(\|g - h\|_{\mathcal{C}(\mathcal{X}, \mathcal{Y}_1)}^p \mu(X) \right)^{1/p} < \epsilon$. Hence, $D \subseteq \bar{L}_p(\mathbb{X}, \mathcal{Y}_1)$ is dense. Therefore, $L_p(\mathbb{X}, \mathcal{Y}_1)$ is separable. By Example 11.179, $L_p(\mathbb{X}, \mathcal{Y}_1)$ is a Banach space. Hence, $L_p(\mathbb{X}, \mathcal{Y}_1)$ is a separable Banach space.

Now, consider the general case where $\mathbb{X}$ is a locally finite σ -compact metric measure space. Then, $\exists (X_n)_{n=1}^\infty \subseteq \mathbb{X}$ such that $\mathbb{X} = \bigcup_{n=1}^\infty X_n$ and X_n is compact, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $X_n \subseteq X_{n+1}$, $\forall n \in \mathbb{N}$. Since $\mathbb{X}$ is locally finite, then $X_n \in \mathcal{B}_B(\mathcal{X})$ and $\mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Fix any $n \in \mathbb{N}$. By Proposition 4.37, let $\mathcal{X}_n$ be the metric subspace of $\mathcal{X}$. Let $\mathbb{X}_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite measure subspace of $\mathbb{X}$. By Proposition 11.29, $\mathcal{B}_n = \mathcal{B}_B(\mathcal{X}_n)$ and $\mathbb{X}_n$ is a finite compact metric measure space. By the special case, $L_p(\mathbb{X}_n, \mathcal{Y}_1)$ is a separable Banach space. Let $D_n \subseteq L_p(\mathbb{X}_n, \mathcal{Y}_1)$ be a countable dense subset. $\forall f \in \bar{L}_p(\mathbb{X}_n, \mathcal{Y}_1)$, define $\bar{f} \in \bar{L}_p(\mathbb{X}, \mathcal{Y}_1)$ by $\bar{f}(x) = \begin{cases} f(x) & \forall x \in X_n \\ \vartheta_{y_1} & \forall x \in X \setminus X_n \end{cases}$, $\forall x \in \mathbb{X}$. Let $D := \{ [\bar{f}] \in L_p(\mathbb{X}, \mathcal{Y}_1) \mid \exists n \in \mathbb{N}, [f] \in D_n \}$. Clearly, $D \subseteq L_p(\mathbb{X}, \mathcal{Y}_1)$ is countable.

We will show that D is dense in $L_p(\mathbb{X}, \mathcal{Y}_1)$. Fix any $f \in \bar{L}_p(\mathbb{X}, \mathcal{Y}_1)$. We have $\int_X (\mathcal{P}_p \circ f) d\mu < \infty$. By Monotone Convergence Theorem 11.81, $\int_X (\mathcal{P}_p \circ f) d\mu = \lim_{n \in \mathbb{N}} \int_X \mathcal{P}_p \circ (f \chi_{X_n, X}) d\mu$. Let $f_n := f \chi_{X_n, X} \in L_p(\mathbb{X}, \mathcal{Y}_1)$, $\forall n \in \mathbb{N}$. Then,

$$\begin{aligned} \lim_{n \in \mathbb{N}} \|f_n - f\|_p &= \lim_{n \in \mathbb{N}} \left(\int_X \mathcal{P}_p \circ (f_n - f) d\mu \right)^{1/p} \\ &= \lim_{n \in \mathbb{N}} \left(\int_X \mathcal{P}_p \circ (f \chi_{X \setminus X_n, X}) d\mu \right)^{1/p} \\ &= \lim_{n \in \mathbb{N}} \left(\int_X (\mathcal{P}_p \circ f - \mathcal{P}_p \circ f_n) d\mu \right)^{1/p} \\ &= \lim_{n \in \mathbb{N}} \left(\int_X \mathcal{P}_p \circ f d\mu - \int_X \mathcal{P}_p \circ f_n d\mu \right)^{1/p} = 0 \end{aligned}$$

where the fourth equality follows from Proposition 11.83. Hence, $\lim_{n \in \mathbb{N}} f_n \doteq f$ in $\bar{L}_p(\mathbb{X}, \mathcal{Y}_1)$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\|f_{n_0} - f\|_p < \epsilon/2$.

Let $\hat{f}_{n_0} := f_{n_0}|_{X_{n_0}}$. Then, $\hat{f}_{n_0} \in \bar{L}_p(\mathbb{X}_{n_0}, \mathcal{Y}_1)$. This implies that $\exists [g] \in D_{n_0}$ with $g \in \bar{L}_p(\mathbb{X}_{n_0}, \mathcal{Y}_1)$ such that $\|\hat{f}_{n_0} - g\|_p < \epsilon/2$. Let $\bar{g} \in \bar{L}_p(\mathbb{X}, \mathcal{Y}_1)$ be defined as in the second paragraph from last. Then, $[\bar{g}] \in D$ and

$$\begin{aligned} \|f - \bar{g}\|_p &\leq \|f - f_{n_0}\|_p + \|f_{n_0} - \bar{g}\|_p \\ &< \epsilon/2 + \left(\int_X \mathcal{P}_p \circ (f_{n_0} - \bar{g}) d\mu \right)^{1/p} \\ &= \epsilon/2 + \left(\int_X \mathcal{P}_p \circ ((f_{n_0} - \bar{g})\chi_{X_{n_0}, X}) d\mu \right)^{1/p} \\ &= \epsilon/2 + \left(\int_{X_{n_0}} \mathcal{P}_p \circ (\hat{f}_{n_0} - g) d\mu_{n_0} \right)^{1/p} \\ &= \epsilon/2 + \|\hat{f}_{n_0} - g\|_p < \epsilon \end{aligned}$$

where the second equality follows from Proposition 11.83. Hence, D is dense in $L_p(\mathbb{X}, \mathcal{Y}_1)$. Then, by Example 11.179 and Proposition 11.187, $L_p(\mathbb{X}, \mathcal{Y}_1)$ is a separable Banach space.

(v) Since $\mathcal{X}_2$ is σ -compact, then $\exists$ compact sets $(K_n)_{n=1}^\infty$ such that $\mathcal{X}_2 = \bigcup_{n=1}^\infty K_n$. $\forall n \in \mathbb{N}$, $\forall i \in \mathbb{N}$, $K_n \subseteq \bigcup_{x \in K_n} \mathcal{B}_{\mathcal{X}_2}(x, 1/i)$. By the compactness of K_n , $\exists$ finite subset $D_{n,i} \subseteq K_n$ such that $K_n \subseteq \bigcup_{x \in D_{n,i}} \mathcal{B}_{\mathcal{X}_2}(x, 1/i)$. Then, $D := \bigcup_{n=1}^\infty \bigcup_{i=1}^\infty D_{n,i} \subseteq \mathcal{X}_2$ is a countable dense set. Hence, $\mathcal{X}_2$ is separable.

This completes the proof of the proposition. $\square$

Definition 11.191 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $\mathcal{Y}$ be a normed linear space, and $(X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space on the same set X . The triple $\mathbb{X} := (\mathcal{X}, \mathcal{B}, \mu)$ is said to be a $\mathcal{Y}$ -valued topological measure space if $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$ and $\bar{\mathbb{X}} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mathcal{P} \circ \mu)$ is a topological measure space. We will say that $\mathbb{X}$ is finite, σ -finite, or locally finite if $\bar{\mathbb{X}}$ is so. We will say that $\mathbb{X}$ is Tychonoff, Hausdorff, regular, completely regular, normal, first countable, second countable, separable, second category everywhere, connected, locally connected, compact, countably compact, sequentially compact, locally compact, σ -compact, or paracompact if $\mathcal{X}$ is so.

Let $\mathcal{X} := (X, \rho)$ be a metric space with the natural topology $\mathcal{O}$, $\mathcal{Y}$ be a normed linear space, and $(X, \mathcal{B}_B(\mathcal{X}), \mu)$ be a $\mathcal{Y}$ -valued measure space on the same set X . The triple $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ is said to be a $\mathcal{Y}$ -valued metric measure space if $((X, \mathcal{O}), \mathcal{B}_B(\mathcal{X}), \mu)$ is a $\mathcal{Y}$ -valued topological measure space. $\mathbb{X}$ is said to be complete or totally bounded if $\mathcal{X}$ is so.

Let $\mathcal{X} := (\mathcal{X}, \mathbb{K}, \|\cdot\|)$ be a normed linear space over the field $\mathbb{K}$, $\mathcal{O}$ be the natural topology on $\mathcal{X}$ generated by the norm $\|\cdot\|$, $\mathcal{Y}$ be a normed linear space, and $(X, \mathcal{B}_B(\mathcal{X}), \mu)$ be a $\mathcal{Y}$ -valued measure space on the same set X . The triple $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ is said to be a $\mathcal{Y}$ -valued normed linear measure

space if $((X, \mathcal{O}), \mathcal{B}_B(\mathcal{X}), \mu)$ is a $\mathcal{Y}$ -valued topological measure space. When $\mathcal{X}$ is a Banach space, then $\mathbb{X}$ is said to be a $\mathcal{Y}$ -valued Banach measure space. Depending on whether $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, we will say that $\mathbb{X}$ is a $\mathcal{Y}$ -valued real or complex Banach measure space.

Proposition 11.192 *Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$. Define $\bar{\mathcal{Z}} := \{\mu \in \mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y}) \mid (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu) \text{ is a finite } \mathcal{Y}\text{-valued topological measure space}\}$. Then, $\bar{\mathcal{Z}}$ is a closed subspace of $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$ and $(\bar{\mathcal{Z}}, \mathbb{K}, \|\cdot\|_{\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})}) =: \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$ is a normed linear space.*

If, in addition, $\mathcal{Y}$ is a Banach space. Then, $\mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$ is a Banach space. Furthermore, define $\mathcal{Z} := \{\mu \in \mathcal{M}_\sigma(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y}) \mid (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu) \text{ is a } \sigma\text{-finite } \mathcal{Y}\text{-valued topological measure space}\} \subseteq \mathcal{M}_\sigma(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$. Let $\mathcal{Z}$ admit the subset topology $\mathcal{O}_{\mathcal{Z}}$. Then, $\mathcal{Z} =: \mathcal{M}_{\sigma t}(\mathcal{X}, \mathcal{Y})$ is a subspace of $\mathcal{M}_\sigma(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$. We will abuse the notation and denote the topological space $(\mathcal{Z}, \mathcal{O}_{\mathcal{Z}})$ by $\mathcal{M}_{\sigma t}(\mathcal{X}, \mathcal{Y})$.

Proof Let $\mathcal{Y}$ be a normed linear space. We will show that $\bar{\mathcal{Z}}$ is a subspace of $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$. $\forall \alpha_1, \alpha_2 \in \mathbb{K}, \forall \mu_1, \mu_2 \in \mathcal{Z}$, by Proposition 11.136, $\mu := \alpha_1 \mu_1 + \alpha_2 \mu_2 \in \mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$. $\forall E \in \mathcal{B}_B(\mathcal{X})$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall i = 1, 2$, by $\mu_i \in \mathcal{Z}$, $\exists O_i \in \mathcal{O}$ with $E \subseteq O_i$ such that $\mathcal{P} \circ \mu_i(O_i \setminus E) < \epsilon/(1 + 2|\alpha_i|)$. Let $O := O_1 \cap O_2 \in \mathcal{O}$. Clearly, $E \subseteq O$ and $\mathcal{P} \circ \mu(O \setminus E) \leq (\sum_{i=1}^2 |\alpha_i| \mathcal{P} \circ \mu_i)(O \setminus E) = \sum_{i=1}^2 |\alpha_i| \mathcal{P} \circ \mu_i(O \setminus E) \leq \sum_{i=1}^2 |\alpha_i| \mathcal{P} \circ \mu_i(O_i \setminus E) < \epsilon$, where the first inequality and the first equality follow from Proposition 11.136. This shows that $(\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mathcal{P} \circ \mu)$ is a topological measure space. Then, $(\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ is a finite $\mathcal{Y}$ -valued topological measure space and $\mu \in \bar{\mathcal{Z}}$. Clearly, $\vartheta_{\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})} \in \bar{\mathcal{Z}} \neq \emptyset$. Hence, $\bar{\mathcal{Z}}$ is a subspace of $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$. Then, $\mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$ is a normed linear space since $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$ is a normed linear space.

Next, we will show that $\bar{\mathcal{Z}}$ is closed. $\forall \mu \in \bar{\mathcal{Z}} \subseteq \mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$, by Proposition 4.13, $\exists (\mu_n)_{n=1}^\infty \subseteq \bar{\mathcal{Z}}$ such that $\lim_{n \in \mathbb{N}} \mu_n = \mu$. Then, $\lim_{n \in \mathbb{N}} \|\mu_n - \mu\|_{\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})} = \lim_{n \in \mathbb{N}} \mathcal{P} \circ (\mu_n - \mu)(X) = 0$. $\forall E \in \mathcal{B}_B(\mathcal{X})$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists n \in \mathbb{N}$ such that $\mathcal{P} \circ (\mu_n - \mu)(X) < \epsilon/2$. By $\mu_n \in \bar{\mathcal{Z}}$, $\exists O \in \mathcal{O}$ with $E \subseteq O$ such that $\mathcal{P} \circ \mu_n(O \setminus E) < \epsilon/2$. Then, $\mathcal{P} \circ \mu(O \setminus E) = \mathcal{P} \circ (\mu_n - \mu_n + \mu)(O \setminus E) \leq (\mathcal{P} \circ \mu_n + \mathcal{P} \circ (\mu_n - \mu))(O \setminus E) = \mathcal{P} \circ \mu_n(O \setminus E) + \mathcal{P} \circ (\mu_n - \mu)(O \setminus E) < \epsilon/2 + \mathcal{P} \circ (\mu_n - \mu)(X) < \epsilon$, where the first equality and the first inequality follow from Proposition 11.136. Hence, $\mu \in \bar{\mathcal{Z}}$. By the arbitrariness of μ , we have $\bar{\bar{\mathcal{Z}}} = \bar{\mathcal{Z}}$ and $\bar{\mathcal{Z}}$ is closed.

Let $\mathcal{Y}$ be a Banach space. By Proposition 11.142, $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$ is a Banach space. By Proposition 4.39, $\mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$ is a Banach space.

Finally, we will show that $\mathcal{Z}$ is a subspace of $\mathcal{M}_\sigma(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$. $\forall \alpha_1, \alpha_2 \in \mathbb{K}, \forall \mu_1, \mu_2 \in \mathcal{Z}$, by Proposition 11.138, $\mu := \alpha_1 \mu_1 + \alpha_2 \mu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$. $\forall E \in \mathcal{B}_B(\mathcal{X})$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall i = 1, 2$, by $\mu_i \in \mathcal{Z}$, $\exists O_i \in \mathcal{O}$ with $E \subseteq O_i$ such that $\mathcal{P} \circ \mu_i(O_i \setminus E) < \epsilon/(1 + 2|\alpha_i|)$. Let $O := O_1 \cap O_2 \in \mathcal{O}$. Clearly, $E \subseteq O$ and $\mathcal{P} \circ \mu(O \setminus E) \leq (\sum_{i=1}^2 |\alpha_i| \mathcal{P} \circ \mu_i)(O \setminus E) < \epsilon$.

$\mu_i)(O \setminus E) = \sum_{i=1}^2 |\alpha_i| \mathcal{P} \circ \mu_i(O \setminus E) \leq \sum_{i=1}^2 |\alpha_i| \mathcal{P} \circ \mu_i(O_i \setminus E) < \epsilon$, where the first inequality and the first equality follow from Proposition 11.138. This shows that $(\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mathcal{P} \circ \mu)$ is a topological measure space. Then, $(\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ is a σ -finite $\mathcal{Y}$ -valued topological measure space and $\mu \in \mathcal{Z}$. Clearly, $\vartheta_{\mathcal{M}_\sigma(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})} \in \mathcal{Z} \neq \emptyset$. Hence, $\mathcal{Z}$ is a subspace of $\mathcal{M}_\sigma(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$.

This completes the proof of the proposition. $\square$

A bit of notation to simplify our presentation. Let $\mathcal{M}_\sigma(X, \mathcal{B})$ denote the set of σ -finite measures on the measurable space $(X, \mathcal{B})$; $\mathcal{M}_f(X, \mathcal{B})$ denote the set of finite measures on the measurable space $(X, \mathcal{B})$; $\mathcal{M}_{\sigma t}(\mathcal{X})$ denote the set of σ -finite topological measures on the topological space $\mathcal{X}$; and $\mathcal{M}_{ft}(\mathcal{X})$ denote the set of finite topological measures on the topological space $\mathcal{X}$.

Proposition 11.193 *Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $\mu_o : \mathcal{O} \rightarrow [0, \infty) \subset \mathbb{R}$. Assume that*

- (i) $\mu_o(\emptyset) = 0$;
- (ii) $\mu_o(O_1) \leq \mu_o(O_2)$, $\forall O_1, O_2 \in \mathcal{O}$ with $O_1 \subseteq O_2$;
- (iii) $\mu_o(\bigcup_{i=1}^\infty O_i) \leq \sum_{i=1}^\infty \mu_o(O_i)$, $\forall (O_i)_{i=1}^\infty \subseteq \mathcal{O}$;
- (iv) $\mu_o(O_1 \cup O_2) = \mu_o(O_1) + \mu_o(O_2)$, $\forall O_1, O_2 \in \mathcal{O}$ with $O_1 \cap O_2 = \emptyset$;
- (v) $\mu_o(O) = \sup_{U \in \mathcal{O}, U \subseteq \overline{U} \subseteq O} \mu_o(U)$, $\forall O \in \mathcal{O}$.

Define $\bar{\mu}_o : {}^X 2 \rightarrow [0, \infty) \subset \mathbb{R}$ by $\bar{\mu}_o(E) = \inf_{O \in \mathcal{O}, E \subseteq O} \mu_o(O)$, $\forall E \subseteq X$. Then, the following statements hold.

1. $\bar{\mu}_o$ is an outer measure. It induces a finite complete measure space $(X, \bar{\mathcal{B}}, \bar{\mu})$, where $\bar{\mathcal{B}} := \{E \subseteq X \mid E \text{ is measurable with respect to } \bar{\mu}_o\}$ and $\bar{\mu} := \bar{\mu}_o|_{\bar{\mathcal{B}}}$.
2. $\mathcal{B}_B(\mathcal{X}) \subseteq \bar{\mathcal{B}}$ and the triple $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu := \bar{\mu}|_{\mathcal{B}_B(\mathcal{X})})$ is a finite topological measure space with $\mu(O) = \mu_o(O)$, $\forall O \in \mathcal{O}$.
3. The measure μ is unique in the sense that if $\hat{\mu}$ be another measure on $(X, \mathcal{B}_B(\mathcal{X}))$ satisfying $\hat{\mu}(O) = \mu_o(O) = \mu(O)$, $\forall O \in \mathcal{O}$, then $\hat{\mu} = \mu$.

Proof 1. By (i), we have $\bar{\mu}_o(\emptyset) = 0$. $\forall A \subseteq B \subseteq X$, we have $0 \leq \bar{\mu}_o(A) = \inf_{O \in \mathcal{O}, A \subseteq O} \mu_o(O) \leq \inf_{O \in \mathcal{O}, B \subseteq O} \mu_o(O) = \bar{\mu}_o(B) \leq \mu_o(X) < \infty$. $\forall E \subseteq \bigcup_{i=1}^\infty E_i \subseteq X$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, $\exists O_i \in \mathcal{O}$ with $E_i \subseteq O_i$ such that $\mu_o(O_i) < \bar{\mu}_o(E_i) + 2^{-i}\epsilon$. Then, $\bar{\mu}_o(E) = \inf_{O \in \mathcal{O}, E \subseteq O} \mu_o(O) \leq \mu_o(\bigcup_{i=1}^\infty O_i) \leq \sum_{i=1}^\infty \mu_o(O_i) < \sum_{i=1}^\infty \bar{\mu}_o(E_i) + \epsilon$, where the second inequality follows from (iii). By the arbitrariness of ϵ , we have $\bar{\mu}_o(E) \leq \sum_{i=1}^\infty \bar{\mu}_o(E_i)$. Hence, $\bar{\mu}_o : {}^X 2 \rightarrow [0, \infty) \subset \mathbb{R}$ is an outer measure. It is easy to see that $\bar{\mu}_o(O) = \mu_o(O)$, $\forall O \in \mathcal{O}$. By Theorem 11.17, $(X, \bar{\mathcal{B}}, \bar{\mu})$ is a finite complete measure space.

2. $\forall O \in \mathcal{O}, \forall E \subseteq X, \forall \epsilon \in (0, +\infty) \subset \mathbb{R}, \exists O_1 \in \mathcal{O}$ with $E \subseteq O_1$ such that $\mu_o(O_1) < \bar{\mu}_o(E) + \epsilon/2$. By (v), $\exists U \in \mathcal{O}$ with $U \subseteq \bar{U} \subseteq O_1 \cap O$ such that $\mu_o(O \cap O_1) < \mu_o(U) + \epsilon/2$. Then, $\bar{\mu}_o(E) > \mu_o(O_1) - \epsilon/2 \geq \mu_o((O_1 \setminus \bar{U}) \cup U) - \epsilon/2 = \mu_o(O_1 \setminus \bar{U}) + \mu_o(U) - \epsilon/2 = \bar{\mu}_o(O_1 \setminus \bar{U}) + \mu_o(U) - \epsilon/2 > \bar{\mu}_o(O_1 \setminus (O_1 \cap O)) + \mu_o(O \cap O_1) - \epsilon = \bar{\mu}_o(O_1 \setminus O) + \bar{\mu}_o(O \cap O_1) - \epsilon \geq \bar{\mu}_o(E \setminus O) + \bar{\mu}_o(E \cap O) - \epsilon \geq \bar{\mu}_o(E) - \epsilon$, where the second inequality follows from (ii); the first equality follows from (iv); and the third, the fourth, and the last inequalities follow from the fact that $\bar{\mu}_o$ is an outer measure. By the arbitrariness of ϵ , we have $\bar{\mu}_o(E) = \bar{\mu}_o(E \setminus O) + \bar{\mu}_o(E \cap O)$. By the arbitrariness of E , O is measurable with respect to $\bar{\mu}_o$ and $O \in \bar{\mathcal{B}}$. By the arbitrariness of O , we have $\mathcal{O} \subseteq \bar{\mathcal{B}}$. Since $\bar{\mathcal{B}}$ is a σ -algebra on X , then $\mathcal{B}_B(\mathcal{X}) \subseteq \bar{\mathcal{B}}$. By Proposition 11.13, $(X, \mathcal{B}_B(\mathcal{X}), \mu)$ is a measure space. Clearly, $\mu(O) = \bar{\mu}(O) = \bar{\mu}_o(O) = \mu_o(O), \forall O \in \mathcal{O}$. This coupled with the definition of $\bar{\mu}_o$ leads to the conclusion that $\mathbb{X}$ is a topological measure space. Clearly, $\mu(X) = \mu_o(X) < \infty$. Hence, $\mathbb{X}$ is a finite topological measure space.

3. Let $\hat{\mu}$ be another measure on $(X, \mathcal{B}_B(\mathcal{X}))$ satisfying $\hat{\mu}(O) = \mu_o(O) = \mu(O), \forall O \in \mathcal{O}$. Then, $\hat{\mu}$ is finite. Suppose $\mu \neq \hat{\mu}$. Then, $\exists E \in \mathcal{B}_B(\mathcal{X})$ such that $\hat{\mu}(E) \neq \mu(E) = \inf_{O \in \mathcal{O}, E \subseteq O} \mu_o(O) = \inf_{O \in \mathcal{O}, E \subseteq O} \mu(O) = \inf_{O \in \mathcal{O}, E \subseteq O} \hat{\mu}(O)$. Since $\hat{\mu}(E) \leq \hat{\mu}(O), \forall O \in \mathcal{O}$ with $E \subseteq O$. Then, we must have $\hat{\mu}(E) < \mu(E)$. Since $\mathbb{X}$ is a topological measure space, then, by Proposition 11.27, $\exists X \setminus F \in \mathcal{O}$ with $F \subseteq E$ such that $\mu(E \setminus F) < (\mu(E) - \hat{\mu}(E))/2$. This implies that $\mu(F) = \mu(E) - \mu(E \setminus F) > \mu(E)/2 + \hat{\mu}(E)/2 > \hat{\mu}(E) \geq \hat{\mu}(F)$ and $\mu_o(X \setminus F) = \mu(X \setminus F) = \mu(X) - \mu(F) < \hat{\mu}(X) - \hat{\mu}(F) = \hat{\mu}(X \setminus F) = \mu_o(X \setminus F)$, where the first equality follows from $X \setminus F \in \mathcal{O}$; the second equality follows from μ being a finite measure; the inequality follows from $X \in \mathcal{O}$; the third equality follows from $\hat{\mu}$ being a measure; and the last equality follows from $X \setminus F \in \mathcal{O}$. This is a contradiction. Hence, we have $\hat{\mu} = \mu$ and μ is unique.

This completes the proof of the proposition. $\square$

Proposition 11.194 *Let $\mathcal{X} := (X, \mathcal{O})$ be a compact Hausdorff topological space, $\mathcal{Y}$ be a Banach space, $\bar{\mu}$ be a function that assigns a vector $\bar{\mu}(F) \in \mathcal{Y}$ for each closed subset $F \subseteq X$, $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \nu)$ be a finite topological measure space. Assume that*

- (i) $\bar{\mu}(F_1 \cup F_2) = \bar{\mu}(F_1) + \bar{\mu}(F_2), \forall X \setminus F_1, X \setminus F_2 \in \mathcal{O}$ with $F_1 \cap F_2 = \emptyset$;
- (ii) $\|\bar{\mu}(F_1) - \bar{\mu}(F_2)\| \leq \nu(O), \forall X \setminus F_1, X \setminus F_2, O \in \mathcal{O}$ with $F_1 \triangle F_2 \subseteq O$.

Then, there exists a unique $\mu \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$ such that $\mu(F) = \bar{\mu}(F), \forall X \setminus F \in \mathcal{O}$. Furthermore, $\mathcal{P} \circ \mu \leq \nu$.

Proof Fix any $E \in \mathcal{B}_B(\mathcal{X})$. Let $\hat{\mathcal{A}}_E := \{F \subseteq E \mid X \setminus F \in \mathcal{O}\}$ and $\tilde{\mathcal{A}}_E := (\hat{\mathcal{A}}_E, \subseteq)$. Clearly, $\tilde{\mathcal{A}}_E$ is a directed system.

Claim 11.194.1 *The net $(\bar{\mu}(F))_{F \in \tilde{\mathcal{A}}_E}$ is Cauchy.*

Proof of claim: $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, by $\mathbb{X}$ being a topological measure space, $\exists V \in \mathcal{O}$ with $E \subseteq V$ such that $\nu(V \setminus E) < \epsilon/2$. By Proposition 11.27, $\exists X \setminus F \in \mathcal{O}$ with $F \subseteq E$ such that $\nu(E \setminus F) < \epsilon/2$. $\forall F_1 \in \tilde{\mathcal{A}}_E$ with $F \subseteq F_1$, we have $F \subseteq F_1 \subseteq E \subseteq V$ and $F \triangle F_1 = F_1 \setminus F \subseteq V \setminus F \in \mathcal{O}$. By (ii), $\|\bar{\mu}(F_1) - \bar{\mu}(F)\| \leq \nu(V \setminus F) = \nu(V \setminus E) + \nu(E \setminus F) < \epsilon$. Hence, the net is Cauchy. This completes the proof of the claim. $\square$

By Proposition 4.44, we may define $\mu(E) = \lim_{F \in \tilde{\mathcal{A}}_E} \bar{\mu}(F) \in \mathcal{Y}$. Thus, we have defined a function $\mu : \mathcal{B}_B(\mathcal{X}) \rightarrow \mathcal{Y}$. We will show that μ is the $\mathcal{Y}$ -valued measure we seek. Clearly, $\mu(F) = \bar{\mu}(F)$, $\forall X \setminus F \in \mathcal{O}$. By (i), $\bar{\mu}(\emptyset) + \bar{\mu}(\emptyset) = \bar{\mu}(\emptyset)$ and then $\mu(\emptyset) = \bar{\mu}(\emptyset) = \vartheta_{\mathcal{Y}}$.

$\forall E \in \mathcal{B}_B(\mathcal{X})$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists V \in \mathcal{O}$ with $E \subseteq V$ such that $\nu(V \setminus E) < \epsilon/2$. By $\mu(E) = \lim_{F \in \tilde{\mathcal{A}}_E} \bar{\mu}(F)$, $\exists F \in \tilde{\mathcal{A}}_E$ such that $\|\mu(E) - \bar{\mu}(F)\| < \epsilon/2$. Clearly, $F \subseteq E \subseteq V$. Then, $\|\mu(E)\| < \|\bar{\mu}(F)\| + \epsilon/2 \leq \nu(V) + \epsilon/2 = \nu(E) + \nu(V \setminus E) + \epsilon/2 < \nu(E) + \epsilon$, where the second inequality follows from (ii). By the arbitrariness of ϵ , we have $\|\mu(E)\| \leq \nu(E)$.

Fix any pairwise disjoint $(E_n)_{n=1}^\infty \subseteq \mathcal{B}_B(\mathcal{X})$, let $E := \bigcup_{n=1}^\infty E_n \in \mathcal{B}_B(\mathcal{X})$. $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\exists V_n \in \mathcal{O}$ with $E_n \subseteq V_n$ such that $\nu(V_n \setminus E_n) < 2^{-n-1}\epsilon/5$. By Proposition 11.27, $\exists X \setminus F_n \in \mathcal{O}$ with $F_n \subseteq E_n$ such that $\nu(E_n \setminus F_n) < 2^{-n-1}\epsilon/5$. By $\mu(E_n) = \lim_{F \in \tilde{\mathcal{A}}_{E_n}} \bar{\mu}(F)$, $\exists \hat{F}_n \in \tilde{\mathcal{A}}_{E_n}$ with $F_n \subseteq \hat{F}_n$ such that $\|\mu(E_n) - \bar{\mu}(\hat{F}_n)\| < 2^{-n}\epsilon/5$. Clearly, $F_n \subseteq \hat{F}_n \subseteq E_n \subseteq V_n$. $\exists V \in \mathcal{O}$ with $E \subseteq V$ such that $\nu(V \setminus E) < \epsilon/5$. By Proposition 11.27, $\exists X \setminus F \in \mathcal{O}$ with $F \subseteq E$ such that $\nu(E \setminus F) < \epsilon/5$. By $\mu(E) = \lim_{F \in \tilde{\mathcal{A}}_E} \bar{\mu}(F)$, $\exists \hat{F} \in \tilde{\mathcal{A}}_E$ with $F \subseteq \hat{F}$ such that $\|\mu(E) - \bar{\mu}(\hat{F})\| < \epsilon/5$. Clearly, $F \subseteq \hat{F} \subseteq E \subseteq V$. By Proposition 5.5, $\hat{F}$ is compact. Since $\hat{F} \subseteq E = \bigcup_{n=1}^\infty E_n \subseteq \bigcup_{n=1}^\infty V_n$, then $\exists n_0 \in \mathbb{N}$ such that $\hat{F} \subseteq \bigcup_{i=1}^{n_0} V_i$. $\forall n \in \mathbb{N}$ with $n_0 \leq n$, $\|\mu(E) - \sum_{i=1}^n \mu(E_i)\| \leq \|\mu(E) - \bar{\mu}(\hat{F})\| + \|\bar{\mu}(\hat{F}) - \sum_{i=1}^n \bar{\mu}(\hat{F}_i)\| + \sum_{i=1}^n \|\bar{\mu}(\hat{F}_i) - \mu(E_i)\| < \epsilon/5 + \|\bar{\mu}(\hat{F}) - \bar{\mu}(\bigcup_{i=1}^n \hat{F}_i)\| + \epsilon/5$, where the second inequality follows from (i). Clearly, $\hat{F}$ and $\bigcup_{i=1}^n \hat{F}_i$ are closed sets. $\hat{F} \triangle (\bigcup_{i=1}^n \hat{F}_i) = (\hat{F} \setminus (\bigcup_{i=1}^n \hat{F}_i)) \cup ((\bigcup_{i=1}^n \hat{F}_i) \setminus \hat{F}) \subseteq ((\bigcup_{i=1}^n V_i) \setminus (\bigcup_{i=1}^n \hat{F}_i)) \cup (E \setminus \hat{F}) \subseteq (\bigcup_{i=1}^n (V_i \setminus \hat{F}_i)) \cup (V \setminus F) \in \mathcal{O}$. By (ii), $\|\mu(E) - \sum_{i=1}^n \mu(E_i)\| < 2\epsilon/5 + \nu((\bigcup_{i=1}^n (V_i \setminus \hat{F}_i)) \cup (V \setminus F)) \leq 2\epsilon/5 + \sum_{i=1}^n \nu(V_i \setminus \hat{F}_i) + \nu(V \setminus F) = 2\epsilon/5 + \sum_{i=1}^n (\nu(V_i \setminus E_i) + \nu(E_i \setminus \hat{F}_i)) + \nu(V \setminus E) + \nu(E \setminus F) < \epsilon$. Then, $\mu(E) = \sum_{n=1}^\infty \mu(E_n) \in \mathcal{Y}$.

$\forall$ pairwise disjoint $(E_n)_{n=1}^\infty \subseteq \mathcal{B}_B(\mathcal{X})$, $\sum_{n=1}^\infty \|\mu(E_n)\| \leq \sum_{n=1}^\infty \nu(E_n) = \nu(\bigcup_{n=1}^\infty E_n) \leq \nu(X) < \infty$. This shows that μ is a $\mathcal{Y}$ -valued pre-measure on $(X, \mathcal{B}_B(\mathcal{X}))$.

$\forall E \in \mathcal{B}_B(\mathcal{X})$, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_i)_{i=1}^n \subseteq \mathcal{B}_B(\mathcal{X})$ with $E = \bigcup_{i=1}^n E_i$, $\sum_{i=1}^n \|\mu(E_i)\| \leq \sum_{i=1}^n \nu(E_i) = \nu(E)$. Hence, $\mathcal{P} \circ \mu(E) \leq \nu(E)$. By the arbitrariness of E , we have $\mathcal{P} \circ \mu \leq \nu$. Hence, μ is a finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B}_B(\mathcal{X}))$. It is easy to show that $(\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ is a $\mathcal{Y}$ -valued topological measure space. Then, $\mu \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$.

Finally, we need to show that $\mu \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$ is unique. Let $\hat{\mu} \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$ be such that: $\hat{\mu}(F) = \bar{\mu}(F) = \mu(F)$, $\forall X \setminus F \in \mathcal{O}$. $\forall E \in \mathcal{B}_B(\mathcal{X})$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, by Proposition 11.27, $\exists X \setminus F \in \mathcal{O}$ with $F \subseteq E$ such that $\mathcal{P} \circ \mu(E \setminus F) < \epsilon/2$. Again by Proposition 11.27, $\exists X \setminus F_1 \in \mathcal{O}$ with $F_1 \subseteq E$ such that $\mathcal{P} \circ \hat{\mu}(E \setminus F) < \epsilon/2$. Let $\bar{F} := F \cup F_1 \subseteq E$, which is clearly closed. Then, $\|\mu(E) - \hat{\mu}(E)\| \leq \|\mu(E) - \mu(\bar{F})\| + \|\mu(\bar{F}) - \hat{\mu}(\bar{F})\| + \|\hat{\mu}(\bar{F}) - \hat{\mu}(E)\| = \|\mu(E \setminus \bar{F})\| + 0 + \|\hat{\mu}(E \setminus \bar{F})\| \leq \mathcal{P} \circ \mu(E \setminus \bar{F}) + \mathcal{P} \circ \hat{\mu}(E \setminus \bar{F}) \leq \mathcal{P} \circ \mu(E \setminus F) + \mathcal{P} \circ \hat{\mu}(E \setminus F_1) < \epsilon$. By the arbitrariness of ϵ , we have $\mu(E) = \hat{\mu}(E)$. Hence, $\mu = \hat{\mu}$.

This completes the proof of the proposition. $\square$

Definition 11.195 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space. $E \subseteq X$ is said to be a G_δ if $\exists (O_i)_{i=1}^\infty \subseteq \mathcal{O}$ such that $E = \bigcap_{i=1}^\infty O_i$. $\bar{E} \subseteq X$ is said to be an F_σ if exists $(\tilde{F}_i)_{i=1}^\infty \subseteq \mathcal{O}$ such that $\bar{E} = \bigcup_{i=1}^\infty \tilde{F}_i$.

Proposition 11.196 Let $\mathcal{X} := (X, \rho)$ be a locally compact separable metric space with the natural topology $\mathcal{O}$, $\mathcal{E} := \{E \subseteq X \mid E \text{ is a compact } G_\delta\}$, and $\mathcal{B}_a(\mathcal{X})$ be the σ -algebra generated by $\mathcal{E}$. Then, $\mathcal{B}_a(\mathcal{X}) = \mathcal{B}_B(\mathcal{X})$.

Proof $\forall E \in \mathcal{E}$, by Proposition 5.5, E is closed. Then, $E \in \mathcal{B}_B(\mathcal{X})$. Hence, $\mathcal{E} \subseteq \mathcal{B}_B(\mathcal{X})$. Since $\mathcal{B}_B(\mathcal{X})$ is a σ -algebra, then $\mathcal{B}_a(\mathcal{X}) \subseteq \mathcal{B}_B(\mathcal{X})$.

$\forall x \in X$, by Definition 5.49, $\exists O_x \in \mathcal{O}$ such that $x \in O_x$ and $\overline{O_x}$ is compact. Then, $X = \bigcup_{x \in X} O_x$. By Propositions 4.4 and 3.24, $\exists$ a countable set $D \subseteq X$ such that $X = \bigcup_{x \in D} O_x = \bigcup_{x \in D} \overline{O_x}$.

$\forall \tilde{F} \in \mathcal{O}$, $\forall x \in D$, $F_x := \tilde{F} \cap \overline{O_x}$ is compact and closed by Propositions 5.5 and 3.5. Then, $F_x = \bigcap_{n=1}^\infty \bigcup_{\bar{x} \in F_x} \mathcal{B}_X(\bar{x}, 1/n)$, by Proposition 4.10. This implies that F_x is a G_δ in addition to being compact. Then, $F_x \in \mathcal{E} \subseteq \mathcal{B}_a(\mathcal{X})$. This implies that $F = \bigcup_{x \in D} F_x \in \mathcal{B}_a(\mathcal{X})$. Hence, $\tilde{F} \in \mathcal{B}_a(\mathcal{X})$. By the arbitrariness of $\tilde{F}$, we have $\mathcal{O} \subseteq \mathcal{B}_a(\mathcal{X})$. By $\mathcal{B}_a(\mathcal{X})$ being a σ -algebra, we have $\mathcal{B}_B(\mathcal{X}) \subseteq \mathcal{B}_a(\mathcal{X})$.

Therefore, $\mathcal{B}_a(\mathcal{X}) = \mathcal{B}_B(\mathcal{X})$. This completes the proof of the proposition. $\square$

Lemma 11.197 Let $\mathcal{X} := (X, \mathcal{O})$ be a normal topological space. Then, $\forall \tilde{E} \in \mathcal{O}$ with E being a G_δ , there exists a continuous function $\phi : X \rightarrow [0, 1] \subset \mathbb{R}$ such that $E = \{x \in X \mid \phi(x) = 1\}$.

Proof Since E is a G_δ , then $\exists (O_i)_{i=1}^\infty \subseteq \mathcal{O}$ such that $E = \bigcap_{i=1}^\infty O_i$. $\forall i \in \mathbb{N}$, by Urysohn's Lemma 3.55, there exists a continuous function $\phi_i : X \rightarrow [0, 1] \subset \mathbb{R}$ such that $\phi_i|_E = 1$ and $\phi_i|_{\widetilde{O_i}} = 0$. Let $\phi := \sum_{i=1}^\infty 2^{-i} \phi_i$. By Proposition 4.26, $\phi : X \rightarrow [0, 1] \subset \mathbb{R}$ is continuous. Clearly $\phi|_E = 1$ and $\phi(x) < 1$, $\forall x \in \bigcup_{i=1}^\infty \widetilde{O_i} = \widetilde{E}$. Hence, the result holds. This completes the proof of the lemma. $\square$

Theorem 11.198 Let $\mathcal{X} := (X, \rho)$ be a locally compact separable metric space with the natural topology $\mathcal{O}$ and μ be a finite measure on $(X, \mathcal{B}_B(\mathcal{X}))$. Then, $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ is a finite metric measure space. Thus, $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X})) = \mathcal{M}_{ft}(\mathcal{X})$. As a consequence, $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y}) = \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$, where $\mathcal{Y}$ is any normed linear space.

Proof By Propositions 4.4, 3.24, and 5.72, $\mathcal{X}$ is σ -compact. Let $\mathcal{R} \subseteq \mathcal{B}_B(\mathcal{X})$ be such that $E \in \mathcal{R}$ if $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, we have

- (i) $\exists O \in \mathcal{O}$ with O being σ -compact and $E \subseteq O$ such that $\mu(O \setminus E) < \epsilon$;
 - (ii) $\exists \tilde{F} \in \mathcal{O}$ with F being a compact G_δ and $F \subseteq \tilde{F}$ such that $\mu(\tilde{F} \setminus F) < \epsilon$.
- We will show that $\mathcal{R} = \mathcal{B}_B(\mathcal{X})$.

Claim 11.198.1 $\forall \tilde{F} \in \mathcal{O}$, $F = \bigcap_{n=1}^{\infty} \bigcup_{x \in F} \mathcal{B}_X(x, 1/n) =: \bar{F}$ and is a G_δ .

Proof of claim: Clearly, $F \subseteq \bar{F}$. $\forall x_0 \in \tilde{F}$, by Proposition 4.10, $\epsilon_0 := \text{dist}(x_0, \tilde{F}) > 0$. Then, $\forall n \in \mathbb{N}$ with $1/n < \epsilon_0$, $x_0 \notin \bigcup_{x \in F} \mathcal{B}_X(x, 1/n)$ and $x_0 \in \bar{F}$. Hence, $\tilde{F} \subseteq \bar{F}$ and $F = \bar{F}$. Clearly, $\bar{F}$ is a G_δ . This completes the proof of the claim. $\square$

Claim 11.198.2 $\forall O \in \mathcal{O}$, O is σ -compact.

Proof of claim: By Claim 11.198.1, $\tilde{O}$ is a G_δ . Then, $\exists (O_i)_{i=1}^{\infty} \subseteq \mathcal{O}$ such that $\tilde{O} = \bigcap_{i=1}^{\infty} O_i$ and $O = \bigcup_{i=1}^{\infty} \tilde{O}_i$. By $\mathcal{X}$ being σ -compact, $\exists$ compact sets $(K_i)_{i=1}^{\infty}$ such that $X = \bigcup_{i=1}^{\infty} K_i$. Then, $O = O \cap X = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} (\tilde{O}_i \cap K_j)$. $\forall i, j \in \mathbb{N}$, by Propositions 3.5 and 5.5, $\tilde{O}_i \cap K_j$ is compact. Hence, O is σ -compact. $\square$

$\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{R}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}_B(\mathcal{X})$. $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, $\exists O_i \in \mathcal{O}$ with O_i being σ -compact and $E_i \subseteq O_i$ such that $\mu(O_i \setminus E_i) < 2^{-i}\epsilon$; $\exists \tilde{F}_i \in \mathcal{O}$ with F_i being a compact G_δ and $F_i \subseteq \tilde{F}_i$ such that $\mu(\tilde{F}_i \setminus F_i) < 2^{-i-1}\epsilon$. Let $O := \bigcup_{i=1}^{\infty} O_i \in \mathcal{O}$, which is σ -compact by Claim 11.198.2. Then, $E \subseteq O$ and $\mu(O \setminus E) = \mu((\bigcup_{i=1}^{\infty} O_i) \setminus (\bigcup_{i=1}^{\infty} E_i)) \leq \mu(\bigcup_{i=1}^{\infty} (O_i \setminus E_i)) \leq \sum_{i=1}^{\infty} \mu(O_i \setminus E_i) < \epsilon$. Hence, E satisfies (i). By Proposition 11.7, $\mu(E) = \lim_{n \in \mathbb{N}} \mu(\bigcup_{i=1}^n E_i)$. Then, $\exists n_0 \in \mathbb{N}$ such that $\mu(E) < \mu(\bigcup_{i=1}^{n_0} E_i) + \epsilon/2$. Let $F := \bigcup_{i=1}^{n_0} F_i$. Clearly, F is compact and closed. By Claim 11.198.1, F is a G_δ . Then, $F \subseteq E$ and $\mu(E \setminus F) = \mu(E) - \mu(F) < \mu(\bigcup_{i=1}^{n_0} E_i) - \mu(\bigcup_{i=1}^{n_0} F_i) + \epsilon/2 = \mu((\bigcup_{i=1}^{n_0} E_i) \setminus (\bigcup_{i=1}^{n_0} F_i)) + \epsilon/2 \leq \mu(\bigcup_{i=1}^{n_0} (E_i \setminus F_i)) + \epsilon/2 \leq \sum_{i=1}^{n_0} \mu(E_i \setminus F_i) + \epsilon/2 < \epsilon$. Hence, E satisfies (ii). Then, $E \in \mathcal{R}$. Thus, $\mathcal{R}$ is closed under countable unions.

$\forall E \in \mathcal{R}$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists O \in \mathcal{O}$ with O being σ -compact and $E \subseteq O$ such that $\mu(O \setminus E) < \epsilon$; $\exists \tilde{F} \in \mathcal{O}$ with F being a compact G_δ and $F \subseteq \tilde{F}$ such that $\mu(\tilde{F} \setminus F) < \epsilon$. Then, $\tilde{O}$ is a closed set, $\tilde{O} \subseteq \tilde{E}$, and $\mu(\tilde{E} \setminus \tilde{O}) = \mu(O \setminus E) < \epsilon$. By $\mathcal{X}$ being σ -compact, $\exists$ compact sets $(K_i)_{i=1}^{\infty}$ such that $X = \bigcup_{i=1}^{\infty} K_i$. Without loss of generality, we may assume that $K_i \subseteq K_{i+1}$, $\forall i \in \mathbb{N}$. Note that $\epsilon > \mu(\tilde{E} \setminus \tilde{O}) = \mu(\tilde{E} \setminus (\tilde{O} \cap \bigcup_{i=1}^{\infty} K_i)) = \mu(\tilde{E} \setminus (\bigcup_{i=1}^{\infty} (\tilde{O} \cap K_i))) = \mu(\bigcap_{i=1}^{\infty} (\tilde{E} \setminus (\tilde{O} \cap K_i))) = \lim_{i \in \mathbb{N}} \mu(\tilde{E} \setminus (\tilde{O} \cap K_i))$,

where the last equality follows from Proposition 11.5. Then, $\exists n \in \mathbb{N}$ such that $\mu(\tilde{E} \setminus (\tilde{O} \cap K_n)) < \epsilon$. By Propositions 3.5 and 5.5, $\tilde{O} \cap K_n$ is closed and compact, then it is a G_δ by Claim 11.198.1. Clearly, $\tilde{O} \cap K_n \subseteq \tilde{E}$. Hence, $\tilde{E}$ satisfies (ii). $\tilde{F} \in \mathcal{O}$ and is σ -compact by Claim 11.198.2. Note that $\tilde{E} \subseteq \tilde{F}$ and $\mu(\tilde{F} \setminus \tilde{E}) = \mu(E \setminus F) < \epsilon$. Then, $\tilde{E}$ satisfies (i). Hence, $\tilde{E} \in \mathcal{R}$. $\mathcal{R}$ is closed under set complements. Clearly, $\emptyset \in \mathcal{R}$. This proves that $\mathcal{R}$ is a σ -algebra.

$\forall \tilde{E} \in \mathcal{O}$ with E being a compact G_δ , $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, then E satisfies (ii) trivially. By Lemma 11.197, there exists a continuous function $\phi : X \rightarrow [0, 1] \subset \mathbb{R}$ such that $E = \{x \in X \mid \phi(x) = 1\}$. $\forall i \in \mathbb{N}$, let $O_i := \{x \in X \mid \phi(x) > 1 - 1/i\} \in \mathcal{O}$. Clearly $E = \bigcap_{i=1}^{\infty} O_i$ and $O_{i+1} \subseteq O_i$, $\forall i \in \mathbb{N}$. By Proposition 11.5, $\mu(E) = \lim_{i \in \mathbb{N}} \mu(O_i)$. $\exists n \in \mathbb{N}$ such that $\mu(O_n) - \epsilon \leq \mu(E) \leq \mu(O_n)$. Clearly, we have $E \subseteq O_n$. Then, $\mu(O_n \setminus E) = \mu(O_n) - \mu(E) < \epsilon$. By Claim 11.198.2, O_n is σ -compact. Then, E satisfies (i) and $E \in \mathcal{R}$.

Thus, $\mathcal{R}$ is a σ -algebra and $\mathcal{E} := \{E \in \mathcal{B}_B(\mathcal{X}) \mid E \text{ is a compact } G_\delta\} \subseteq \mathcal{R}$. By Proposition 11.196, $\mathcal{R} = \mathcal{B}_B(\mathcal{X})$. Then, $\mathbb{X}$ is a finite metric measure space.

Thus, we have $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X})) \subseteq \mathcal{M}_{ft}(\mathcal{X})$. Clearly, $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X})) \supseteq \mathcal{M}_{ft}(\mathcal{X})$. Hence, $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X})) = \mathcal{M}_{ft}(\mathcal{X})$.

$\forall \bar{\mu} \in \mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y})$ then $\mathcal{P} \circ \bar{\mu} \in \mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X})) = \mathcal{M}_{ft}(\mathcal{X})$. This implies that $\bar{\mu} \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$. Hence, $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y}) \subseteq \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$. Clearly, $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y}) \supseteq \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$. Then, $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y}) = \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y})$.

This completes the proof of the theorem. $\square$

Definition 11.199 Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $\mathcal{Y}$ be a normed linear space, $f : X \rightarrow \mathcal{Y}$. The support of f is the set $\text{supp}(f) := \{x \in X \mid f(x) \neq \vartheta_{\mathcal{Y}}\}$.

Lemma 11.200 Let $\mathcal{X} := (X, \mathcal{O})$ be compact Hausdorff topological space, $\mathbb{Z} := \mathcal{C}(\mathcal{X}, \mathbb{R})$, and $f \in \mathbb{Z}^*$. $\bar{f} \in \mathbb{Z}^*$ is said to be a positive linear functional if $\bar{f}(z) \geq 0$, $\forall z \in P := \{h \in \mathbb{Z} \mid h : X \rightarrow [0, \infty) \subset \mathbb{R}\}$. Then, $f = f_+ - f_-$, where $f_+, f_- \in \mathbb{Z}^*$ are positive linear functionals.

Proof $\forall z \in P$, define $f_+(z) := \sup_{\phi \in \mathbb{Z}, 0 \leq \phi(x) \leq z(x), \forall x \in X} f(\phi)$. Then, $0 \leq f_+(z) \leq \|f\| \|z\| < \infty$ and $f_+(z) \geq f(z)$, $\forall z \in P$. Clearly,

(i) $f_+(\alpha z) = \alpha f_+(z)$, $\forall z \in P$ and $\forall \alpha \in [0, \infty) \subset \mathbb{R}$.

$\forall z_1, z_2 \in P$, $\forall i \in \{1, 2\}$, $\forall \phi_i \in \mathbb{Z}$ with $0 \leq \phi_i(x) \leq z_i(x)$, $\forall x \in X$, we have $f(\phi_1) + f(\phi_2) = f(\phi_1 + \phi_2) \leq f_+(z_1 + z_2)$. Then, $f_+(z_1) + f_+(z_2) \leq f_+(z_1 + z_2)$. On the other hand, $\forall \phi \in \mathbb{Z}$ with $0 \leq \phi(x) \leq z_1(x) + z_2(x)$, $\forall x \in X$, we have $f(\phi) = f(\phi \wedge z_1) + f(\phi - \phi \wedge z_1) \leq f_+(z_1) + f_+(z_2)$. Then, $f_+(z_1 + z_2) \leq f_+(z_1) + f_+(z_2)$. Therefore,

(ii) $f_+(z_1) + f_+(z_2) = f_+(z_1 + z_2)$, $\forall z_1, z_2 \in P$.

$\forall z \in \mathcal{Z}$, define $z_+ := z \vee 0 \in P$ and $z_- := (-z) \vee 0 \in P$. Clearly, $z = z_+ - z_-$. Define $f_+(z) := f_+(z_+) - f_+(z_-) \in \mathbb{R}$. Then, $f_+ : \mathcal{Z} \rightarrow \mathbb{R}$ is well-defined. We will show that $f_+ \in \mathcal{Z}^*$.

$\forall z \in \mathcal{Z}$, $|f_+(z)| = |f_+(z_+) - f_+(z_-)| \leq f_+(z_+) + f_+(z_-) = f_+(z_+ + z_-) \leq \|f\| \|z_+ + z_-\| = \|f\| \|z\|$.

$\forall z_1, z_2 \in \mathcal{Z}$, let $z := z_1 + z_2$. Then, $f_+(z_1) + f_+(z_2) = f_+(z_{1+}) - f_+(z_{1-}) + f_+(z_{2+}) - f_+(z_{2-}) = f_+(z_{1+} + z_{2+}) - f_+(z_{1-} + z_{2-}) = f_+(z_+) + f_+(z_{1+} + z_{2+} - z_+) - f_+(z_-) - f_+(z_{1-} + z_{2-} - z_-) = f_+(z) + f_+(z_{1+} + z_{2+} - z_+) - f_+(z_{1-} + z_{2-} - z_-)$, where the first equality follows from the definition of f_+ ; the second equality follows from (ii); the third equality follows from (ii); and the last equality follows from the definition of f_+ . Note that, $\forall x \in X$,

$$\begin{aligned} (z_{1+} + z_{2+} - z_+)(x) &= z_1(x) \vee 0 + z_2(x) \vee 0 - (z_1(x) + z_2(x)) \vee 0 \\ &= \begin{cases} 0 & z_1(x) \geq 0 \text{ and } z_2(x) \geq 0 \\ -z_2(x) & z_1(x) \geq 0 > z_2(x) \geq -z_1(x) \\ z_1(x) & z_1(x) \geq 0 \geq -z_1(x) > z_2(x) \\ -z_1(x) & z_2(x) \geq 0 > z_1(x) \geq -z_2(x) \\ z_2(x) & z_2(x) \geq 0 \geq -z_2(x) > z_1(x) \\ 0 & z_1(x) < 0 \text{ and } z_2(x) < 0 \end{cases} \\ &= (-z_1(x)) \vee 0 + (-z_2(x)) \vee 0 - (-z_1(x) - z_2(x)) \vee 0 \\ &= z_{1-}(x) + z_{2-}(x) - z_-(x) \end{aligned}$$

Hence, (iii) $f_+(z_1) + f_+(z_2) = f_+(z) = f_+(z_1 + z_2)$, $\forall z_1, z_2 \in \mathcal{Z}$.

$\forall z \in \mathcal{Z}$, $\forall \alpha \in \mathbb{R}$. If $\alpha = 0$, then $f_+(\alpha z) = 0 = \alpha f_+(z)$. If $\alpha > 0$, then $f_+(\alpha z) = f_+(\alpha z_+) - f_+(\alpha z_-) = \alpha f_+(z_+) - \alpha f_+(z_-) = \alpha f_+(z)$, where the first equality follows from the definition of f_+ ; the second equality follows from (i); and the third equality follows from the definition of f_+ . If $\alpha < 0$, then $f_+(\alpha z) = f_+(-\alpha z_-) - f_+(-\alpha z_+) = -\alpha f_+(z_-) + \alpha f_+(z_+) = \alpha f_+(z)$, where the first equality follows from the definition of f_+ ; the second equality follows from (i); and the third equality follows from the definition of f_+ .

Hence, we have

(iv) $f_+(\alpha z) = \alpha f_+(z)$, $\forall z \in \mathcal{Z}$ and $\forall \alpha \in \mathbb{R}$.

Hence, $f_+ \in \mathcal{Z}^*$ and is a positive linear functional.

Let $f_- := f_+ - f \in \mathcal{Z}^*$. Clearly, $f_-(z) \geq 0$, $\forall z \in P$. Hence, f_- is also a positive linear functional.

This completes the proof of the lemma. $\square$

Theorem 11.201 (Riesz Representation Theorem) Let $\mathcal{X} := (X, \mathcal{O})$ be a compact Hausdorff topological space, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $\mathcal{Z} := \mathcal{C}(\mathcal{X}, \mathcal{Y})$, and $Z := \{z \in \mathcal{Z} \mid z = hy, h \in \mathcal{C}(\mathcal{X}, \mathbb{R}), y \in \mathcal{Y}\}$. Assume that $\text{span}(Z) = \mathcal{Z}$. Then, $\forall f \in \mathcal{Z}^*$, $\exists! \mu \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$ such that $f(z) = \int_X \langle d\mu(x), z(x) \rangle =: \langle \mu, z \rangle$, $\forall z \in \mathcal{Z}$. Furthermore, $\mathcal{Z}^* = \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$ isometrically isomorphically.

Proof By Proposition 5.14, $\mathcal{X}$ is normal. By Example 7.31, $\mathcal{Z}$ is a normed linear space over $\mathbb{K}$. Fix any $f \in \mathcal{Z}^*$. Define $\nu_o : \mathcal{O} \rightarrow [0, \|f\|] \subset \mathbb{R}$ by $\nu_o(O) := \sup_{z \in \mathcal{Z}, \|z\| \leq 1, \text{supp}(z) \subseteq O} |f(z)|$, $\forall O \in \mathcal{O}$. Clearly, we have

- (i) $\nu_o(\emptyset) = 0$;
- (ii) $0 \leq \nu_o(O) \leq \|f\| < \infty$, $\forall O \in \mathcal{O}$;
- (iii) $\nu_o(O_1) \leq \nu_o(O_2)$, $\forall O_1, O_2 \in \mathcal{O}$ with $O_1 \subseteq O_2$.

$\forall (O_i)_{i=1}^\infty \subseteq \mathcal{O}$, let $O := \bigcup_{i=1}^\infty O_i \in \mathcal{O}$. $\forall z \in \mathcal{Z}$ with $K := \text{supp}(z) \subseteq O$ and $\|z\| \leq 1$, by Proposition 5.5, K is compact. By Corollary 5.65 of Partition of Unity, $\exists m \in \mathbb{Z}_+$, $\exists n_1, \dots, n_m \in \mathbb{N}$, which may be taken to be distinct, $\exists (\varphi_{n_i})_{i=1}^m \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ such that $\varphi_{n_i} : X \rightarrow [0, 1] \subset \mathbb{R}$, $\text{supp}(\varphi_{n_i}) \subseteq O_{n_i}$, $i = 1, \dots, m$, $\sum_{i=1}^m \varphi_{n_i}(x) = 1$, $\forall x \in K$, and $0 \leq \sum_{i=1}^m \varphi_{n_i}(x) \leq 1$, $\forall x \in X$. Then, $z = \sum_{i=1}^m z\varphi_{n_i}$. $\forall i \in \{1, \dots, m\}$, $z\varphi_{n_i} \in \mathcal{Z}$, $\|z\varphi_{n_i}\| \leq 1$, and $\text{supp}(z\varphi_{n_i}) \subseteq \text{supp}(z) \cap \text{supp}(\varphi_{n_i}) \subseteq \text{supp}(\varphi_{n_i}) \subseteq O_{n_i}$. Then, $|f(z)| = |\sum_{i=1}^m f(z\varphi_{n_i})| \leq \sum_{i=1}^m |f(z\varphi_{n_i})| \leq \sum_{i=1}^m \nu_o(O_{n_i}) \leq \sum_{i=1}^\infty \nu_o(O_i)$. By the arbitrariness of z , we have

- (iv) $\nu_o(O) = \nu_o(\bigcup_{i=1}^\infty O_i) \leq \sum_{i=1}^\infty \nu_o(O_i)$, $\forall (O_i)_{i=1}^\infty \subseteq \mathcal{O}$.

$\forall O_1, O_2 \in \mathcal{O}$ with $O_1 \cap O_2 = \emptyset$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists \bar{z}_i \in \mathcal{Z}$ with $\text{supp}(\bar{z}_i) \subseteq O_i$ and $\|\bar{z}_i\| \leq 1$, such that $|f(\bar{z}_i)| > \nu_o(O_i) - \epsilon/2$, $i = 1, 2$. $\forall i \in \{1, 2\}$, when $f(\bar{z}_i) \neq 0$, take $z_i := \frac{|f(\bar{z}_i)|}{f(\bar{z}_i)} \bar{z}_i \in \mathcal{Z}$; and when $f(\bar{z}_i) = 0$, take $z_i := \bar{z}_i \in \mathcal{Z}$. Then, $z_i \in \mathcal{Z}$, $\|z_i\| = \|\bar{z}_i\| \leq 1$, $\text{supp}(z_i) = \text{supp}(\bar{z}_i) \subseteq O_i$, and $f(z_i) = |f(\bar{z}_i)| > \nu_o(O_i) - \epsilon/2$. Let $z := z_1 + z_2 \in \mathcal{Z}$. Clearly, $\|z\| \leq 1$ since $O_1 \cap O_2 = \emptyset$. Note that $\text{supp}(z) \subseteq \text{supp}(z_1) \cup \text{supp}(z_2) = \text{supp}(z_1) \cup \text{supp}(z_2) \subseteq O_1 \cup O_2$, where the equality follows from Proposition 3.3. Then, $\nu_o(O_1) + \nu_o(O_2) - \epsilon < f(z_1) + f(z_2) = f(z) \leq \nu_o(O_1 \cup O_2) \leq \nu_o(O_1) + \nu_o(O_2)$, where the last inequality follows from (iv). By the arbitrariness of ϵ , we have

- (v) $\nu_o(O_1 \cup O_2) = \nu_o(O_1) + \nu_o(O_2)$, $\forall O_1, O_2 \in \mathcal{O}$ with $O_1 \cap O_2 = \emptyset$.

$\forall O \in \mathcal{O}$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists z \in \mathcal{Z}$ with $K := \text{supp}(z) \subseteq O$ and $\|z\| \leq 1$, such that $|f(z)| > \nu_o(O) - \epsilon$. By Proposition 5.5, K is compact. By Proposition 3.35, $\exists U \in \mathcal{O}$, such that $K \subseteq U \subseteq \bar{U} \subseteq O$. Then, $\nu_o(O) \geq \nu_o(U) \geq |f(z)| > \nu_o(O) - \epsilon$, where the first inequality follows from (iii). Hence, we have

- (vi) $\nu_o(O) = \sup_{U \in \mathcal{O}, U \subseteq \bar{U} \subseteq O} \nu_o(U)$, $\forall O \in \mathcal{O}$.

By Proposition 11.193, there exists a finite topological measure space $\bar{\mathcal{X}} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \nu)$ such that $\nu(O) = \nu_o(O)$, $\forall O \in \mathcal{O}$ and $\nu(E) = \inf_{O \in \mathcal{O}, E \subseteq O} \nu_o(O)$, $\forall E \in \mathcal{B}_B(\mathcal{X})$. Furthermore, ν is unique among measures $\hat{\nu}$ on $(\mathcal{X}, \mathcal{B}_B(\mathcal{X}))$ with $\hat{\nu}|_{\mathcal{O}} = \nu_o$. Clearly, $\nu \in \mathcal{M}_{ft}(\mathcal{X})$.

Fix any $\tilde{F} \in \mathcal{O}$. Let $\mathcal{A}_F := \{(U, V, h) \in \mathcal{O} \times \mathcal{O} \times \mathcal{C}(\mathcal{X}, \mathbb{R}) \mid F \subseteq U \subseteq \bar{U} \subseteq V, h : X \rightarrow [0, 1] \subset \mathbb{R}, h|_{\bar{U}} = 1, \text{supp}(h) \subseteq V\}$. Define a relation $\prec$ on $\mathcal{A}_F$ by $(U_1, V_1, h_1) \prec (U_2, V_2, h_2)$ if $V_1 \supseteq V_2$. By Proposition 3.35 and Urysohn's Lemma 3.55, $\bar{\mathcal{A}}_F := (\mathcal{A}_F, \prec)$ is a directed system. $\forall (U, V, h) \in \bar{\mathcal{A}}_F$, the function $f_h : \mathcal{Y} \rightarrow \mathbb{K}$, defined by $f_h(y) = f(hy)$, $\forall y \in \mathcal{Y}$, is a bounded linear functional since $f \in \mathcal{Z}^*$. Then, $f_h \in \mathcal{Y}^*$. This defines a net $(f_h)_{(U, V, h) \in \bar{\mathcal{A}}_F} \subseteq \mathcal{Y}^*$.

Claim 11.201.1 *The net $(f_h)_{(U,V,h) \in \bar{\mathcal{A}}_F} \subseteq \mathcal{Y}^*$ is Cauchy.*

Proof of claim: $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}, \exists V \in \mathcal{O}$ with $F \subseteq V$ such that $\nu(V \setminus F) < \epsilon$. By Proposition 3.35, $\exists U, \tilde{U} \in \mathcal{O}$ such that $F \subseteq U \subseteq \bar{U} \subseteq \tilde{U} \subseteq \bar{\tilde{U}} \subseteq V$. By Urysohn's Lemma 3.55, $\exists h \in \mathcal{C}(\mathcal{X}, \mathbb{R})$ with $h : X \rightarrow [0, 1] \subset \mathbb{R}$ such that $h|_{\bar{U}} = 1$ and $h|_{\tilde{U}} = 0$. Then, $\text{supp}(h) \subseteq \bar{U} \subseteq V$ and $(U, V, h) \in \bar{\mathcal{A}}_F$. $\forall (U_1, V_1, h_1) \in \bar{\mathcal{A}}_F$ with $(U, V, h) \prec (U_1, V_1, h_1)$, we have $V_1 \subseteq V$ and $\|f_h - f_{h_1}\| = \sup_{y \in \mathcal{Y}, \|y\| \leq 1} |\langle f_h, y \rangle - \langle f_{h_1}, y \rangle| = \sup_{y \in \mathcal{Y}, \|y\| \leq 1} |f(hy) - f(h_1y)| = \sup_{y \in \mathcal{Y}, \|y\| \leq 1} |f((h - h_1)y)|$. Note that, $\forall y \in \mathcal{Y}$ with $\|y\| \leq 1$, $(h - h_1)y \in \mathcal{Z}$, $\|(h - h_1)y\| \leq 1$, and

$$\begin{aligned} \text{supp}((h - h_1)y) &\subseteq \text{supp}(h - h_1) \\ &\subseteq \overline{(\text{supp}(h) \cup \text{supp}(h_1)) \setminus \{x \in X \mid h(x) = 1 = h_1(x)\}} \\ &\subseteq \overline{(\text{supp}(h) \cup \text{supp}(h_1)) \setminus (\bar{U} \cap \bar{U}_1)} \\ &= \overline{\text{supp}(h) \setminus (\bar{U} \cap \bar{U}_1) \cup \text{supp}(h_1) \setminus (\bar{U} \cap \bar{U}_1)} \\ &\subseteq \overline{\text{supp}(h) \setminus (U \cap U_1) \cup \text{supp}(h_1) \setminus (U \cap U_1)} \\ &= (\text{supp}(h) \setminus (U \cap U_1)) \cup (\text{supp}(h_1) \setminus (U \cap U_1)) \\ &\subseteq (V \setminus F) \cup (V_1 \setminus F) \subseteq V \setminus F \in \mathcal{O} \end{aligned}$$

where the third containment follows from the fact that $h|_{\bar{U}} = 1$ and $h_1|_{\bar{U}_1} = 1$; and the first equality follows from Proposition 3.3. Then, $\|f_h - f_{h_1}\| \leq \nu_o(V \setminus F) = \nu(V \setminus F) < \epsilon$. Hence, the net is Cauchy. This completes the proof of the claim. $\square$

By Propositions 4.44 and 7.72,

(vii) $\lim_{(U,V,h) \in \bar{\mathcal{A}}_F} f_h =: \bar{\mu}(F) \in \mathcal{Y}^*, \forall F \in \mathcal{O}$.

Thus, we have defined a function $\bar{\mu}$ that assigns a vector $\bar{\mu}(F) \in \mathcal{Y}^*$ to each closed subset $F \subseteq X$. We will next show that $\bar{\mu}$ satisfies the assumption of Proposition 11.194.

Fix any $\tilde{F}_1, \tilde{F}_2 \in \mathcal{O}$ with $F_1 \cap F_2 = \emptyset$. Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. By the normality of $\mathcal{X}$, $\exists V_1, V_2 \in \mathcal{O}$ with $V_1 \cap V_2 = \emptyset$ such that $F_i \subseteq V_i, i = 1, 2$. Without loss of generality, we may assume that $\nu(V_i \setminus F_i) < \epsilon/4, i = 1, 2$. Fix any $i \in \{1, 2\}$. By $\lim_{(U,V,h) \in \bar{\mathcal{A}}_{F_i}} f_h = \bar{\mu}(F_i)$, $\exists(\hat{U}_i, \hat{V}_i, \hat{h}_i) \in \bar{\mathcal{A}}_{F_i}$ with $\hat{V}_i \subseteq V_i$ such that $\|\bar{\mu}(F_i) - f_{\hat{h}_i}\| < \epsilon/6$. By $\lim_{(U,V,h) \in \bar{\mathcal{A}}_{F_1 \cup F_2}} f_h = \bar{\mu}(F_1 \cup F_2)$, $\exists(\hat{U}, \hat{V}, \hat{h}) \in \bar{\mathcal{A}}_{F_1 \cup F_2}$ with $\hat{V} \subseteq V_1 \cup V_2$ such that $\|\bar{\mu}(F_1 \cup F_2) - f_{\hat{h}}\| < \epsilon/6$. This implies that $\|\bar{\mu}(F_1 \cup F_2) - \bar{\mu}(F_1) - \bar{\mu}(F_2)\| \leq \|\bar{\mu}(F_1 \cup F_2) - f_{\hat{h}}\| + \|f_{\hat{h}} - f_{\hat{h}_1} - f_{\hat{h}_2}\| + \|f_{\hat{h}_1} - \bar{\mu}(F_1)\| + \|f_{\hat{h}_2} - \bar{\mu}(F_2)\| < \epsilon/2 + \sup_{y \in \mathcal{Y}, \|y\| \leq 1} |\langle f_{\hat{h}}, y \rangle - \langle f_{\hat{h}_1}, y \rangle - \langle f_{\hat{h}_2}, y \rangle| = \epsilon/2 + \sup_{y \in \mathcal{Y}, \|y\| \leq 1} |f(\hat{h}y) - f(\hat{h}_1y) - f(\hat{h}_2y)| = \epsilon/2 + \sup_{y \in \mathcal{Y}, \|y\| \leq 1} |f((\hat{h} - \hat{h}_1 - \hat{h}_2)y)|$. Note that, $\forall y \in \mathcal{Y}$ with $\|y\| \leq 1$, $\hat{z} := (\hat{h} - \hat{h}_1 - \hat{h}_2)y \in \mathcal{Z}$, $\|\hat{z}\| \leq 1$, since $\text{supp}(\hat{h}_1) \cap \text{supp}(\hat{h}_2) \subseteq \hat{V}_1 \cap \hat{V}_2 = \emptyset$, and $\text{supp}(\hat{z}) \subseteq$

$(\text{supp}(\hat{h}) \cup \text{supp}(\hat{h}_1) \cup \text{supp}(\hat{h}_2)) \setminus ((\overline{\hat{U} \cap \hat{U}_1}) \cup (\overline{\hat{U} \cap \hat{U}_2}))$. Then, $\text{supp}(\hat{z}) \subseteq \text{supp}(\hat{h}) \cup \text{supp}(\hat{h}_1) \cup \text{supp}(\hat{h}_2) = \text{supp}(\hat{h}) \cup \text{supp}(\hat{h}_1) \cup \text{supp}(\hat{h}_2) \subseteq \hat{V} \cup \hat{V}_1 \cup \hat{V}_2 \subseteq V_1 \cup V_2$. In addition, $\text{supp}(\hat{z}) \subseteq \left((\overline{\hat{U} \cap \hat{U}_1}) \cup (\overline{\hat{U} \cap \hat{U}_2}) \right)^\sim \subseteq \left((\hat{U} \cap \hat{U}_1) \cup (\hat{U} \cap \hat{U}_2) \right)^\sim = \left(\hat{U} \cap (\hat{U}_1 \cup \hat{U}_2) \right)^\sim \subseteq \widetilde{F_1 \cup F_2}$. This yields $\text{supp}(\hat{z}) \subseteq (V_1 \cup V_2) \cap \widetilde{F_1 \cup F_2} = (V_1 \setminus F_1) \cup (V_2 \setminus F_2) =: \hat{O} \in \mathcal{O}$. Then, we have $\|\bar{\mu}(F_1 \cup F_2) - \bar{\mu}(F_1) - \bar{\mu}(F_2)\| < \epsilon/2 + \nu_o(\hat{O}) = \epsilon/2 + \nu(\hat{O}) = \epsilon/2 + \nu(V_1 \setminus F_1) + \nu(V_2 \setminus F_2) < \epsilon$. By the arbitrariness of ϵ , we have (viii) $\bar{\mu}(F_1) + \bar{\mu}(F_2) = \bar{\mu}(F_1 \cup F_2)$, $\forall \widetilde{F_1}, \widetilde{F_2} \in \mathcal{O}$ with $F_1 \cap F_2 = \emptyset$.

Fix any $\widetilde{F_1}, \widetilde{F_2}, O \in \mathcal{O}$ with $F_1 \triangle F_2 \subseteq O$. $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall i \in \{1, 2\}$, $\exists V_i \in \mathcal{O}$ with $F_i \subseteq V_i$ such that $\nu(V_i \setminus F_i) < \epsilon/4$. By $\lim_{(U, V, h) \in \bar{\mathcal{A}}_{F_i}} f_h = \bar{\mu}(F_i)$, $\exists (\hat{U}_i, \hat{V}_i, \hat{h}_i) \in \bar{\mathcal{A}}_{F_i}$ with $\hat{V}_i \subseteq V_i$ such that $\|\bar{\mu}(F_i) - f_{\hat{h}_i}\| < \epsilon/4$. This leads to $\|\bar{\mu}(F_1) - \bar{\mu}(F_2)\| \leq \|\bar{\mu}(F_1) - f_{\hat{h}_1}\| + \|f_{\hat{h}_1} - f_{\hat{h}_2}\| + \|f_{\hat{h}_2} - \bar{\mu}(F_2)\| < \epsilon/2 + \sup_{y \in \mathcal{Y}, \|y\| \leq 1} |\langle f_{\hat{h}_1}, y \rangle - \langle f_{\hat{h}_2}, y \rangle| = \epsilon/2 + \sup_{y \in \mathcal{Y}, \|y\| \leq 1} |f((\hat{h}_1 - \hat{h}_2)y)|$. Note that, $\forall y \in \mathcal{Y}$ with $\|y\| \leq 1$, $\hat{z} := (\hat{h}_1 - \hat{h}_2)y \in \mathcal{Z}$, $\|\hat{z}\| \leq 1$, and $\text{supp}(\hat{z}) \subseteq (\text{supp}(\hat{h}_1) \cup \text{supp}(\hat{h}_2)) \setminus (\overline{\hat{U}_1 \cap \hat{U}_2}) \subseteq (\text{supp}(\hat{h}_1) \cup \text{supp}(\hat{h}_2)) \setminus (\hat{U}_1 \cap \hat{U}_2) = (\text{supp}(\hat{h}_1) \cup \text{supp}(\hat{h}_2)) \setminus (\hat{U}_1 \cap \hat{U}_2) \subseteq (V_1 \cup V_2) \setminus (F_1 \cap F_2) = (V_1 \setminus (F_1 \cap F_2)) \cup (V_2 \setminus (F_1 \cap F_2)) \subseteq (V_1 \setminus F_1) \cup (F_1 \setminus F_2) \cup (F_2 \setminus F_1) \cup (V_2 \setminus F_2) = (V_1 \setminus F_1) \cup (V_2 \setminus F_2) \cup (F_1 \triangle F_2) \subseteq (V_1 \setminus F_1) \cup (V_2 \setminus F_2) \cup O =: \hat{O} \in \mathcal{O}$. Thus, $\|\bar{\mu}(F_1) - \bar{\mu}(F_2)\| < \epsilon/2 + \nu_o(\hat{O}) = \epsilon/2 + \nu(\hat{O}) \leq \epsilon/2 + \nu(V_1 \setminus F_1) + \nu(V_2 \setminus F_2) + \nu(O) < \epsilon + \nu(O)$. By the arbitrariness of ϵ , we have (ix) $\|\bar{\mu}(F_1) - \bar{\mu}(F_2)\| \leq \nu(O)$, $\forall \widetilde{F_1}, \widetilde{F_2}, O \in \mathcal{O}$ with $F_1 \triangle F_2 \subseteq O$.

By Proposition 11.194, $\exists! \mu \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$ such that $\mu(F) = \bar{\mu}(F)$, $\forall \widetilde{F} \in \mathcal{O}$. Furthermore, $\mathcal{P} \circ \mu \leq \nu$. Then, $\|\mu\| = \mathcal{P} \circ \mu(X) \leq \nu(X) \leq \|f\| < \infty$. This defines a mapping $\Phi_{\mathcal{Z}} : \mathcal{Z}^* \rightarrow \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$ by $\Phi_{\mathcal{Z}}(f) = \mu$, $\forall f \in \mathcal{Z}^*$. The above can be expressed as

(x) $\|\Phi_{\mathcal{Z}}(f)\| \leq \|f\|$, $\forall f \in \mathcal{Z}^*$.

$\forall f_1, f_2 \in \mathcal{Z}^*$, $\forall \alpha \in \mathbb{K}$, let $\mu_i := \Phi_{\mathcal{Z}}(f_i) \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$, $i = 1, 2$, $f := f_1 + f_2 \in \mathcal{Z}^*$, and $\mu := \Phi_{\mathcal{Z}}(f) \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$. $\forall \widetilde{F} \in \mathcal{O}$, $\mu(F) = \bar{\mu}(F) = \lim_{(U, V, h) \in \bar{\mathcal{A}}_F} f_h = \lim_{(U, V, h) \in \bar{\mathcal{A}}_F} (f_1 h + f_2 h) = \lim_{(U, V, h) \in \bar{\mathcal{A}}_F} f_1 h + \lim_{(U, V, h) \in \bar{\mathcal{A}}_F} f_2 h = \bar{\mu}_1(F) + \bar{\mu}_2(F) = \mu_1(F) + \mu_2(F) = (\mu_1 + \mu_2)(F)$. By Proposition 11.192, $\mu_1 + \mu_2 \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$. Then, by Proposition 11.194, $\Phi_{\mathcal{Z}}(f_1 + f_2) = \mu = \mu_1 + \mu_2 = \Phi_{\mathcal{Z}}(f_1) + \Phi_{\mathcal{Z}}(f_2)$. Let $\hat{\mu} := \Phi_{\mathcal{Z}}(\alpha f_1) \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$. $\forall \widetilde{F} \in \mathcal{O}$, we have $\hat{\mu}(F) = \tilde{\mu}(F) = \lim_{(U, V, h) \in \bar{\mathcal{A}}_F} (\alpha f_1)_h = \lim_{(U, V, h) \in \bar{\mathcal{A}}_F} \alpha f_1 h = \alpha \bar{\mu}_1(F) = (\alpha \mu_1)(F)$. By Proposition 11.192, $\alpha \mu_1 \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$. Then, by Proposition 11.194, $\Phi_{\mathcal{Z}}(\alpha f_1) = \hat{\mu} = \alpha \mu_1 = \alpha \Phi_{\mathcal{Z}}(f_1)$. This shows that $\Phi_{\mathcal{Z}}$ is a bound linear operator.

Fix any $\mu \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$. $\forall z \in \mathcal{Z}$. By Proposition 11.37, z is $\mathcal{B}_{\mathcal{B}}(\mathcal{X})$ -measurable. Clearly, $\mathcal{P} \circ z(x) \leq \|z\|$, $\forall x \in X$. Then, $\int_X \mathcal{P} \circ z d\mathcal{P} \circ \mu \leq$

$\|z\| \mathcal{P} \circ \mu(X) = \|z\| \|\mu\| < \infty$. Hence, z is absolutely integrable over $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$. By Proposition 5.7, $z(X) \subseteq \mathcal{Y}$ is compact. Then, by Propositions 11.190, 7.35, and 7.17, $H := \overline{\text{span}(z(X))} \subseteq \mathcal{Y} \subseteq \mathcal{Y}^{**}$ is a separable subspace. By Proposition 11.132, $\langle\langle \mu, z \rangle\rangle = \int_X \langle d\mu(x), z(x) \rangle \in \mathbb{K}$ and $|\langle\langle \mu, z \rangle\rangle| = \left| \int_X \langle d\mu(x), z(x) \rangle \right| \leq \|z\| \|\mu\| < \infty$. Hence, we have (xi) $\forall \mu \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$, $\forall z \in \mathcal{Z}$, we have z is absolutely integrable and integrable over $(\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$, and $|\langle\langle \mu, z \rangle\rangle| \leq \|z\| \|\mu\| < \infty$.

We will show that $f(z) = \int_X \langle d\mu(x), z(x) \rangle = \langle\langle \mu, z \rangle\rangle$, $\forall z \in \mathcal{Z}$, $\forall f \in \mathcal{Z}^*$, where $\mu = \Phi_{\mathcal{Z}}(f)$, in four steps. In the first step, we consider the special case: $\mathcal{Y} = \mathbb{R}$ and f is a positive linear functional. $\forall \tilde{F} \in \mathcal{O}$, $\forall (U, V, h) \in \tilde{\mathcal{A}}_F$, $f_h = f(h) \in [0, \|f\|] \subset \mathbb{R}$, since $h \in P := \{z \in \mathcal{C}(\mathcal{X}, \mathbb{R}) \mid z : X \rightarrow [0, \infty) \subset \mathbb{R}\} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R}) = \mathcal{Z}$. Then, $\bar{\mu}(F) = \lim_{(U, V, h) \in \tilde{\mathcal{A}}_F} f_h \in [0, \|f\|] \subset \mathbb{R}$. Hence, by Proposition 11.194 and its proof, $\mu : \mathcal{B}_B(\mathcal{X}) \rightarrow [0, \|f\|] \subset \mathbb{R}$ and $\mu \in \mathcal{M}_{ft}(\mathcal{X})$. Fix any $z \in \mathcal{Z}$. We will distinguish three exhaustive cases: Case 1: $z = \vartheta_{\mathcal{Z}}$; Case 2: $z \neq \vartheta_{\mathcal{Z}}$ and $z \in P$; Case 3: $z \in \mathcal{Z}$. Case 1: $z = \vartheta_{\mathcal{Z}}$. Then, $f(z) = 0 = \int_X z(x) d\mu(x) = \langle\langle \mu, z \rangle\rangle$.

Case 2: $z \neq \vartheta_{\mathcal{Z}}$ and $z \in P$. Let $\bar{z} := z/\|z\| \in P$. Then, $\bar{z} : X \rightarrow [0, 1] \subset \mathbb{R}$. Fix any $n \in \mathbb{N}$ and fix any $i \in \{1, \dots, n\}$. Define $\phi_i := ((n\bar{z} - i + 1) \vee 0) \wedge 1 \in P$, $O_i := \{x \in X \mid n\bar{z}(x) > i - 1\}$, and $F_i := \{x \in X \mid n\bar{z}(x) \geq i - 1\}$. Let $O_0 := \{x \in X \mid n\bar{z}(x) > -1\} = X$, $O_{n+1} := \{x \in X \mid n\bar{z}(x) > n\} = \emptyset$, $F_{n+1} := \{x \in X \mid n\bar{z}(x) \geq n\}$, and $F_{n+2} := \{x \in X \mid n\bar{z}(x) \geq n + 1\} = \emptyset$. Clearly, $F_{i+2} \subseteq O_{i+1} \subseteq \overline{O_{i+1}} \subseteq F_{i+1} \subseteq O_i \subseteq \overline{O_i} \subseteq F_i \subseteq O_{i-1}$, $\phi_i : X \rightarrow [0, 1] \subset \mathbb{R}$, $\phi_i|_{F_{i+1}} = 1$, and $\text{supp}(\phi_i) \subseteq \overline{O_i} \subseteq F_i \subseteq O_{i-1}$. Clearly $O_i \in \mathcal{O}$ and $\widetilde{F_{i+1}} \in \mathcal{O}$, $i = 0, \dots, n+1$, and $\bar{z} = \frac{1}{n} \sum_{i=1}^n \phi_i$. Fix any $i \in \{1, \dots, n\}$. By the definition of $\bar{\mu}$ and the fact that f is a positive linear functional, $\forall (U, V, h) \in \tilde{\mathcal{A}}_{F_{i+2}}$ with $V \subseteq O_{i+1}$, we have $h(x) \leq \chi_{O_{i+1}, X}(x) \leq \chi_{F_{i+1}, X}(x) \leq \phi_i(x)$, $\forall x \in X$, and $f_h = f(h) \leq f(\phi_i)$. This implies that $\mu(F_{i+2}) \leq f(\phi_i)$. $\forall (U, V, h) \in \tilde{\mathcal{A}}_{F_i}$, we have $\phi_i(x) \leq \chi_{F_i, X}(x) \leq h(x)$, $\forall x \in X$, and $f_h = f(h) \geq f(\phi_i)$. This leads to $f(\phi_i) \leq \mu(F_i)$. By Propositions 11.75 and 11.83, we have $\mu(F_{i+2}) \leq \mu(F_{i+1}) = \int_X \chi_{F_{i+1}, X} d\mu \leq \int_X \phi_i d\mu \leq \int_X \chi_{F_i, X} d\mu = \mu(F_i)$. Then, $|f(\phi_i) - \int_X \phi_i d\mu| \leq \mu(F_i) - \mu(F_{i+2})$. Summing both side from 1 to n , we have $|f(\bar{z}) - \int_X \bar{z} d\mu| = \left| \frac{1}{n} \sum_{i=1}^n (f(\phi_i) - \int_X \phi_i d\mu) \right| \leq \frac{1}{n} \sum_{i=1}^n |f(\phi_i) - \int_X \phi_i d\mu| \leq \frac{1}{n} (\mu(F_1) + \mu(F_2)) \leq 2\mu(X)/n = 2\|\mu\|/n \leq 2\|f\|/n$, where the first equality follows from (xi) and Proposition 11.92; and the last inequality follows from (x). By the arbitrariness of n , we have $f(\bar{z}) = \int_X \bar{z} d\mu$. By (xi) and Proposition 11.92, $f(z) = f(\|z\| \bar{z}) = \|z\| f(\bar{z}) = \|z\| \int_X \bar{z} d\mu = \int_X z d\mu = \langle\langle \mu, z \rangle\rangle$.

Case 3: $z \in \mathcal{Z}$. Let $z_+ := z \vee 0 \in P$ and $z_- := (-z) \vee 0 \in P$. Clearly, $z = z_+ - z_-$. By Cases 1 and 2, $f(z) = f(z_+) - f(z_-) = \int_X z_+ d\mu - \int_X z_- d\mu = \int_X z d\mu = \langle\langle \mu, z \rangle\rangle$, where the third equality follows from (xi) and Proposition 11.92.

Hence, $f(z) = \int_X z d\mu = \langle\langle \mu, z \rangle\rangle$, $\forall z \in \mathcal{Z}$. This completes the first step.

In the second step, we consider the special case: $\mathcal{Y} = \mathbb{R}$. By Lemma 11.200, $f = f_+ - f_-$, where $f_+, f_- \in \mathcal{Z}^*$ are positive linear functionals. Let $\mu_+ := \Phi_{\mathcal{Z}}(f_+)$ and $\mu_- := \Phi_{\mathcal{Z}}(f_-)$. By the first step, $f_+(z) = \int_X z d\mu_+$ and $f_-(z) = \int_X z d\mu_-$, $\forall z \in \mathcal{Z}$. Then, by linearity of $\Phi_{\mathcal{Z}}$, $\mu = \Phi_{\mathcal{Z}}(f) = \Phi_{\mathcal{Z}}(f_+ - f_-) = \Phi_{\mathcal{Z}}(f_+) - \Phi_{\mathcal{Z}}(f_-) = \mu_+ - \mu_-$. By Propositions 11.145 and 11.146, $f(z) = f_+(z) - f_-(z) = \int_X z d\mu_+ - \int_X z d\mu_- = \int_X z d\mu = \langle \mu, z \rangle$. This completes the second step.

In the third step, we consider the special case when $\mathcal{Y}$ is a real normed linear space. $\forall y \in \mathcal{Y}$, define $\mu_y : \mathcal{B}_{\mathcal{B}}(\mathcal{X}) \rightarrow \mathbb{R}$ by $\mu_y(E) = \langle \mu(E), y \rangle$, $\forall E \in \mathcal{B}_{\mathcal{B}}(\mathcal{X})$. Then, $\mu_y \in \mathcal{M}_{ft}(\mathcal{X}, \mathbb{R})$ by Proposition 11.136. Define $f_y : \mathcal{C}(\mathcal{X}, \mathbb{R}) \rightarrow \mathbb{R}$ by $f_y(h) = \langle \mu, hy \rangle$, $\forall h \in \mathcal{C}(\mathcal{X}, \mathbb{R})$. By $f \in \mathcal{Z}^*$, we have $f_y \in (\mathcal{C}(\mathcal{X}, \mathbb{R}))^*$. $\forall \tilde{F} \in \mathcal{O}$, $\mu_y(F) = \langle \mu(F), y \rangle = \langle \langle \lim_{(U,V,h) \in \tilde{\mathcal{A}}_F} f_h, y \rangle \rangle = \lim_{(U,V,h) \in \tilde{\mathcal{A}}_F} \langle \langle f_h, y \rangle \rangle = \lim_{(U,V,h) \in \tilde{\mathcal{A}}_F} f(hy) = \lim_{(U,V,h) \in \tilde{\mathcal{A}}_F} f_y(h) = \Phi_{\mathcal{C}(\mathcal{X}, \mathbb{R})}(f_y)(F)$. By Proposition 11.194, we have $\mu_y = \Phi_{\mathcal{C}(\mathcal{X}, \mathbb{R})}(f_y)$. By the second step, we have $f(hy) = f_y(h) = \int_X h d\mu_y = \int_X \langle d\mu(x), h(x)y \rangle = \langle \mu, hy \rangle$, $\forall h \in \mathcal{C}(\mathcal{X}, \mathbb{R})$, where the third equality follows from Proposition 11.146. $\forall z \in \mathcal{Z}$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists \bar{z} \in \text{span}(Z)$ such that $\|z - \bar{z}\| < \epsilon/(2\|f\| + 1)$. Then, $\bar{z} = \sum_{i=1}^n h_i y_i$, where $n \in \mathbb{Z}_+$, $h_1, \dots, h_n \in \mathcal{C}(\mathcal{X}, \mathbb{R})$, and $y_1, \dots, y_n \in \mathcal{Y}$. Then, $|f(z) - \langle \mu, z \rangle| \leq |f(z) - f(\bar{z})| + |f(\bar{z}) - \langle \mu, \bar{z} \rangle| + |\langle \mu, \bar{z} \rangle - \langle \mu, z \rangle| \leq \|f\| \|z - \bar{z}\| + \left| \sum_{i=1}^n f(h_i y_i) - \int_X \langle d\mu(x), \sum_{i=1}^n h_i(x) y_i \rangle \right| + \left| \int_X \langle d\mu(x), \bar{z}(x) \rangle - \int_X \langle d\mu(x), z(x) \rangle \right| \leq \epsilon/2 + \left| \sum_{i=1}^n (f(h_i y_i) - \int_X \langle d\mu(x), h_i(x) y_i \rangle) \right| + \left| \int_X \langle d\mu(x), \bar{z}(x) - z(x) \rangle \right| \leq \epsilon/2 + 0 + \int_X \mathcal{P} \circ (\bar{z} - z)(x) d\mathcal{P} \circ \mu(x) \leq \epsilon/2 + \|z - \bar{z}\| \mathcal{P} \circ \mu(X) \leq \epsilon/2 + \frac{\epsilon}{2\|f\|+1} \|\mu\| \leq \epsilon/2 + \frac{\epsilon}{2\|f\|+1} \|f\| < \epsilon$, where the second inequality follows from Proposition 7.72; the third inequality follows from (xi) and Proposition 11.132; the fourth inequality follows from Proposition 11.132; the fifth inequality follows from Propositions 11.75 and 11.83; and the seventh inequality follows from (x). By the arbitrariness of ϵ , we have $f(z) = \langle \mu, z \rangle = \int_X \langle d\mu(x), z(x) \rangle$. This completes the third step.

In the fourth step, we consider the special case when $\mathcal{Y}$ is a complex normed linear space. Let $g := \text{Re} \circ f$. By Lemmas 7.40 and 7.81, g is a bounded linear functional of the real normed linear space $\mathcal{Z}_{\mathbb{R}} := (\mathcal{C}(\mathcal{X}, \mathcal{Y}_{\mathbb{R}}), \mathbb{R}, \|\cdot\|_{\mathcal{C}(\mathcal{X}, \mathcal{Y}_{\mathbb{R}})})$ and $f(z) = g(z) - i g(iz)$, $\forall z \in \mathcal{Z}_{\mathbb{R}}$. Let $\bar{g} \in \mathcal{Z}_{\mathbb{R}}^*$ be defined by $\bar{g}(z) = g(iz)$, $\forall z \in \mathcal{Z}_{\mathbb{R}}$, $\mu_{gr} := \Phi_{\mathcal{Z}_{\mathbb{R}}}(g) \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}_{\mathbb{R}}^*)$, and $\mu_{gi} := \Phi_{\mathcal{Z}_{\mathbb{R}}}(\bar{g}) \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}_{\mathbb{R}}^*)$. $\forall \tilde{F} \in \mathcal{O}$, $\mu(F) = \lim_{(U,V,h) \in \tilde{\mathcal{A}}_F} f_h = \lim_{(U,V,h) \in \tilde{\mathcal{A}}_F} (g_h - i \bar{g}_h) = \lim_{(U,V,h) \in \tilde{\mathcal{A}}_F} g_h - i \lim_{(U,V,h) \in \tilde{\mathcal{A}}_F} \bar{g}_h = \Phi_{\mathcal{Z}_{\mathbb{R}}}(g)(F) - i \Phi_{\mathcal{Z}_{\mathbb{R}}}(\bar{g})(F) = \mu_{gr}(F) - i \mu_{gi}(F) = (\mu_{gr} - i \mu_{gi})(F)$. By the definition of $\bar{g}$, we have $\mu_{gi}(E)(y) = \mu_{gr}(E)(iy) \in \mathbb{C}$, $\forall E \in \mathcal{B}_{\mathcal{B}}(\mathcal{X})$, $\forall y \in \mathcal{Y}_{\mathbb{R}}$. Then, by Lemma 7.81, $\mu_{gr}(E) - i \mu_{gi}(E) \in \mathcal{Y}^*$, $\forall E \in \mathcal{B}_{\mathcal{B}}(\mathcal{X})$. By Proposition 11.136, $\mu_{gr} - i \mu_{gi} \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$. By Proposition 11.194, we have $\mu = \mu_{gr} - i \mu_{gi}$. By the third step, $g(z) = \int_X \langle d\mu_{gr}(x), z(x) \rangle$ and $\bar{g}(z) = \int_X \langle d\mu_{gi}(x), z(x) \rangle$, $\forall z \in \mathcal{Z}_{\mathbb{R}}$. This implies, $\forall z \in \mathcal{Z}$, $f(z) = g(z) - i g(iz) = g(z) - i \bar{g}(z) = \int_X \langle d\mu_{gr}(x), z(x) \rangle - i \int_X \langle d\mu_{gi}(x), z(x) \rangle =$

$\int_X \langle \langle d(\mu_{gr} - i\mu_{gi})(x), z(x) \rangle \rangle = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle = \langle \langle \mu, z \rangle \rangle$, where the fourth equality follows from Propositions 11.145 and 11.146. This completes the fourth step.

Thus, we have the representation

(xii) $f(z) = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle = \langle \langle \mu, z \rangle \rangle$, $\forall z \in \mathcal{Z}$, $\forall f \in \mathcal{Z}^*$, where $\mu = \Phi_{\mathcal{Z}}(f) \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$.

Fix any $\mu \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$. Define $f : \mathcal{Z} \rightarrow \mathbb{K}$ by $f(z) = \langle \langle \mu, z \rangle \rangle = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle$, $\forall z \in \mathcal{Z}$. We will show that $f \in \mathcal{Z}^*$, $\|f\| \leq \|\mu\|$, and $\Phi_{\mathcal{Z}}(f) = \mu$. $\forall z \in \mathcal{Z}$. By Proposition 5.7, $z(X) \subseteq \mathcal{Y}$ is compact. Then, by Propositions 11.190, 7.35, and 7.17, $H := \overline{\text{span}(z(X))} \subseteq \mathcal{Y}$ is a separable subspace. By (xi) and Proposition 11.132, $f(z) = \langle \langle \mu, z \rangle \rangle = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle \in \mathbb{K}$ is well-defined. By Proposition 11.132, $f \in \mathcal{Z}^*$ and $\|f\| = \sup_{z \in \mathcal{Z}, \|z\| \leq 1} |\langle \langle \mu, z \rangle \rangle| = \sup_{z \in \mathcal{Z}, \|z\| \leq 1} \left| \int_X \langle \langle d\mu(x), z(x) \rangle \rangle \right| \leq \sup_{z \in \mathcal{Z}, \|z\| \leq 1} \|z\| \|\mu\| \leq \|\mu\|$.

Let $\hat{\mu} := \Phi_{\mathcal{Z}}(f) \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$. $\forall \tilde{F} \in \mathcal{O}$, $\forall y \in \mathcal{Y}$, $\langle \langle \hat{\mu}(F), y \rangle \rangle = \langle \langle \lim_{(U,V,h) \in \bar{\mathcal{A}}_F} f_h, y \rangle \rangle = \lim_{(U,V,h) \in \bar{\mathcal{A}}_F} \langle \langle f_h, y \rangle \rangle = \lim_{(U,V,h) \in \bar{\mathcal{A}}_F} f(hy) = \lim_{(U,V,h) \in \bar{\mathcal{A}}_F} \int_X \langle \langle d\mu(x), h(x)y \rangle \rangle$. $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists O \in \mathcal{O}$ with $F \subseteq O$ such that $\mathcal{P} \circ \mu(O \setminus F) < \epsilon/(1 + \|y\|)$. $\forall (U, V, h) \in \bar{\mathcal{A}}_F$ with $V \subseteq O$, $\left| \int_X \langle \langle d\mu(x), h(x)y \rangle \rangle - \langle \langle \mu(F), y \rangle \rangle \right| = \left| \int_F \langle \langle d\mu(x), h(x)y \rangle \rangle - \langle \langle \mu(F), y \rangle \rangle + \int_{O \setminus F} \langle \langle d\mu(x), h(x)y \rangle \rangle + \int_{\bar{O}} \langle \langle d\mu(x), h(x)y \rangle \rangle \right| = \left| \langle \langle \mu(F), y \rangle \rangle - \langle \langle \mu(F), y \rangle \rangle + \int_{O \setminus F} \langle \langle d\mu(x), h(x)y \rangle \rangle + 0 \right| \leq \int_{O \setminus F} (\|y\| \mathcal{P} \circ h) d\mathcal{P} \circ \mu \leq \|y\| \mathcal{P} \circ \mu(O \setminus F) < \epsilon$, where the first equality follows from (xi) and Proposition 11.132; the second equality follows from Proposition 11.125; the first inequality follows from Proposition 11.132; and the second inequality follows from Proposition 11.83. This implies that $|\langle \langle \hat{\mu}(F), y \rangle \rangle - \langle \langle \mu(F), y \rangle \rangle| = \lim_{(U,V,h) \in \bar{\mathcal{A}}_F} \left| \int_X \langle \langle d\mu(x), h(x)y \rangle \rangle - \langle \langle \mu(F), y \rangle \rangle \right| \leq \epsilon$. By the arbitrariness of ϵ and y , we have $\hat{\mu}(F) = \mu(F)$. By Proposition 11.194, $\hat{\mu} = \mu$. This shows that $\Phi_{\mathcal{Z}}$ is surjective.

$\forall f \in \mathcal{Z}^*$, let $\mu := \Phi_{\mathcal{Z}}(f) \in \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$. Then, by the above and (x), we have $\|\mu\| = \|\Phi_{\mathcal{Z}}(f)\| \leq \|f\| \leq \|\mu\|$. Hence, $\|\Phi_{\mathcal{Z}}(f)\| = \|f\|$. Then, $\Phi_{\mathcal{Z}}$ is injective. This shows that $\Phi_{\mathcal{Z}}$ is bijective, continuous, linear, and norm preserving. Hence, $\Phi_{\mathcal{Z}} : \mathcal{Z}^* \rightarrow \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$ is an isometrical isomorphism.

This completes the proof of the theorem. $\square$

Definition 11.202 Let $m \in \mathbb{Z}_+$, $I := [0, 1] \subset \mathbb{R}$, $\mathcal{Y}$ be a normed linear space, $f : I^m \rightarrow \mathcal{Y}$, and $n \in \mathbb{N}$. The n th Bernstein function for f , $B_n : I^m \rightarrow \mathcal{Y}$, is defined by, $\forall x := (x_1, \dots, x_m) \in I^m$,

$$B_n(x) = B_n(x; f) = \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n f\left(\frac{k_1}{n}, \dots, \frac{k_m}{n}\right) \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1 - x_i)^{n-k_i}$$

where, for notational consistency, we have made the arbitrary assignment of $0^0 = 1$ in this definition as well as the following theorem.

Theorem 11.203 (Bernsteĭn Approximation Theorem) *Let $m \in \mathbb{Z}_+$, $I := [0, 1] \subset \mathbb{R}$, $\mathcal{Y}$ be a normed linear space, $f : I^m \rightarrow \mathcal{Y}$ be continuous, and $(B_n)_{n=1}^\infty$ be the sequence of Bernsteĭn functions for f . Then, $\lim_{n \in \mathbb{N}} \|f - B_n\|_{\mathcal{C}(I^m, \mathcal{Y})} = 0$.*

Proof $\forall n \in \mathbb{Z}_+, \forall y \in I$, we have

$$1 = (y + 1 - y)^n = \sum_{k=0}^n \binom{n}{k} y^k (1 - y)^{n-k} \quad (11.6)$$

Then, $\forall n \in \mathbb{N}, \forall y \in I$, we have

$$\begin{aligned} y &= y \cdot 1 = y \left(\sum_{k=0}^{n-1} \binom{n-1}{k} y^k (1 - y)^{n-1-k} \right) \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} y^{k+1} (1 - y)^{n-1-k} = \sum_{k=1}^n \binom{n-1}{k-1} y^k (1 - y)^{n-k} \\ &= \sum_{k=1}^n \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k} = \sum_{k=0}^n \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k} \end{aligned} \quad (11.7)$$

where the second equality follows from (11.6). Similarly, $\forall n - 1 \in \mathbb{N}, \forall y \in I$,

$$\begin{aligned} y^2 &= y \left(\sum_{k=0}^{n-1} \frac{k}{n-1} \binom{n-1}{k} y^k (1 - y)^{n-1-k} \right) \\ &= \sum_{k=0}^{n-1} \frac{k}{n-1} \binom{n-1}{k} y^{k+1} (1 - y)^{n-1-k} \\ &= \sum_{k=1}^n \frac{k-1}{n-1} \binom{n-1}{k-1} y^k (1 - y)^{n-k} = \sum_{k=1}^n \frac{k(k-1)}{n(n-1)} \binom{n}{k} y^k (1 - y)^{n-k} \\ &= \sum_{k=0}^n \frac{k(k-1)}{n(n-1)} \binom{n}{k} y^k (1 - y)^{n-k} \end{aligned}$$

where the first equality follows from (11.7). This implies that

$$\begin{aligned} \left(1 - \frac{1}{n}\right) y^2 &= \sum_{k=0}^n \frac{k^2 - k}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} \\ &= \sum_{k=0}^n \frac{k^2}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} - \frac{1}{n} \sum_{k=0}^n \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k} \\ &= \sum_{k=0}^n \frac{k^2}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} - \frac{1}{n} y \end{aligned}$$

where the last equality follows from (11.7). Rearranging terms in the above yields, $\forall y \in I, \forall n-1 \in \mathbb{N}$,

$$\begin{aligned}
\frac{1}{n}y(1-y) &= \sum_{k=0}^n \frac{k^2}{n^2} \binom{n}{k} y^k (1-y)^{n-k} + y^2 - 2y^2 \\
&= \sum_{k=0}^n \frac{k^2}{n^2} \binom{n}{k} y^k (1-y)^{n-k} + y^2 \sum_{k=0}^n \binom{n}{k} y^k (1-y)^{n-k} \\
&\quad - 2y \sum_{k=0}^n \frac{k}{n} \binom{n}{k} y^k (1-y)^{n-k} \\
&= \sum_{k=0}^n (y^2 - 2y \frac{k}{n} + \frac{k^2}{n^2}) \binom{n}{k} y^k (1-y)^{n-k} \\
&= \sum_{k=0}^n (y - \frac{k}{n})^2 \binom{n}{k} y^k (1-y)^{n-k} \tag{11.8}
\end{aligned}$$

where the second equality follows from (11.6) and (11.7). $\forall n \in \mathbb{Z}_+, \forall x := (x_1, \dots, x_m) \in I^m$, we have

$$\begin{aligned}
1 &= \prod_{i=1}^m (x_i + 1 - x_i)^n = \prod_{i=1}^m \left(\sum_{k=0}^n \binom{n}{k} x_i^k (1-x_i)^{n-k} \right) \\
&= \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \tag{11.9}
\end{aligned}$$

Then, $\forall x := (x_1, \dots, x_m) \in I^m, \forall n-1 \in \mathbb{N}$, we have

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^m x_i (1-x_i) &= \sum_{i=1}^m \sum_{k_i=0}^n (x_i - \frac{k_i}{n})^2 \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \\
&= \sum_{i=1}^m \left(\sum_{k_i=0}^n (x_i - \frac{k_i}{n})^2 \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \right) \\
&\quad \cdot \prod_{\substack{j=1 \\ j \neq i}}^m \left(\sum_{k_j=0}^n \binom{n}{k_j} x_j^{k_j} (1-x_j)^{n-k_j} \right) \\
&= \sum_{i=1}^m \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n (x_i - \frac{k_i}{n})^2 \prod_{j=1}^m \binom{n}{k_j} x_j^{k_j} (1-x_j)^{n-k_j} \\
&= \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n \left(\prod_{j=1}^m \binom{n}{k_j} x_j^{k_j} (1-x_j)^{n-k_j} \right) \left(\sum_{i=1}^m (x_i - \frac{k_i}{n})^2 \right) \\
&= \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n \left| x - \left(\frac{k_1}{n}, \dots, \frac{k_m}{n} \right) \right|^2 \prod_{j=1}^m \binom{n}{k_j} x_j^{k_j} (1-x_j)^{n-k_j} \tag{11.10}
\end{aligned}$$

where the first equality follows from (11.8); and the second equality follows from (11.9).

Note that $I^m \subseteq \mathbb{R}^m$ is a compact metric space. By Propositions 5.22, 5.29, 7.21, 3.12, and 3.9, $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|f(x)\| \leq M$, $\forall x \in I^m$. By Proposition 5.39, f is uniformly continuous. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x, \bar{x} \in I^m$ with $|x - \bar{x}| < \delta$, we have $\|f(x) - f(\bar{x})\| < \epsilon/2$. Let $n_0 := \lceil \max\{\delta^{-4}, m^2 M^2 / \epsilon^2, 2\} \rceil \in \mathbb{N}$. $\forall n \in \mathbb{N}$ with $n_0 \leq n$, $\forall x := (x_1, \dots, x_m) \in I^m$, we have

$$\begin{aligned} & \|f(x) - B_n(x)\| \\ &= \left\| f(x) \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \right. \\ &\quad \left. - \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n f\left(\frac{k_1}{n}, \dots, \frac{k_m}{n}\right) \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \right\| \\ &= \left\| \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n (f(x) - f(\frac{k_1}{n}, \dots, \frac{k_m}{n})) \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \right\| \\ &\leq \sum_{k_1=0}^n \cdots \sum_{k_m=0}^n \|f(x) - f(\frac{k_1}{n}, \dots, \frac{k_m}{n})\| \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \end{aligned}$$

where the first equality follows from (11.9) and Definition 11.202. Let $J := \{0, \dots, n\}^m$ and $J_x := \{(k_1, \dots, k_m) \in J \mid |x - (\frac{k_1}{n}, \dots, \frac{k_m}{n})| < n^{-1/4}\}$. Then,

$$\begin{aligned} & \|f(x) - B_n(x)\| \\ &\leq \sum_{(k_1, \dots, k_m) \in J_x} \|f(x) - f(\frac{k_1}{n}, \dots, \frac{k_m}{n})\| \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \\ &\quad + \sum_{(k_1, \dots, k_m) \in J \setminus J_x} \|f(x) - f(\frac{k_1}{n}, \dots, \frac{k_m}{n})\| \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \end{aligned}$$

Note that, $\forall (k_1, \dots, k_m) \in J_x$, $|(x_1, \dots, x_m) - (\frac{k_1}{n}, \dots, \frac{k_m}{n})| < n^{-1/4} \leq \delta$, then $\|f(x) - f(\frac{k_1}{n}, \dots, \frac{k_m}{n})\| < \epsilon/2$. This implies that

$$\begin{aligned} & \sum_{(k_1, \dots, k_m) \in J_x} \|f(x) - f(\frac{k_1}{n}, \dots, \frac{k_m}{n})\| \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \\ &< \frac{\epsilon}{2} \sum_{(k_1, \dots, k_m) \in J_x} \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \\ &\leq \frac{\epsilon}{2} \sum_{(k_1, \dots, k_m) \in J} \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} = \frac{\epsilon}{2} \end{aligned}$$

where the equality follows from (11.9). Note also that, $\forall(k_1, \dots, k_m) \in J \setminus J_x$, $\|f(x) - f(\frac{k_1}{n}, \dots, \frac{k_m}{n})\| \leq 2M = 2M \frac{|x - (\frac{k_1}{n}, \dots, \frac{k_m}{n})|^2}{|x - (\frac{k_1}{n}, \dots, \frac{k_m}{n})|^2} \leq 2M\sqrt{n} |x - (\frac{k_1}{n}, \dots, \frac{k_m}{n})|^2$. Then,

$$\begin{aligned} & \sum_{(k_1, \dots, k_m) \in J \setminus J_x} \|f(x) - f(\frac{k_1}{n}, \dots, \frac{k_m}{n})\| \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \\ & \leq \sum_{(k_1, \dots, k_m) \in J \setminus J_x} 2M\sqrt{n} |x - (\frac{k_1}{n}, \dots, \frac{k_m}{n})|^2 \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \\ & \leq 2M\sqrt{n} \sum_{(k_1, \dots, k_m) \in J} |x - (\frac{k_1}{n}, \dots, \frac{k_m}{n})|^2 \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i} (1-x_i)^{n-k_i} \\ & = 2M\sqrt{n} \frac{1}{n} \sum_{i=1}^m x_i (1-x_i) \leq \frac{Mm}{2\sqrt{n}} \leq \epsilon/2 \end{aligned}$$

Hence, we have $\|f(x) - B_n(x)\| < \epsilon$. By the arbitrariness of $x \in I^m$, we have $\|f - B_n\|_{\mathcal{C}(I^m, \mathcal{Y})} < \epsilon$. Therefore, $\lim_{n \in \mathbb{N}} \|f - B_n\|_{\mathcal{C}(I^m, \mathcal{Y})} = 0$.

This completes the proof of the theorem. $\square$

Theorem 11.204 (Riesz Representation Theorem) *Let $m \in \mathbb{Z}_+$, $I_1, \dots, I_m \subset \mathbb{R}$ be compact intervals, $X := \prod_{i=1}^m I_i \subset \mathbb{R}^m$ with subset topology $\mathcal{O}$, $\mathcal{X} := (X, \mathcal{O})$, $\mathcal{Y}$ be a normed linear space, and $\mathcal{Z} := \mathcal{C}(\mathcal{X}, \mathcal{Y})$. Then, $\mathcal{Z}^* = \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*) = \mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y}^*)$.*

Proof By Tychonoff Theorem 5.47 and Proposition 4.37, $\mathcal{X}$ is a compact metric space. By Proposition 5.5, X is closed in $\mathbb{R}^m$. By Proposition 4.39, $\mathcal{X}$ is a complete metric space. Since $\mathbb{R}^m$ is separable, by Proposition 4.38, $\mathcal{X}$ is separable. Hence, $\mathcal{X}$ is a separable compact complete metric space.

Let $Z := \{z \in \mathcal{Z} \mid z = hy, h \in \mathcal{C}(\mathcal{X}, \mathbb{R}), y \in \mathcal{Y}\}$. We will show that $\mathcal{Z} = \overline{\text{span}(Z)}$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\exists i_0 \in \{1, \dots, m\}$ such that $I_{i_0} = \emptyset$; Case 2: $I_1, \dots, I_m$ are nonempty. Case 1: $\exists i_0 \in \{1, \dots, m\}$ such that $I_{i_0} = \emptyset$. Then, $X = \emptyset$ and $\mathcal{Z}$ is the trivial Banach space with a singleton element. Clearly, $Z = \mathcal{Z}$. Hence, $\mathcal{Z} = \overline{\text{span}(Z)}$.

Case 2: $I_1, \dots, I_m$ are nonempty. Without loss of generality, assume $I_i = [a_i, b_i] \subset \mathbb{R}$ with $a_i, b_i \in \mathbb{R}$ and $a_i \leq b_i$, $i = 1, \dots, m$, and $\exists \bar{m} \in \mathbb{Z}_+$ with $\bar{m} \leq m$ such that $a_i < b_i$, $i = 1, \dots, \bar{m}$, and $a_i = b_i$, $i = \bar{m} + 1, \dots, m$. Let $I := [0, 1] \subset \mathbb{R}$. Then, we may define a homeomorphism $\psi : \mathcal{X} \rightarrow I^{\bar{m}}$ by $\psi(x) = (\frac{x_1 - a_1}{b_1 - a_1}, \dots, \frac{x_{\bar{m}} - a_{\bar{m}}}{b_{\bar{m}} - a_{\bar{m}}})$, $\forall x := (x_1, \dots, x_m) \in X$. $\forall f \in \mathcal{Z}$, let $\bar{f} : I^{\bar{m}} \rightarrow \mathcal{Y}$ be defined by $\bar{f} := f \circ \psi_{\text{inv}}$. By Proposition 3.12, $\bar{f} \in \mathcal{C}(I^{\bar{m}}, \mathcal{Y})$. By Bernstein Approximation Theorem 11.203, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n \in \mathbb{N}$ such that $\|\bar{f} - B_n\|_{\mathcal{C}(I^{\bar{m}}, \mathcal{Y})} < \epsilon$, where $B_n : I^{\bar{m}} \rightarrow \mathcal{Y}$ is the n th Bernstein function

for $\bar{f}$. Then, $\|f - B_n \circ \psi\|_{\mathcal{Z}} = \|(\bar{f} - B_n) \circ \psi\|_{\mathcal{Z}} = \|\bar{f} - B_n\|_{\mathcal{C}(I^m, \mathcal{Y})} < \epsilon$. Clearly, $B_n \circ \psi \in \text{span}(Z)$. By the arbitrariness of ϵ , we have $f \in \overline{\text{span}(Z)}$. By the arbitrariness of f , we have $\mathcal{Z} = \overline{\text{span}(Z)}$.

Hence, in both cases, we have $\mathcal{Z} = \overline{\text{span}(Z)}$. By Riesz Representation Theorem 11.201, we have $\mathcal{Z}^* = \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$. By Theorem 11.198, $\mathcal{Z}^* = \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*) = \mathcal{M}_f(X, \mathcal{B}_{\mathbb{B}}(\mathcal{X}), \mathcal{Y}^*)$. This completes the proof of the theorem. $\square$

Theorem 11.205 (Riesz Representation Theorem) $\mathcal{X} := (X, \mathcal{O})$ be a compact Hausdorff topological space, $\mathcal{Y}$ be a finite dimensional Banach space over $\mathbb{K}$, and $\mathcal{Z} := \mathcal{C}(\mathcal{X}, \mathcal{Y})$. Then, $\mathcal{Z}^* = \mathcal{M}_{ft}(\mathcal{X}, \mathcal{Y}^*)$.

Proof Let $Z := \{z \in \mathcal{Z} \mid z = hy, h \in \mathcal{C}(\mathcal{X}, \mathbb{R}), y \in \mathcal{Y}\}$. We will show that $\mathcal{Z} = \overline{\text{span}(Z)}$. Then, the theorem is a direct consequence of Riesz Representation Theorem 11.201.

Let the dimension of $\mathcal{Y}$ be $m \in \mathbb{Z}_+$ and $\{y_1, \dots, y_m\}$ be a basis of $\mathcal{Y}$. Then, there is an invertible bounded linear mapping $\phi : \mathcal{Y} \rightarrow \mathbb{K}^m$ such that $y = \sum_{i=1}^m (\pi_i \circ \phi(y))y_i$, $\forall y \in \mathcal{Y}$, where $\pi_i : \mathbb{K}^m \rightarrow \mathbb{K}$ is the i th coordinate projection function, $i = 1, \dots, m$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathbb{K} = \mathbb{R}$; Case 2: $\mathbb{K} = \mathbb{C}$.

Case 1: $\mathbb{K} = \mathbb{R}$. $\forall f \in \mathcal{Z}$, let $g_i := \pi_i \circ \phi \circ f \in \mathcal{C}(\mathcal{X}, \mathbb{R})$, $i = 1, \dots, m$. Then, $f(x) = \sum_{i=1}^m (\pi_i \circ \phi(f(x)))y_i = \sum_{i=1}^m g_i(x)y_i$, $\forall x \in \mathcal{X}$. Hence, $f \in \text{span}(Z)$. This case is proved.

Case 2: $\mathbb{K} = \mathbb{C}$. $\forall f \in \mathcal{Z}$, let $g_i := \text{Re} \circ \pi_i \circ \phi \circ f \in \mathcal{C}(\mathcal{X}, \mathbb{R})$ and $h_i := \text{Im} \circ \pi_i \circ \phi \circ f \in \mathcal{C}(\mathcal{X}, \mathbb{R})$, $i = 1, \dots, m$. Then, $f(x) = \sum_{i=1}^m (\pi_i \circ \phi(f(x)))y_i = \sum_{i=1}^m (\text{Re} \circ \pi_i \circ \phi \circ f(x) + i \text{Im} \circ \pi_i \circ \phi \circ f(x))y_i = \sum_{i=1}^m (g_i(x)y_i + h_i(x)iy_i)$, $\forall x \in \mathcal{X}$. Hence, $f \in \text{span}(Z)$. This case is also proved.

This completes the proof of the theorem. $\square$

Proposition 11.206 Let $m \in \mathbb{Z}_+$, $I_1, \dots, I_m \subset \mathbb{R}$ be compact intervals, $X := \prod_{i=1}^m I_i \subset \mathbb{R}^m$ with subset topology $\mathcal{O}$, $\mathcal{X} := (X, \mathcal{O})$, and $\mathcal{Y}$ be a separable normed linear space. Then, $\mathcal{Z} := \mathcal{C}(\mathcal{X}, \mathcal{Y})$ is separable.

Proof Let $Z := \{z \in \mathcal{Z} \mid z = hy, h \in \mathcal{C}(\mathcal{X}, \mathbb{R}), y \in \mathcal{Y}\}$. By Riesz Representation Theorem 11.204 and its proof, we have $\mathcal{Z} = \overline{\text{span}(Z)}$. By Proposition 11.190, $\mathcal{C}(\mathcal{X}, \mathbb{R})$ is separable. Let $D_1 \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ be a countable dense subset. Let $D_2 \subseteq \mathcal{Y}$ be a countable dense subset. Then, $Z_0 := \{z \in Z \mid z = h_1y_1, h_1 \in D_1, y_1 \in D_2\}$ is a countable set. Clearly, $\overline{\text{span}(Z_0)} \supseteq Z$. Then, Z is separable. By Proposition 7.35, $\mathcal{Z}$ is separable. This completes the proof of the proposition. $\square$

Chapter 12

Differentiation and Integration

12.1 Carathéodory Extension Theorem

Definition 12.1 Let X be a set, $\mathcal{A} \subseteq 2^X$ be an algebra on X , $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $\nu : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R}_e$ be a measure on the algebra $\mathcal{A}$. A $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}$ is a function μ from $\mathcal{A}$ to $\mathcal{Y}$ that satisfies

- (i) $\mu(\emptyset) = \vartheta_{\mathcal{Y}}$;
- (ii) $\forall A \in \mathcal{A}$ with $\nu(A) = \infty$, $\mu(A)$ is undefined;
- (iii) $\forall A \in \mathcal{A}$ with $\nu(A) < \infty$, $\mu(A) \in \mathcal{Y}$ and $\forall$ pairwise disjoint $(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ with $A = \bigcup_{i=1}^{\infty} A_i$, we have $\sum_{i=1}^{\infty} \|\mu(A_i)\| < +\infty$ and $\mu(A) = \sum_{i=1}^{\infty} \mu(A_i)$;
- (iv) $\forall A \in \mathcal{A}$ with $\nu(A) < \infty$, we have

$$\nu(A) = \sup_{n \in \mathbb{Z}_+, (A_i)_{i=1}^n \subseteq \mathcal{A}, A = \bigcup_{i=1}^n A_i, A_i \cap A_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(A_i)\|$$

Then, ν is said to be the total variation of μ and denoted by $\mathcal{P} \circ \mu$. μ is said to be finite if ν is finite; and μ is said to be σ -finite if ν is σ -finite.

Proposition 12.2 Let X be a set, $\mathcal{A} \subseteq 2^X$ be an algebra on X , $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, μ be a $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}$, and $E \in \mathcal{A}$. Then, $\mathcal{A}_E := \{C \in \mathcal{A} \mid C \subseteq E\}$ is an algebra on E , $\mu_E := \mu|_{\mathcal{A}_E}$ is a $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}_E$, and $\mathcal{P} \circ \mu_E = (\mathcal{P} \circ \mu)|_{\mathcal{A}_E}$.

Proof It is easy to show that $\mathcal{A}_E$ is an algebra on E . Note that $\mathcal{P} \circ \mu$ is a measure on the algebra $\mathcal{A}$. Define $\nu_E := (\mathcal{P} \circ \mu)|_{\mathcal{A}_E}$. Then, $\nu_E(\emptyset) = \mathcal{P} \circ \mu(\emptyset) = 0$. $\forall$ pairwise disjoint $(A_i)_{i=1}^\infty \subseteq \mathcal{A}_E$ with $A := \bigcup_{i=1}^\infty A_i \in \mathcal{A}_E$, we have $\nu_E(A) = \mathcal{P} \circ \mu(A) = \sum_{i=1}^\infty \mathcal{P} \circ \mu(A_i) = \sum_{i=1}^\infty \nu_E(A_i)$. Then, ν_E is a measure on the algebra $\mathcal{A}_E$.

$\mu_E(\emptyset) = \mu(\emptyset) = \vartheta_{\mathcal{Y}}$. $\forall A \in \mathcal{A}_E$ with $\nu_E(A) = \infty$, we have $\mathcal{P} \circ \mu(A) = \infty$, $\mu(A)$ is undefined, and therefore, $\mu_E(A)$ is undefined. $\forall A \in \mathcal{A}_E$ with $\nu_E(A) < \infty$, we have $\mathcal{P} \circ \mu(A) < \infty$ and $\mu_E(A) = \mu(A) \in \mathcal{Y}$. $\forall$ pairwise disjoint $(A_i)_{i=1}^\infty \subseteq \mathcal{A}_E$ with $A = \bigcup_{i=1}^\infty A_i$, we have $\sum_{i=1}^\infty \|\mu_E(A_i)\| = \sum_{i=1}^\infty \|\mu(A_i)\| < \infty$ and $\sum_{i=1}^\infty \mu_E(A_i) = \sum_{i=1}^\infty \mu(A_i) = \mu(A) = \mu_E(A)$. $\forall A \in \mathcal{A}_E$ with $\nu_E(A) < \infty$, we have $\nu_E(A) = \mathcal{P} \circ \mu(A) = \sup_{n \in \mathbb{Z}_+, (A_i)_{i=1}^n \subseteq \mathcal{A}, A = \bigcup_{i=1}^n A_i, A_i \cap A_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(A_i)\| = \sup_{n \in \mathbb{Z}_+, (A_i)_{i=1}^n \subseteq \mathcal{A}_E, A = \bigcup_{i=1}^n A_i, A_i \cap A_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu_E(A_i)\|$. Hence, μ_E is a $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}_E$ with $\mathcal{P} \circ \mu_E = \mathcal{P} \circ (\mu|_{\mathcal{A}_E}) = \nu_E = (\mathcal{P} \circ \mu)|_{\mathcal{A}_E}$.

This completes the proof of the proposition. $\square$

Proposition 12.3 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space. Define $\rho : \mathcal{B} \times \mathcal{B} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\rho(E_1, E_2) = \mu(E_1 \triangle E_2)$, $\forall E_1, E_2 \in \mathcal{B}$. Then, ρ defines a pseudo-metric on $\mathcal{B}$.

Proof $\forall E_1, E_2, E_3 \in \mathcal{B}$, $\rho(E_1, E_2) = \mu(E_1 \triangle E_2) \in [0, +\infty) \subset \mathbb{R}$ since μ is a finite measure on $(X, \mathcal{B})$. Clearly, $\rho(E_1, E_1) = \mu(E_1 \triangle E_1) = \mu(\emptyset) = 0$ and $\rho(E_1, E_2) = \mu(E_1 \triangle E_2) = \mu(E_2 \triangle E_1) = \rho(E_2, E_1)$. Note that $\rho(E_1, E_2) + \rho(E_2, E_3) = \mu(E_1 \triangle E_2) + \mu(E_2 \triangle E_3) = \mu(E_1 \setminus E_2) + \mu(E_2 \setminus E_1) + \mu(E_2 \setminus E_3) + \mu(E_3 \setminus E_2) \geq \mu((E_1 \setminus E_2) \cup (E_2 \setminus E_3)) + \mu((E_2 \setminus E_1) \cup (E_3 \setminus E_2))$. Clearly, we have $(E_1 \setminus E_2) \cup (E_2 \setminus E_3) \supseteq E_1 \setminus E_3$ and $(E_2 \setminus E_1) \cup (E_3 \setminus E_2) \supseteq E_3 \setminus E_1$. Then, $\rho(E_1, E_2) + \rho(E_2, E_3) \geq \mu(E_1 \setminus E_3) + \mu(E_3 \setminus E_1) = \mu(E_1 \triangle E_3) = \rho(E_1, E_3)$. Hence, ρ defines a pseudo-metric on $\mathcal{B}$. This completes the proof of the proposition. $\square$

Theorem 12.4 (Carathéodory Extension Theorem) Let $\mathcal{A}$ be an algebra on a set X , $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, μ be a σ -finite $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}$, and $\mathcal{B}$ be the σ -algebra on X generated by $\mathcal{A}$. Then, there is a unique $\mathcal{Y}$ -valued measure $\bar{\mu}$ on $(X, \mathcal{B})$ such that $\bar{\mu}|_{\mathcal{A}} = \mu$ and $(\mathcal{P} \circ \bar{\mu})|_{\mathcal{A}} = \mathcal{P} \circ \mu$. Furthermore, $\bar{\mu}$ is σ -finite.

Proof We first consider the special case where μ is finite. Then, $\mu : \mathcal{A} \rightarrow \mathcal{Y}$ and $\mathcal{P} \circ \mu : \mathcal{A} \rightarrow [0, \mathcal{P} \circ \mu(X)] \subset \mathbb{R}$. By Carathéodory Extension Theorem 11.19, there is a unique finite measure $\nu : \mathcal{B} \rightarrow [0, \mathcal{P} \circ \mu(X)] \subset \mathbb{R}$ on the measurable space $(X, \mathcal{B})$ such that $\nu(A) = \mathcal{P} \circ \mu(A)$, $\forall A \in \mathcal{A}$, and $\nu(E) = \inf_{(A_i)_{i=1}^\infty \subseteq \mathcal{A}, \bigcup_{i=1}^\infty A_i \supseteq E} \sum_{i=1}^\infty \mathcal{P} \circ \mu(A_i)$, $\forall E \in \mathcal{B}$. Define $\rho : \mathcal{B} \times \mathcal{B} \rightarrow [0, \mathcal{P} \circ \mu(X)] \subset \mathbb{R}$ by $\rho(E_1, E_2) = \nu(E_1 \triangle E_2)$, $\forall E_1, E_2 \in \mathcal{B}$. By Proposition 12.3, ρ defines a pseudo-metric on $\mathcal{B}$. Define an equivalence relation $\equiv$ on $\mathcal{B}$ by $E_1 \equiv E_2$ if $\rho(E_1, E_2) = \nu(E_1 \triangle E_2) = 0$. By Lemma 4.48, $(\mathcal{B}/\equiv, \bar{\rho})$ is a metric space, where $\bar{\rho} : (\mathcal{B}/\equiv) \times (\mathcal{B}/\equiv) \rightarrow [0, \mathcal{P} \circ \mu(X)] \subset \mathbb{R}$

is defined by $\bar{\rho}([E_1], [E_2]) = \rho(E_1, E_2)$, $\forall [E_1], [E_2] \in \mathcal{B}/ \equiv$ with $E_1, E_2 \in \mathcal{B}$. Let $\mathcal{A}_{\equiv} := \{[A] \in \mathcal{B}/ \equiv \mid A \in \mathcal{A}\}$.

Claim 12.4.1 $\mathcal{A}_{\equiv}$ is dense in $\mathcal{B}/ \equiv$, that is $\overline{\mathcal{A}_{\equiv}} = \mathcal{B}/ \equiv$.

Proof of claim: $\forall [E] \in \mathcal{B}/ \equiv$ with $E \in \mathcal{B}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ with $E \subseteq \bigcup_{i=1}^{\infty} A_i$ such that $\nu(E) \leq \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(A_i) < \nu(E) + \epsilon/2 < +\infty$. Then, $\exists n_0 \in \mathbb{N}$ such that $\sum_{i=n_0+1}^{\infty} \mathcal{P} \circ \mu(A_i) < \epsilon/2$. Let $A := \bigcup_{i=1}^{n_0} A_i \in \mathcal{A}$. Note that $E \setminus A \subseteq (\bigcup_{i=1}^{\infty} A_i) \setminus A \subseteq \bigcup_{i=n_0+1}^{\infty} A_i$. This implies that $0 \leq \nu(E \setminus A) \leq \nu(\bigcup_{i=n_0+1}^{\infty} A_i) \leq \sum_{i=n_0+1}^{\infty} \nu(A_i) = \sum_{i=n_0+1}^{\infty} \mathcal{P} \circ \mu(A_i) < \epsilon/2$. Note also that $A \setminus E \subseteq (\bigcup_{i=1}^{\infty} A_i) \setminus E$. This further implies that $0 \leq \nu(A \setminus E) \leq \nu((\bigcup_{i=1}^{\infty} A_i) \setminus E) = \nu(\bigcup_{i=1}^{\infty} A_i) - \nu(E) \leq \sum_{i=1}^{\infty} \nu(A_i) - \nu(E) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(A_i) - \nu(E) < \epsilon/2$. Hence, $\bar{\rho}([E], [A]) = \rho(E, A) = \nu(E \triangle A) = \nu(E \setminus A) + \nu(A \setminus E) < \epsilon$. Hence, $\overline{\mathcal{A}_{\equiv}} = \mathcal{B}/ \equiv$. This completes the proof of the claim. $\square$

Define $g : \mathcal{A}_{\equiv} \rightarrow \mathcal{Y}$ by $g([A]) = \mu(A)$, $\forall [A] \in \mathcal{A}_{\equiv}$ with $A \in \mathcal{A}$. We will show that g is well defined. $\forall A_1, A_2 \in \mathcal{A}$ with $A_1 \equiv A_2$, we have $\rho(A_1, A_2) = \nu(A_1 \triangle A_2) = \mathcal{P} \circ \mu(A_1 \triangle A_2) = 0$. Then, $\mathcal{P} \circ \mu(A_1 \setminus A_2) = 0 = \mathcal{P} \circ \mu(A_2 \setminus A_1)$. This implies that $\mu(A_1 \setminus A_2) = \vartheta_{\mathcal{Y}} = \mu(A_2 \setminus A_1)$ and $\mu(A_1) = \mu(A_1 \cap A_2) + \mu(A_1 \setminus A_2) = \mu(A_1 \cap A_2) + \mu(A_2 \setminus A_1) = \mu(A_2) \in \mathcal{Y}$. Hence, g is well defined.

Claim 12.4.2 g is uniformly continuous.

Proof of claim: $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta = \epsilon$. $\forall [A_1], [A_2] \in \mathcal{A}_{\equiv}$ with $A_1, A_2 \in \mathcal{A}$ and $\bar{\rho}([A_1], [A_2]) < \delta$, we have $\|g([A_1]) - g([A_2])\| = \|\mu(A_1) - \mu(A_2)\| = \|\mu(A_1 \cap A_2) + \mu(A_1 \setminus A_2) - \mu(A_2 \cap A_1) - \mu(A_2 \setminus A_1)\| = \|\mu(A_1 \setminus A_2) - \mu(A_2 \setminus A_1)\| \leq \|\mu(A_1 \setminus A_2)\| + \|\mu(A_2 \setminus A_1)\| \leq \mathcal{P} \circ \mu(A_1 \setminus A_2) + \mathcal{P} \circ \mu(A_2 \setminus A_1) = \mathcal{P} \circ \mu(A_1 \triangle A_2) = \nu(A_1 \triangle A_2) = \rho(A_1, A_2) = \bar{\rho}([A_1], [A_2]) < \delta = \epsilon$. Hence, g is uniformly continuous. This completes the proof of the claim. $\square$

By Proposition 4.46 and the completeness of $\mathcal{Y}$, there exists a unique continuous function $\bar{g} : \mathcal{B}/ \equiv \rightarrow \mathcal{Y}$ such that $\bar{g}|_{\mathcal{A}_{\equiv}} = g$, and $\bar{g}$ is uniformly continuous. Define $\bar{\mu} : \mathcal{B} \rightarrow \mathcal{Y}$ by $\bar{\mu}(E) = \bar{g}([E])$, $\forall E \in \mathcal{B}$. We will show that $\bar{\mu}$ is the $\mathcal{Y}$ -valued measure we seek. Clearly, $\bar{\mu}(A) = \bar{g}([A]) = g([A]) = \mu(A)$, $\forall A \in \mathcal{A}$. Then, $\bar{\mu}(\emptyset) = \mu(\emptyset) = \vartheta_{\mathcal{Y}}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the uniform continuity of $\bar{g}$, $\exists \delta(\epsilon) \in (0, \epsilon] \subset \mathbb{R}$ such that, $\forall [E_1], [E_2] \in \mathcal{B}/ \equiv$ with $E_1, E_2 \in \mathcal{B}$ and $\bar{\rho}([E_1], [E_2]) = \rho(E_1, E_2) = \nu(E_1 \triangle E_2) < \delta(\epsilon)$, we have $\|\bar{\mu}(E_1) - \bar{\mu}(E_2)\| = \|\bar{g}([E_1]) - \bar{g}([E_2])\| < \epsilon$.

Claim 12.4.3 $\bar{\mu}$ is finitely additive, that is, $\bar{\mu}(E_1 \cup E_2) = \bar{\mu}(E_1) + \bar{\mu}(E_2)$, $\forall E_1, E_2 \in \mathcal{B}$ with $E_1 \cap E_2 = \emptyset$.

Proof of claim: $\forall E_1, E_2 \in \mathcal{B}$ with $E_1 \cap E_2 = \emptyset$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Claim 12.4.1, $\exists A_1, A_2 \in \mathcal{A}$ such that $\bar{\rho}([A_i], [E_i]) = \rho(A_i, E_i) = \nu(A_i \triangle E_i) < \delta(\epsilon/4)/2$, $i = 1, 2$. Then, $\|\bar{\mu}(E_i) - \bar{\mu}(A_i)\| < \epsilon/4$, $i = 1, 2$. Note that

$(E_1 \cup E_2) \setminus (A_1 \cup A_2) = (E_1 \cup E_2) \cap \widetilde{A_1} \cap \widetilde{A_2} = (E_1 \cap \widetilde{A_1} \cap \widetilde{A_2}) \cup (E_2 \cap \widetilde{A_1} \cap \widetilde{A_2}) \subseteq$
 $(E_1 \setminus A_1) \cup (E_2 \setminus A_2)$ and $(A_1 \cup A_2) \setminus (E_1 \cup E_2) \subseteq (A_1 \setminus E_1) \cup (A_2 \setminus E_2)$. This
leads to $(E_1 \cup E_2) \triangle (A_1 \cup A_2) \subseteq (E_1 \setminus A_1) \cup (E_2 \setminus A_2) \cup (A_1 \setminus E_1) \cup (A_2 \setminus E_2) =$
 $(E_1 \triangle A_1) \cup (E_2 \triangle A_2)$. This implies that $\bar{\rho}([E_1 \cup E_2], [A_1 \cup A_2]) =$
 $\nu((E_1 \cup E_2) \triangle (A_1 \cup A_2)) \leq \nu(E_1 \triangle A_1) + \nu(E_2 \triangle A_2) < \delta(\epsilon/4)$. Hence,
 $\|\bar{\mu}(E_1 \cup E_2) - \bar{\mu}(A_1 \cup A_2)\| < \epsilon/4$. Note also that $(A_1 \cap A_2) \triangle (E_1 \cap E_2) =$
 $((A_1 \cap A_2) \setminus (E_1 \cap E_2)) \cup ((E_1 \cap E_2) \setminus (A_1 \cap A_2)) = (A_1 \cap A_2 \cap \widetilde{E_1}) \cup (A_1 \cap A_2 \cap$
 $\widetilde{E_2}) \cup (E_1 \cap E_2 \cap \widetilde{A_1}) \cup (E_1 \cap E_2 \cap \widetilde{A_2}) \subseteq (A_1 \setminus E_1) \cup (A_2 \setminus E_2) \cup (E_1 \setminus A_1) \cup$
 $(E_2 \setminus A_2) = (E_1 \triangle A_1) \cup (E_2 \triangle A_2)$. This leads to $\bar{\rho}([E_1 \cap E_2], [A_1 \cap A_2]) =$
 $\bar{\rho}(\emptyset, [A_1 \cap A_2]) = \nu((E_1 \cap E_2) \triangle (A_1 \cap A_2)) \leq \nu((E_1 \triangle A_1) \cup (E_2 \triangle A_2)) \leq$
 $\nu(E_1 \triangle A_1) + \nu(E_2 \triangle A_2) < \delta(\epsilon/4)$. This implies that $\|\bar{\mu}(A_1 \cap A_2) - \bar{\mu}(E_1 \cap E_2)\| < \epsilon/4$. Therefore,
we have $\|\bar{\mu}(E_1 \cup E_2) - \bar{\mu}(E_1) - \bar{\mu}(E_2)\| \leq \|\bar{\mu}(E_1 \cup E_2) - \bar{\mu}(A_1 \cup A_2)\| +$
 $\|\bar{\mu}(A_1 \cup A_2) - \bar{\mu}(A_1) - \bar{\mu}(A_2)\| + \|\bar{\mu}(A_1) - \bar{\mu}(E_1)\| + \|\bar{\mu}(A_2) - \bar{\mu}(E_2)\| < \epsilon$.
By the arbitrariness of ϵ , we have $\bar{\mu}(E_1 \cup E_2) = \bar{\mu}(E_1) + \bar{\mu}(E_2)$. Hence, $\bar{\mu}$
is finitely additive. This completes the proof of the claim. $\square$

Claim 12.4.4 $\forall E \in \mathcal{B}, \|\bar{\mu}(E)\| \leq \nu(E)$.

Proof of claim: $\forall E \in \mathcal{B}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Claim 12.4.1, $\exists A \in \mathcal{A}$
such that $\bar{\rho}([E], [A]) = \nu(E \triangle A) < \delta(\epsilon/2) \leq \epsilon/2$. Then, $\|\bar{\mu}(E) - \bar{\mu}(A)\| <$
 $\epsilon/2$. Note that $\|\bar{\mu}(A)\| = \|\mu(A)\| \leq \mathcal{P} \circ \mu(A) = \nu(A)$. Then, $\|\bar{\mu}(E)\| \leq$
 $\|\bar{\mu}(E) - \bar{\mu}(A)\| + \|\bar{\mu}(A)\| < \epsilon/2 + \nu(A) \leq \epsilon/2 + \nu(A \cap E) + \nu(A \setminus E) \leq$
 $\epsilon/2 + \nu(E) + \nu(A \triangle E) < \nu(E) + \epsilon$. By the arbitrariness of ϵ , we have
 $\|\bar{\mu}(E)\| \leq \nu(E)$. This completes the proof of the claim. $\square$

$\forall$ pairwise disjoint $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$, let $E := \bigcup_{n=1}^\infty E_n \in \mathcal{B}$. By
Claim 12.4.4, we have $\sum_{n=1}^\infty \|\bar{\mu}(E_n)\| \leq \sum_{n=1}^\infty \nu(E_n) = \nu(E) < +\infty$. Let
 $\bar{E}_n := \bigcup_{i=1}^n E_i \in \mathcal{B}, \forall n \in \mathbb{N}$. By Claim 12.4.3, $\bar{\mu}(\bar{E}_n) = \sum_{i=1}^n \bar{\mu}(E_i)$. We
will show that $\lim_{n \in \mathbb{N}} \bar{\mu}(\bar{E}_n) = \bar{\mu}(E)$, which then implies that $\bar{\mu}$ is countably
additive. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n_0 \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have
 $0 \leq \sum_{i=n+1}^\infty \nu(E_i) < \delta(\epsilon)$. Then, $\bar{\rho}([\bar{E}_n], [E]) = \rho(\bar{E}_n, E) = \nu(\bar{E}_n \triangle E) =$
 $\nu(E \setminus \bar{E}_n) = \sum_{i=n+1}^\infty \nu(E_i) < \delta(\epsilon)$. This implies that $\|\bar{\mu}(E) - \bar{\mu}(\bar{E}_n)\| < \epsilon$.
Hence, $\bar{\mu}(E) = \lim_{n \in \mathbb{N}} \bar{\mu}(\bar{E}_n) = \sum_{n=1}^\infty \bar{\mu}(E_n)$. Therefore, $\bar{\mu}$ is countably
additive. This shows that $\bar{\mu}$ is a $\mathcal{Y}$ -valued pre-measure on $(X, \mathcal{B})$.

$\forall E \in \mathcal{B}, \forall n \in \mathbb{Z}_+, \forall$ pairwise disjoint $(E_i)_{i=1}^n \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^n E_i$,
by Claim 12.4.4, we have $\sum_{i=1}^n \|\bar{\mu}(E_i)\| \leq \sum_{i=1}^n \nu(E_i) = \nu(E)$. Therefore,
 $\mathcal{P} \circ \bar{\mu}(E) \leq \nu(E)$. By Proposition 11.100, $\mathcal{P} \circ \bar{\mu}$ is a measure on $(X, \mathcal{B})$.
By the above argument, $\mathcal{P} \circ \bar{\mu}(X) \leq \nu(X) = \mathcal{P} \circ \mu(X) < +\infty$. Hence,
 $\mathcal{P} \circ \bar{\mu}$ is a finite measure on $(X, \mathcal{B})$. $\forall A \in \mathcal{A}$, by Definitions 11.108 and 12.1,
 $\nu(A) = \mathcal{P} \circ \mu(A) \leq \mathcal{P} \circ \bar{\mu}(A) \leq \nu(A)$. Then, $\mathcal{P} \circ \bar{\mu}(A) = \nu(A) = \mathcal{P} \circ \mu(A)$,
 $\forall A \in \mathcal{A}$. By Carathéodory Extension Theorem 11.19, $\mathcal{P} \circ \bar{\mu} = \nu$.

To show the uniqueness of $\bar{\mu}$, let $\tilde{\mu}$ be another $\mathcal{Y}$ -valued measure on
 $(X, \mathcal{B})$ such that $\tilde{\mu}|_{\mathcal{A}} = \mu$ and $(\mathcal{P} \circ \tilde{\mu})|_{\mathcal{A}} = \mathcal{P} \circ \mu$. By Carathéodory Ex-
tension Theorem 11.19, $\mathcal{P} \circ \tilde{\mu} = \nu$. Hence, $\tilde{\mu}$ is a finite $\mathcal{Y}$ -valued measure

and $\tilde{\mu} : \mathcal{B} \rightarrow \mathcal{Y}$. Define $\tilde{g} : \mathcal{B}/ \equiv \rightarrow \mathcal{Y}$ by $\tilde{g}([E]) = \tilde{\mu}(E)$, $\forall [E] \in \mathcal{B}/ \equiv$ with $E \in \mathcal{B}$. We need to show that $\tilde{g}$ is well defined. $\forall E_1, E_2 \in \mathcal{B}$ with $E_1 \equiv E_2$, we have $\rho(E_1, E_2) = \nu(E_1 \triangle E_2) = 0$. This leads to $\nu(E_1 \setminus E_2) = 0 = \nu(E_2 \setminus E_1)$ and $\tilde{\mu}(E_1 \setminus E_2) = \vartheta_{\mathcal{Y}} = \tilde{\mu}(E_2 \setminus E_1)$. Then, $\tilde{\mu}(E_1) = \tilde{\mu}(E_1 \cap E_2) + \tilde{\mu}(E_1 \setminus E_2) = \tilde{\mu}(E_1 \cap E_2) + \tilde{\mu}(E_2 \setminus E_1) = \tilde{\mu}(E_2)$. Hence, $\tilde{g}$ is well defined. Clearly, $\tilde{g}([A]) = \tilde{\mu}(A) = \mu(A) = g([A])$, $\forall [A] \in \mathcal{A}_{\equiv}$ with $A \in \mathcal{A}$. Hence, $\tilde{g}|_{\mathcal{A}_{\equiv}} = g$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta = \epsilon$. $\forall [E_1], [E_2] \in \mathcal{B}/ \equiv$ with $E_1, E_2 \in \mathcal{B}$ and $\bar{\rho}([E_1], [E_2]) = \nu(E_1 \triangle E_2) < \delta$, we have $\|\tilde{g}([E_1]) - \tilde{g}([E_2])\| = \|\tilde{\mu}(E_1) - \tilde{\mu}(E_2)\| = \|\tilde{\mu}(E_1 \cap E_2) + \tilde{\mu}(E_1 \setminus E_2) - \tilde{\mu}(E_2 \cap E_1) - \tilde{\mu}(E_2 \setminus E_1)\| \leq \|\tilde{\mu}(E_1 \setminus E_2)\| + \|\tilde{\mu}(E_2 \setminus E_1)\| \leq \mathcal{P} \circ \tilde{\mu}(E_1 \setminus E_2) + \mathcal{P} \circ \tilde{\mu}(E_2 \setminus E_1) = \mathcal{P} \circ \tilde{\mu}(E_1 \triangle E_2) = \nu(E_1 \triangle E_2) < \delta = \epsilon$. Hence, $\tilde{g}$ is uniformly continuous. By Proposition 4.46, $\tilde{g} = \bar{g}$. Then, $\tilde{\mu}(E) = \tilde{g}([E]) = \bar{g}([E]) = \bar{\mu}(E)$, $\forall E \in \mathcal{B}$. Hence, $\tilde{\mu} = \bar{\mu}$. This establishes the uniqueness and completes the proof for the special case.

Next, we consider the general case where μ is σ -finite. Then, $\exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{A}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mathcal{P} \circ \mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $(X_n)_{n=1}^{\infty}$ is pairwise disjoint. Fix any $n \in \mathbb{N}$. Let $\mathcal{A}_n := \{E \in \mathcal{A} \mid E \subseteq X_n\}$ and $\mu_n := \mu|_{\mathcal{A}_n}$. By Proposition 12.2, μ_n is a finite $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}_n$ and $\mathcal{P} \circ \mu_n = (\mathcal{P} \circ \mu)|_{\mathcal{A}_n}$. By the special case, there exists a unique $\mathcal{Y}$ -valued measure $\bar{\mu}_n$ on $(X_n, \mathcal{B}_n)$, where $\mathcal{B}_n$ is the σ -algebra generated by $\mathcal{A}_n$, such that $\bar{\mu}_n(A) = \mu_n(A)$ and $\mathcal{P} \circ \bar{\mu}_n(A) = \mathcal{P} \circ \mu_n(A)$, $\forall A \in \mathcal{A}_n$. Furthermore, $\bar{\mu}_n$ is finite. By Proposition 11.118, the generation process on $(\mathcal{X}_n := (X_n, \mathcal{B}_n, \bar{\mu}_n))_{n=1}^{\infty}$ yields a unique σ -finite $\mathcal{Y}$ -valued measure space $\mathcal{X} := (X, \bar{\mathcal{B}}, \bar{\mu})$ on X , where $\bar{\mathcal{B}} := \{E \subseteq X \mid E \cap X_n \in \mathcal{B}_n, \forall n \in \mathbb{N}\}$, $\forall E \in \bar{\mathcal{B}}, \mathcal{P} \circ \bar{\mu}(E) = \sum_{n=1}^{\infty} \mathcal{P} \circ \bar{\mu}_n(E \cap X_n)$, $\bar{\mu}(E)$ is undefined if $\mathcal{P} \circ \bar{\mu}(E) = \infty$, $\bar{\mu}(E) = \sum_{n=1}^{\infty} \bar{\mu}_n(E \cap X_n) \in \mathcal{Y}$ if $\mathcal{P} \circ \bar{\mu}(E) < \infty$, and $\mathcal{X}_n$ is the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$, $\forall n \in \mathbb{N}$.

Claim 12.4.5 $\mathcal{B} = \bar{\mathcal{B}}$.

Proof of claim: $\forall A \in \mathcal{A}, \forall n \in \mathbb{N}, A \cap X_n \in \mathcal{A}_n \subseteq \mathcal{B}_n$. This implies that $A \in \bar{\mathcal{B}}$. Then, $\mathcal{A} \subseteq \bar{\mathcal{B}}$. Since $\bar{\mathcal{B}}$ is a σ -algebra, then $\mathcal{B} \subseteq \bar{\mathcal{B}}$. On the other hand, $\forall E \in \bar{\mathcal{B}}, \forall n \in \mathbb{N}$, we have $E \cap X_n \in \mathcal{B}_n \subseteq \mathcal{B}$. Then, $E = \bigcup_{n=1}^{\infty} (E \cap X_n) \in \mathcal{B}$ and, hence, $\bar{\mathcal{B}} \subseteq \mathcal{B}$. This shows that $\mathcal{B} = \bar{\mathcal{B}}$ and completes the proof of the claim. $\square$

$\forall A \in \mathcal{A}, \forall n \in \mathbb{N}$, we have $A \cap X_n \in \mathcal{A}_n \subseteq \mathcal{B}_n$ and $\mathcal{P} \circ \bar{\mu}(A) = \sum_{n=1}^{\infty} \mathcal{P} \circ \bar{\mu}_n(A \cap X_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(A \cap X_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu(A \cap X_n) = \mathcal{P} \circ \mu(A)$. Hence, $(\mathcal{P} \circ \bar{\mu})|_{\mathcal{A}} = \mathcal{P} \circ \mu$. $\forall A \in \mathcal{A}$ with $\mathcal{P} \circ \bar{\mu}(A) = \infty$, then $\bar{\mu}(A)$ is undefined and $\mathcal{P} \circ \mu(A) = \infty$, which implies that $\mu(A)$ is undefined. $\forall A \in \mathcal{A}$ with $\mathcal{P} \circ \bar{\mu}(A) < \infty$, then $\mathcal{P} \circ \mu(A) < \infty$ and $\bar{\mu}(A) = \sum_{n=1}^{\infty} \bar{\mu}_n(A \cap X_n) = \sum_{n=1}^{\infty} \mu_n(A \cap X_n) = \sum_{n=1}^{\infty} \mu(A \cap X_n) = \mu(A) \in \mathcal{Y}$. Hence, $\bar{\mu}|_{\mathcal{A}} = \mu$. This shows that $\bar{\mu}$ is the σ -finite $\mathcal{Y}$ -valued measure on $(X, \mathcal{B})$ that we seek.

This completes the proof of the theorem. $\square$

Definition 12.5 Let X be a set, $\mathcal{A} \subseteq 2^X$ be an algebra on X , $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $\mu : \mathcal{A} \rightarrow \mathcal{Y}$. μ is said to be a $\mathcal{Y}$ -valued pre-measure on the algebra $\mathcal{A}$ if

$$(i) \quad \mu(\emptyset) = \vartheta_{\mathcal{Y}};$$

$$(ii) \quad \forall \text{ pairwise disjoint } (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \text{ with } A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}, \text{ we have } \sum_{i=1}^{\infty} \|\mu(A_i)\| < \infty \text{ and } \mu(A) = \sum_{i=1}^{\infty} \mu(A_i) \in \mathcal{Y}.$$

Define $\mathcal{P} \circ \mu : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R}_e$ by, $\forall A \in \mathcal{A}$, $\mathcal{P} \circ \mu(A) := \sup_{n \in \mathbb{Z}_+, (A_i)_{i=1}^n \subseteq \mathcal{A}, A = \bigcup_{i=1}^n A_i, A_i \cap A_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(A_i)\|$. $\mathcal{P} \circ \mu$ is said to be the total variation of μ .

μ is said to be finite if $\mathcal{P} \circ \mu(X) < \infty$.

Proposition 12.6 Let X be a set, $\mathcal{A} \subseteq 2^X$ be an algebra on X , $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $\mu : \mathcal{A} \rightarrow \mathcal{Y}$ be a $\mathcal{Y}$ -valued pre-measure on the algebra $\mathcal{A}$. Then, $\mathcal{P} \circ \mu : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R}_e$ defines a measure on the algebra $\mathcal{A}$.

Proof Clearly, $\mathcal{P} \circ \mu(\emptyset) = 0$. $\forall$ pairwise disjoint $(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ with $A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, $\forall i \in \mathbb{N}$ with $\mathcal{P} \circ \mu(A_i) > 0$, $\forall 0 \leq t_i < \mathcal{P} \circ \mu(A_i)$, $\exists n_i \in \mathbb{Z}_+$ and $\exists$ pairwise disjoint $(A_{i,j})_{j=1}^{n_i} \subseteq \mathcal{A}$ with $A_i = \bigcup_{j=1}^{n_i} A_{i,j}$ such that $t_i < \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| \leq \mathcal{P} \circ \mu(A_i)$. $\forall i \in \mathbb{N}$ with $\mathcal{P} \circ \mu(A_i) = 0$, let $t_i = 0$, $n_i = 1$, $A_{i,1} = A_i$. Then, $0 = t_i \leq \|\mu(A_{i,1})\| \leq \mathcal{P} \circ \mu(A_i) = 0$.

Claim 12.6.1 $0 \leq t := \sum_{i=1}^{\infty} t_i \leq \mathcal{P} \circ \mu(A)$.

Proof of claim: Clearly, $t \geq 0$ since $t_i \geq 0$, $\forall i \in \mathbb{N}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $t < +\infty$; Case 2: $t = +\infty$.

Case 1: $t < +\infty$. Then, $t \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| \geq 0$. Therefore, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $t - \epsilon < \sum_{i=1}^N \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| \leq \sum_{i=1}^N \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| + \left\| \mu(A \setminus \left(\bigcup_{i=1}^N \bigcup_{j=1}^{n_i} A_{i,j} \right)) \right\| \leq \mathcal{P} \circ \mu(A)$. By the arbitrariness of ϵ , we have $t \leq \mathcal{P} \circ \mu(A)$.

Case 2: $t = +\infty$. Then $t = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| = \infty$. $\forall M \in [0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that

$$\begin{aligned} M &\leq \sum_{i=1}^N \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| \\ &\leq \sum_{i=1}^N \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| + \left\| \mu(A \setminus \left(\bigcup_{i=1}^N \bigcup_{j=1}^{n_i} A_{i,j} \right)) \right\| \leq \mathcal{P} \circ \mu(A) \end{aligned}$$

This implies that $\mathcal{P} \circ \mu(A) = +\infty = t$.

This completes the proof of the claim. $\square$

Hence, $\sum_{i=1}^{\infty} t_i \leq \mathcal{P} \circ \mu(A)$. By the arbitrariness of t_i 's, we have $\sum_{i=1}^{\infty} \mathcal{P} \circ \mu(A_i) \leq \mathcal{P} \circ \mu(A)$.

On the other hand, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_j)_{j=1}^n \subseteq \mathcal{A}$ with $A = \bigcup_{j=1}^n E_j$, we have

$$\begin{aligned} \sum_{j=1}^n \|\mu(E_j)\| &= \sum_{j=1}^n \left\| \mu \left(\bigcup_{i=1}^{\infty} (A_i \cap E_j) \right) \right\| = \sum_{j=1}^n \left\| \sum_{i=1}^{\infty} \mu(A_i \cap E_j) \right\| \\ &\leq \sum_{j=1}^n \sum_{i=1}^{\infty} \|\mu(A_i \cap E_j)\| = \sum_{i=1}^{\infty} \sum_{j=1}^n \|\mu(A_i \cap E_j)\| \leq \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(A_i) \end{aligned}$$

By the arbitrariness of n and $(E_j)_{j=1}^n$, we have $\mathcal{P} \circ \mu(A) \leq \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(A_i)$.

Hence, $\mathcal{P} \circ \mu(A) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(A_i)$. This implies that $\mathcal{P} \circ \mu$ is a measure on the algebra $\mathcal{A}$. This completes the proof of the proposition. $\square$

Clearly, $\mu : \mathcal{A} \rightarrow \mathcal{Y}$ is a finite $\mathcal{Y}$ -valued pre-measure on the algebra $\mathcal{A}$ if, and only if, μ is a finite $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}$. In this case, the definitions of total variations of μ coincide considering μ as a $\mathcal{Y}$ -valued pre-measure on the algebra $\mathcal{A}$ or as a $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}$.

12.2 Change of Variable

Definition 12.7 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a ($\mathcal{Y}$ -valued) measure space, $\bar{\mathcal{X}} := (\bar{X}, \bar{\mathcal{B}}, \nu)$ be a ($\mathcal{Y}$ -valued) measure space, ($\mathcal{Y}$ be a normed linear space,) and $g : X \rightarrow \bar{X}$. g is said to be an isomeasure between $\mathcal{X}$ and $\bar{\mathcal{X}}$ if g is bijective; $\forall E \in \mathcal{B}$, $g(E) \in \bar{\mathcal{B}}$; $\forall \bar{E} \in \bar{\mathcal{B}}$, $g_{\text{inv}}(\bar{E}) \in \mathcal{B}$; $(\forall \bar{E} \in \bar{\mathcal{B}}, \mathcal{P} \circ \nu(\bar{E}) = \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})))$; $\forall E \in \text{dom}(\mu)$, we have $g(E) \in \text{dom}(\nu)$ and $\nu(g(E)) = \mu(E)$; and $\forall \bar{E} \in \text{dom}(\nu)$, we have $g_{\text{inv}}(\bar{E}) \in \text{dom}(\mu)$ and $\nu(\bar{E}) = \mu(g_{\text{inv}}(\bar{E}))$. We will say $\mathcal{X}$ and $\bar{\mathcal{X}}$ are isomeasuric if there is an isomeasure between $\mathcal{X}$ and $\bar{\mathcal{X}}$.

Two ($\mathcal{Y}$ -valued) measure spaces are isomeasuric, then they are equivalent up to relabeling of elements. Thus, isomeasure preserves the completeness, finiteness, and σ -finiteness of the ($\mathcal{Y}$ -valued) measure spaces.

Definition 12.8 Let $\mathbb{X} := (\mathcal{X}, \mathcal{B}, \mu)$ be a ($\mathcal{Y}$ -valued) topological measure space, $\bar{\mathbb{X}} := (\bar{\mathcal{X}}, \bar{\mathcal{B}}, \nu)$ be a ($\mathcal{Y}$ -valued) topological measure space, ($\mathcal{Y}$ be a normed linear space,) and $g : \mathbb{X} \rightarrow \bar{\mathbb{X}}$. g is said to be an homeomorphical isomeasure between $\mathbb{X}$ and $\bar{\mathbb{X}}$ if $g : \mathcal{X} \rightarrow \bar{\mathcal{X}}$ is a homeomorphism and g is an isomeasure between $\mathbb{X}$ and $\bar{\mathbb{X}}$. We will say $\mathbb{X}$ and $\bar{\mathbb{X}}$ are homeomorphically isomeasuric if there is an homeomorphic isomeasure between $\mathbb{X}$ and $\bar{\mathbb{X}}$.

Two ($\mathcal{Y}$ -valued) topological measure spaces are homeomorphically isomeasuric, then they are equivalent up to relabeling of elements.

Proposition 12.9 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $(\bar{X}, \bar{\mathcal{B}})$ be a measurable space, and $g : X \rightarrow \bar{X}$ be bijective and such that $\forall \bar{E} \in \bar{\mathcal{B}}$, $g_{\text{inv}}(\bar{E}) \in \mathcal{B}$ and, $\forall E \in \mathcal{B}$, $g(E) \in \bar{\mathcal{B}}$. Then, we may define the induced measure $g_*\mu : \bar{\mathcal{B}} \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $g_*\mu(\bar{E}) = \mu(g_{\text{inv}}(\bar{E}))$, $\forall \bar{E} \in \bar{\mathcal{B}}$. Then, $\bar{\mathcal{X}} := (\bar{X}, \bar{\mathcal{B}}, g_*\mu)$ is a measure space and g is an isomeasure between $\mathcal{X}$ and $\bar{\mathcal{X}}$.

Proof This is straightforward, and therefore omitted. □

Proposition 12.10 Let $\mathcal{Y}$ be a normed linear space, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, $(\bar{X}, \bar{\mathcal{B}})$ be a measurable space, and $g : X \rightarrow \bar{X}$ be bijective and such that $\forall \bar{E} \in \bar{\mathcal{B}}$, $g_{\text{inv}}(\bar{E}) \in \mathcal{B}$ and, $\forall E \in \mathcal{B}$, $g(E) \in \bar{\mathcal{B}}$. Then, we may define the induced measure $g_*\mu$ from $\bar{\mathcal{B}}$ to $\mathcal{Y}$ by, $\forall \bar{E} \in \bar{\mathcal{B}}$ with $g_{\text{inv}}(\bar{E}) \in \text{dom}(\mu)$, $g_*\mu(\bar{E}) = \mu(g_{\text{inv}}(\bar{E})) \in \mathcal{Y}$; and $\forall \bar{E} \in \bar{\mathcal{B}}$ with $g_{\text{inv}}(\bar{E}) \in \mathcal{B} \setminus \text{dom}(\mu)$, $g_*\mu(\bar{E})$ is undefined. Then, $\bar{\mathcal{X}} := (\bar{X}, \bar{\mathcal{B}}, g_*\mu)$ is a $\mathcal{Y}$ -valued measure space, $\mathcal{P} \circ (g_*\mu) = g_*(\mathcal{P} \circ \mu)$, and g is an isomeasure between $\mathcal{X}$ and $\bar{\mathcal{X}}$.

Proof Since $\mathcal{X}$ is a $\mathcal{Y}$ -valued measure space, then $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ is a measure space. By Proposition 12.9, $(\bar{X}, \bar{\mathcal{B}}, g_*(\mathcal{P} \circ \mu))$ is a measure space and g is an isomeasure between $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ and $(\bar{X}, \bar{\mathcal{B}}, g_*(\mathcal{P} \circ \mu))$.

(i) $g_*\mu(\emptyset) = \mu(g_{\text{inv}}(\emptyset)) = \mu(\emptyset) = \emptyset_{\mathcal{Y}} \in \mathcal{Y}$.
(ii) $\forall \bar{E} \in \bar{\mathcal{B}}$ with $g_*(\mathcal{P} \circ \mu)(\bar{E}) = +\infty$, we have $\mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) = +\infty$ and $g_{\text{inv}}(\bar{E}) \in \mathcal{B} \setminus \text{dom}(\mu)$. This implies that $g_*\mu(\bar{E})$ is undefined.
(iii) $\forall \bar{E} \in \bar{\mathcal{B}}$ with $g_*(\mathcal{P} \circ \mu)(\bar{E}) < \infty$, we have $\mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) < \infty$ and $g_{\text{inv}}(\bar{E}) \in \text{dom}(\mu)$. This implies that $g_*\mu(\bar{E}) = \mu(g_{\text{inv}}(\bar{E})) \in \mathcal{Y}$. $\forall$ pairwise disjoint $(\bar{E}_i)_{i=1}^{\infty} \subseteq \bar{\mathcal{B}}$ with $\bar{E} = \bigcup_{i=1}^{\infty} \bar{E}_i$, $\forall i \in \mathbb{N}$, $\bar{E}_i \subseteq \bar{E}$, $g_{\text{inv}}(\bar{E}_i) \subseteq g_{\text{inv}}(\bar{E})$, $\mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E}_i)) \leq \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) < \infty$, and $g_{\text{inv}}(\bar{E}_i) \in \text{dom}(\mu)$. Then, $\sum_{i=1}^{\infty} \|g_*\mu(\bar{E}_i)\| = \sum_{i=1}^{\infty} \|\mu(g_{\text{inv}}(\bar{E}_i))\| \leq \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E}_i)) = \mathcal{P} \circ \mu(\bigcup_{i=1}^{\infty} g_{\text{inv}}(\bar{E}_i)) = \mathcal{P} \circ \mu(g_{\text{inv}}(\bigcup_{i=1}^{\infty} \bar{E}_i)) = \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) < \infty$, where the first inequality follows from Definition 11.108; the second equality follows from the fact that $\mathcal{P} \circ \mu$ is a measure; and the third equality follows from Proposition 2.5. Furthermore, $\sum_{i=1}^{\infty} g_*\mu(\bar{E}_i) = \sum_{i=1}^{\infty} \mu(g_{\text{inv}}(\bar{E}_i)) = \mu(\bigcup_{i=1}^{\infty} g_{\text{inv}}(\bar{E}_i)) = \mu(g_{\text{inv}}(\bigcup_{i=1}^{\infty} \bar{E}_i)) = \mu(g_{\text{inv}}(\bar{E})) = g_*\mu(\bar{E}) \in \mathcal{Y}$, where the second equality follows from Definition 11.108; and the third equality follows from Proposition 2.5.

(iv) $\forall \bar{E} \in \bar{\mathcal{B}}$ with $g_*(\mathcal{P} \circ \mu)(\bar{E}) < \infty$, by the previous paragraph, we have $\sup_{n \in \mathbb{Z}_+} (\bar{E}_i)_{i=1}^n \subseteq \bar{\mathcal{B}}$, $\bar{E} = \bigcup_{i=1}^{\infty} \bar{E}_i$, $\bar{E}_i \cap \bar{E}_j = \emptyset$, $\forall 1 \leq i < j \leq n$, $\sum_{i=1}^n \|g_*\mu(\bar{E}_i)\| =: s_{\bar{E}} \leq \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) = g_*(\mathcal{P} \circ \mu)(\bar{E})$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n \in \mathbb{Z}_+$, $\exists$ pairwise disjoint $(\bar{E}_i)_{i=1}^n \subseteq \bar{\mathcal{B}}$ with $g_{\text{inv}}(\bar{E}) = \bigcup_{i=1}^n \bar{E}_i$, such that $\mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) - \epsilon < \sum_{i=1}^n \mu(\bar{E}_i) \leq \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) = g_*(\mathcal{P} \circ \mu)(\bar{E})$. Then, $(\bar{E}_i)_{i=1}^n := (g(\bar{E}_i))_{i=1}^n \subseteq \bar{\mathcal{B}}$ is pairwise disjoint, $\bigcup_{i=1}^n \bar{E}_i = g(\bigcup_{i=1}^n \bar{E}_i) = g(g_{\text{inv}}(\bar{E})) = \bar{E}$, and $\sum_{i=1}^n \|g_*\mu(\bar{E}_i)\| = \sum_{i=1}^n \|\mu(g_{\text{inv}}(\bar{E}_i))\| = \sum_{i=1}^n \mu(\bar{E}_i) > g_*(\mathcal{P} \circ \mu)(\bar{E}) - \epsilon$. Hence, $s_{\bar{E}} > g_*(\mathcal{P} \circ \mu)(\bar{E}) - \epsilon$. By the arbitrariness of ϵ , we have $s_{\bar{E}} \geq g_*(\mathcal{P} \circ \mu)(\bar{E})$ and $s_{\bar{E}} = g_*(\mathcal{P} \circ \mu)(\bar{E})$.

Therefore, by Definition 11.108, $(\bar{X}, \bar{\mathcal{B}}, g_*\mu)$ is a $\mathcal{Y}$ -valued measure space with $\mathcal{P} \circ (g_*\mu) = g_*(\mathcal{P} \circ \mu)$. It is straightforward to show that g is an isomeasure between $\mathcal{X}$ and $\bar{\mathcal{X}}$. This completes the proof of the proposition. $\square$

Proposition 12.11 *Let $\mathcal{X} := (X, \mathcal{O})$ and $\bar{\mathcal{X}} := (\bar{X}, \bar{\mathcal{O}})$ be topological spaces, $g : \mathcal{X} \rightarrow \bar{\mathcal{X}}$ be a homeomorphism, $(\mathcal{Y}, \|\cdot\|)$ be a normed linear space, and $\mathbb{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ be a $(\mathcal{Y}$ -valued) topological measure space. Then, $\bar{\mathbb{X}} := (\bar{\mathcal{X}}, \mathcal{B}_B(\bar{\mathcal{X}}), g_*\mu)$ is a $(\mathcal{Y}$ -valued) topological measure space and g is a homeomorphical isomeasure between $\mathbb{X}$ and $\bar{\mathbb{X}}$.*

Proof Clearly, g is bijective and g and g_{inv} are continuous. By Proposition 11.34, $\forall E \in \mathcal{B}_B(\mathcal{X})$, we have $g(E) \in \mathcal{B}_B(\bar{\mathcal{X}})$; and $\forall \bar{E} \in \mathcal{B}_B(\bar{\mathcal{X}})$, we have $g_{\text{inv}}(\bar{E}) \in \mathcal{B}_B(\mathcal{X})$. Then, by Propositions 12.9 and 12.10, the $(\mathcal{Y}$ -valued) induced measure $g_*\mu$ is well defined, $(\bar{X}, \mathcal{B}_B(\bar{\mathcal{X}}), g_*\mu)$ is a $(\mathcal{Y}$ -valued) measure space, and g is an isomeasure between $\mathbb{X}$ and $\bar{\mathbb{X}}$.

We will show that $\bar{\mathbb{X}}$ is a $(\mathcal{Y}$ -valued) topological measure space. $\forall \bar{E} \in \mathcal{B}_B(\bar{\mathcal{X}})$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, we have $g_{\text{inv}}(\bar{E}) \in \mathcal{B}_B(\mathcal{X})$. Since $\mathbb{X}$ is a $(\mathcal{Y}$ -valued) topological measure space, then $\exists U \in \mathcal{O}$ with $g_{\text{inv}}(\bar{E}) \subseteq U$ such that $\mu(U \setminus g_{\text{inv}}(\bar{E})) < \epsilon$ ($\mathcal{P} \circ \mu(U \setminus g_{\text{inv}}(\bar{E})) < \epsilon$). Let $\bar{U} := g(U) \in \bar{\mathcal{O}}$. Then, $\bar{E} \subseteq g(U) = \bar{U}$ and $g_*\mu(\bar{U} \setminus \bar{E}) = \mu(g_{\text{inv}}(\bar{U} \setminus \bar{E})) = \mu(g_{\text{inv}}(\bar{U}) \setminus g_{\text{inv}}(\bar{E})) = \mu(U \setminus g_{\text{inv}}(\bar{E})) < \epsilon$ (in the case when μ is a $\mathcal{Y}$ -valued measure, we have $\mathcal{P} \circ (g_*\mu)(\bar{U} \setminus \bar{E}) = g_*(\mathcal{P} \circ \mu)(\bar{U} \setminus \bar{E}) = \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{U} \setminus \bar{E})) = \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{U}) \setminus g_{\text{inv}}(\bar{E})) = \mathcal{P} \circ \mu(U \setminus g_{\text{inv}}(\bar{E})) < \epsilon$, where the first equality follows from Proposition 12.10). Hence, $\bar{\mathbb{X}}$ is a $(\mathcal{Y}$ -valued) topological measure space. Then, g is a homeomorphical isomeasure between $\mathbb{X}$ and $\bar{\mathbb{X}}$. $\square$

Proposition 12.12 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\bar{\mathcal{X}} := (\bar{X}, \bar{\mathcal{B}}, \nu)$ be a measure space, $\mathcal{Y}$ be a normed linear space, $f : \bar{\mathcal{X}} \rightarrow \mathcal{Y}$ be $\bar{\mathcal{B}}$ -measurable, and $g : \mathcal{X} \rightarrow \bar{\mathcal{X}}$ be an isomeasure between $\mathcal{X}$ and $\bar{\mathcal{X}}$. Then, $\int_{\bar{\mathcal{X}}} (f \circ g) d\mu = \int_{\bar{\mathcal{X}}} f d\nu$, whenever one of the integrals exists.*

Proof Clearly, $f \circ g$ is $\mathcal{B}$ -measurable. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu(X) < +\infty$; Case 2: $\mu(X) = +\infty$.

Case 1: $\mu(X) < +\infty$. $\nu(\bar{X}) = \mu(g_{\text{inv}}(\bar{X})) = \mu(X) < +\infty$. Let $\mathcal{J}(\mathcal{Y})$ be the integration system on $\mathcal{Y}$. $\forall R = \{(y_\alpha, U_\alpha) \mid \alpha \in \Lambda\} \in \mathcal{J}(\mathcal{Y})$, let $\bar{F}_R := \sum_{\alpha \in \Lambda} y_\alpha \nu(f_{\text{inv}}(U_\alpha)) \in \mathcal{Y}$ and $F_R := \sum_{\alpha \in \Lambda} y_\alpha \mu((f \circ g)_{\text{inv}}(U_\alpha)) \in \mathcal{Y}$. Note that $\bar{F}_R = \sum_{\alpha \in \Lambda} y_\alpha \mu(g_{\text{inv}}(f_{\text{inv}}(U_\alpha))) = F_R$. Then, by Definition 11.70, we have $\int_{\bar{\mathcal{X}}} f d\nu = \lim_{R \in \mathcal{J}(\mathcal{Y})} \bar{F}_R = \lim_{R \in \mathcal{J}(\mathcal{Y})} F_R = \int_X (f \circ g) d\mu$, whenever one of the integrals exists.

Case 2: $\mu(X) = +\infty$. $\nu(\bar{X}) = \mu(g_{\text{inv}}(\bar{X})) = \mu(X) = +\infty$. Let $(\bar{F}_A)_{A \in \mathcal{M}(\bar{\mathcal{X}})}$ be the net for $\int_{\bar{\mathcal{X}}} f d\nu$ and $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X (f \circ g) d\mu$ as defined in Definition 11.71. Without loss of generality,

assume that $\int_{\bar{X}} f \, d\nu$ exists. Then,

$$\begin{aligned} \int_{\bar{X}} f \, d\nu &= \lim_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})} \bar{F}_{\bar{A}} = \lim_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})} \int_{\bar{A}} f|_{\bar{A}} \, d\nu_{\bar{A}} \\ &= \lim_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})} \int_{g_{\text{inv}}(\bar{A})} (f \circ g)|_{g_{\text{inv}}(\bar{A})} \, d\mu_{g_{\text{inv}}(\bar{A})} \end{aligned}$$

where the first two equalities follow from Definition 11.71; and the third equality follows from Case 1. Fix any open set U ($U \subseteq \mathbb{R}_e$ if $\mathcal{Y} = \mathbb{R}$ or $U \subseteq \mathcal{Y}$ if $\mathcal{Y} \neq \mathbb{R}$) with $\int_{\bar{X}} f \, d\nu \in U$, $\exists \bar{A}_0 \in \mathcal{M}(\bar{\mathcal{X}})$, $\forall \bar{A} \in \mathcal{M}(\bar{\mathcal{X}})$ with $\bar{A}_0 \subseteq \bar{A}$, $\bar{F}_{\bar{A}} = \int_{g_{\text{inv}}(\bar{A})} (f \circ g)|_{g_{\text{inv}}(\bar{A})} \, d\mu_{g_{\text{inv}}(\bar{A})} \in U$. $\forall A \in \mathcal{M}(\mathcal{X})$ with $g_{\text{inv}}(\bar{A}_0) \subseteq A$, we have $A \in \mathcal{B}$ and $\mu(A) < +\infty$. Then, $g(A) \in \bar{\mathcal{B}}$ and $\mu(A) = \nu(g(A)) < +\infty$. Hence, $g(A) \in \mathcal{M}(\bar{\mathcal{X}})$ and $g(A) \supseteq g(g_{\text{inv}}(\bar{A}_0)) = \bar{A}_0$. Then, $F_A = \int_A (f \circ g)|_A \, d\mu_A = \int_{g_{\text{inv}}(g(A))} (f \circ g)|_{g_{\text{inv}}(g(A))} \, d\mu_{g_{\text{inv}}(g(A))} \in U$. Hence, $\int_X (f \circ g) \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_A = \int_{\bar{X}} f \, d\nu$.

This completes the proof of the proposition. $\square$

Proposition 12.13 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\bar{\mathcal{X}} := (\bar{X}, \bar{\mathcal{B}}, \nu)$ be a measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $f : \bar{X} \rightarrow \mathcal{W}$ be absolutely integrable over $\bar{\mathcal{X}}$, and $g : X \rightarrow \bar{X}$ be an isomeasure between $\mathcal{X}$ and $\bar{\mathcal{X}}$. Then, $f \circ g$ is absolutely integrable over $\mathcal{X}$ and $\int_X (f \circ g) \, d\mu = \int_{\bar{X}} f \, d\nu \in \mathcal{Y}$.*

Proof Clearly, $f \circ g$ is $\mathcal{B}$ -measurable. By the assumption, $0 \leq \int_{\bar{X}} (\mathcal{P} \circ f) \, d\nu < +\infty$. By Proposition 12.12, $\int_X (\mathcal{P} \circ f \circ g) \, d\mu = \int_{\bar{X}} (\mathcal{P} \circ f) \, d\nu \in [0, \infty) \subset \mathbb{R}$. Hence, $f \circ g$ is absolutely integrable over $\mathcal{X}$. By Proposition 11.92, f is integrable over $\bar{\mathcal{X}}$. By Proposition 12.12, $\int_X (f \circ g) \, d\mu = \int_{\bar{X}} f \, d\nu \in \mathcal{Y}$. This completes the proof of the proposition. $\square$

Proposition 12.14 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\bar{\mathcal{X}} := (\bar{X}, \bar{\mathcal{B}}, \nu)$ be a $\mathcal{Y}$ -valued measure space, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $\mathcal{W} := \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $f : \bar{X} \rightarrow \mathcal{W}$ be $\bar{\mathcal{B}}$ -measurable, and $g : X \rightarrow \bar{X}$ be an isomeasure between $\mathcal{X}$ and $\bar{\mathcal{X}}$. Then, $\int_{\bar{X}} f \, d\nu = \int_X (f \circ g) \, d\mu$, whenever one of the integrals exists.*

Proof Clearly, $f \circ g$ is $\mathcal{B}$ -measurable. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{P} \circ \mu(X) < +\infty$; Case 2: $\mathcal{P} \circ \mu(X) = +\infty$.

Case 1: $\mathcal{P} \circ \mu(X) < +\infty$. $\mathcal{P} \circ \nu(\bar{X}) = \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{X})) = \mathcal{P} \circ \mu(X) < +\infty$. Let $\mathcal{I}(\mathcal{W})$ be the integration system on $\mathcal{W}$. $\forall R = \{(w_\alpha, U_\alpha) \mid \alpha \in \Lambda\} \in \mathcal{I}(\mathcal{W})$, let $\bar{F}_R := \sum_{\alpha \in \Lambda} w_\alpha \nu(f_{\text{inv}}(U_\alpha)) \in \mathcal{Z}$ and $F_R := \sum_{\alpha \in \Lambda} w_\alpha \mu((f \circ g)_{\text{inv}}(U_\alpha)) \in \mathcal{Z}$. Then, $\bar{F}_R = \sum_{\alpha \in \Lambda} w_\alpha \mu(g_{\text{inv}}(f_{\text{inv}}(U_\alpha))) = F_R$. Then, by Definition 11.119, we have $\int_{\bar{X}} f \, d\nu = \lim_{R \in \mathcal{I}(\mathcal{W})} \bar{F}_R = \lim_{R \in \mathcal{I}(\mathcal{W})} F_R = \int_X (f \circ g) \, d\mu$, whenever one of the integrals exists.

Case 2: $\mathcal{P} \circ \mu(X) = +\infty$. $\mathcal{P} \circ \nu(\bar{X}) = \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{X})) = \mathcal{P} \circ \mu(X) = +\infty$. Let $(\bar{F}_{\bar{A}})_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})}$ be the net for $\int_{\bar{X}} f \, d\nu$ and $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ be the net for

$\int_X (f \circ g) d\mu$ as defined in Definition 11.120. Without loss of generality, assume that $\int_{\bar{X}} f d\nu$ exists. Then,

$$\begin{aligned} \int_{\bar{X}} f d\nu &= \lim_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})} \bar{F}_{\bar{A}} = \lim_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})} \int_{\bar{A}} f|_{\bar{A}} d\nu_{\bar{A}} \\ &= \lim_{\bar{A} \in \mathcal{M}(\bar{\mathcal{X}})} \int_{g_{\text{inv}}(\bar{A})} (f \circ g)|_{g_{\text{inv}}(\bar{A})} d\mu_{g_{\text{inv}}(\bar{A})} \end{aligned}$$

where the first two equalities follow from Definition 11.120; and the third equality follows from Case 1. Fix any open set U ($U \subseteq \mathbb{R}_e$ if $\mathcal{Z} = \mathbb{R}$ or $U \subseteq \mathcal{Z}$ if $\mathcal{Z} \neq \mathbb{R}$) with $\int_{\bar{X}} f d\nu \in U$, $\exists \bar{A}_0 \in \mathcal{M}(\bar{\mathcal{X}})$, $\forall \bar{A} \in \mathcal{M}(\bar{\mathcal{X}})$ with $\bar{A}_0 \subseteq \bar{A}$, $\bar{F}_{\bar{A}} = \int_{g_{\text{inv}}(\bar{A})} (f \circ g)|_{g_{\text{inv}}(\bar{A})} d\mu_{g_{\text{inv}}(\bar{A})} \in U$. $\forall \bar{A} \in \mathcal{M}(\bar{\mathcal{X}})$ with $g_{\text{inv}}(\bar{A}_0) \subseteq A$, we have $A \in \mathcal{B}$ and $\mathcal{P} \circ \mu(A) < +\infty$. Then, $g(A) \in \bar{\mathcal{B}}$ and $\mathcal{P} \circ \mu(A) = \mathcal{P} \circ \nu(g(A)) < +\infty$. Hence, $g(A) \in \mathcal{M}(\bar{\mathcal{X}})$ and $g(A) \supseteq g(g_{\text{inv}}(\bar{A}_0)) = \bar{A}_0$. Then, $F_A = \int_A (f \circ g)|_A d\mu_A = \int_{g_{\text{inv}}(g(A))} (f \circ g)|_{g_{\text{inv}}(g(A))} d\mu_{g_{\text{inv}}(g(A))} \in U$. Hence, $\int_X (f \circ g) d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_A = \int_{\bar{X}} f d\nu$.

This completes the proof of the proposition. $\square$

Proposition 12.15 *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\bar{\mathcal{X}} := (\bar{X}, \bar{\mathcal{B}}, \nu)$ be a $\mathcal{Y}$ -valued measure space, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, $f : \bar{X} \rightarrow \mathcal{W}$ be absolutely integrable over $\bar{\mathcal{X}}$, and $g : X \rightarrow \bar{X}$ be an isomeasure between $\mathcal{X}$ and $\bar{\mathcal{X}}$. Then, $f \circ g$ is absolutely integrable over $\mathcal{X}$, and $\int_{\bar{X}} f d\nu = \int_X (f \circ g) d\mu \in \mathcal{Z}$.*

Proof Clearly, $f \circ g$ is $\mathcal{B}$ -measurable. By Definition 12.7, g is an isomeasure between $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ and $(\bar{X}, \bar{\mathcal{B}}, \mathcal{P} \circ \nu)$. By Proposition 12.12, $\int_X (\mathcal{P} \circ f \circ g) d\mathcal{P} \circ \mu = \int_{\bar{X}} (\mathcal{P} \circ f) d\mathcal{P} \circ \nu \in \mathbb{R}$, where the last step follows from the assumption that f is absolutely integrable over $\bar{\mathcal{X}}$. Hence, $f \circ g$ is absolutely integrable over $\mathcal{X}$. By Proposition 11.132, f is integrable over $\bar{\mathcal{X}}$. By Proposition 12.14, $\int_X (f \circ g) d\mu = \int_{\bar{X}} f d\nu \in \mathcal{Z}$. $\square$

Theorem 12.16 (Change of Variable) *Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite $\mathbb{K}$ -valued measure space, $\bar{\mathcal{X}} := (\bar{X}, \bar{\mathcal{B}}, \nu)$ be a σ -finite $\mathcal{Y}$ -valued measure space, where $\mathcal{Y}$ is a separable Banach space over $\mathbb{K}$, $g : \bar{X} \rightarrow X$ be an isomeasure between $\bar{\mathcal{X}}$ and $\hat{\mathcal{X}} := (X, \mathcal{B}, g_*\nu)$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $B(\mathcal{Y}, \mathcal{Z})$, $f : \bar{X} \rightarrow \mathcal{W}$ be $\bar{\mathcal{B}}$ -measurable, and $\frac{d(g_*\nu)}{d\mu} =: h : X \rightarrow \mathcal{Y}$. Then, f is absolutely integrable over $\bar{\mathcal{X}}$ if, and only if, $(\mathcal{P} \circ f \circ g_{\text{inv}})(\mathcal{P} \circ h)$ is integrable over $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. In which case, $(f \circ g_{\text{inv}})h$ is absolutely integrable over $\mathcal{X}$ and $\int_{\bar{X}} f d\nu = \int_X (f \circ g_{\text{inv}})h d\mu \in \mathcal{Z}$.*

Proof By Proposition 12.15, f is absolutely integrable over $\bar{\mathcal{X}}$ if, and only if, $f \circ g_{\text{inv}}$ is absolutely integrable over $\hat{\mathcal{X}}$. By (ii) of Proposition 11.168, $f \circ g_{\text{inv}}$ is absolutely integrable over $\hat{\mathcal{X}}$ if, and only if, $(\mathcal{P} \circ f \circ g_{\text{inv}})(\mathcal{P} \circ h)$ is integrable over $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. Hence, f is absolutely integrable over $\bar{\mathcal{X}}$ if,

and only if, $(\mathcal{P} \circ f \circ g_{\text{inv}})(\mathcal{P} \circ h)$ is integrable over $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. In this case, $(f \circ g_{\text{inv}})h$ is absolutely integrable over $\mathcal{X}$ and, by Proposition 12.15 and (ii) of Proposition 11.168, $\int_{\bar{X}} f \, d\nu = \int_X (f \circ g_{\text{inv}}) \, d(g_*\nu) = \int_X (f \circ g_{\text{inv}})h \, d\mu \in \mathbb{Z}$. This completes the proof of the theorem. $\square$

12.3 Product Measure

Definition 12.17 Let X be a set. $\mathcal{A} \subseteq {}^X 2$ is said to be a π -system on X if: (i) $\emptyset, X \in \mathcal{A}$; and (ii) $\forall A, B \in \mathcal{A}$, we have $A \cap B \in \mathcal{A}$. $\mathcal{D} \subseteq {}^X 2$ is said to be a monotone class on X if

- (i) $\emptyset, X \in \mathcal{D}$;
- (ii) $\forall A, B \in \mathcal{D}$ with $A \subseteq B$, we have $B \setminus A \in \mathcal{D}$;
- (iii) $\forall (A_i)_{i=1}^\infty \subseteq \mathcal{D}$ with $A_i \subseteq A_{i+1}$, $\forall i \in \mathbb{N}$, we have $\bigcup_{i=1}^\infty A_i \in \mathcal{D}$.

Let X be a set and $\mathcal{A} \subseteq {}^X 2$. Then, there exists the smallest monotone class on X that contains $\mathcal{A}$, which is the intersection of all monotone classes on X that contains $\mathcal{A}$.

Proposition 12.18 Let X be a set. $\mathcal{B} \subseteq {}^X 2$ is a σ -algebra on X if, and only if, $\mathcal{B}$ is a π -system on X and is a monotone class on X .

Proof “Necessity” Let $\mathcal{B}$ be a σ -algebra on X . Then, $\emptyset, X \in \mathcal{B}$. $\forall A, B \in \mathcal{B}$, we have $A \cap B \in \mathcal{B}$. Hence, $\mathcal{B}$ is a π -system on X . $\forall A, B \in \mathcal{B}$ with $A \subseteq B$, $B \setminus A \in \mathcal{B}$. $\forall (A_i)_{i=1}^\infty \subseteq \mathcal{B}$ with $A_i \subseteq A_{i+1}$, $\forall i \in \mathbb{N}$, we have $\bigcup_{i=1}^\infty A_i \in \mathcal{B}$. Hence, $\mathcal{B}$ is a monotone class on X .

“Sufficiency” Let $\mathcal{B}$ be a monotone class on X and a π -system on X . Then, $\emptyset, X \in \mathcal{B}$. $\forall A \subseteq \mathcal{B}$, we have $A \subseteq X$ and $X \setminus A = \widetilde{A} \in \mathcal{B}$, since $\mathcal{B}$ is a monotone class on X . $\forall \underline{A_1}, \underline{A_2} \in \mathcal{B}$, we have $A_1 \cup A_2 = \widetilde{\widetilde{A_1} \cap \widetilde{A_2}}$. By the above, $\widetilde{A_1}, \widetilde{A_2} \in \mathcal{B}$. Then, $\widetilde{A_1} \cap \widetilde{A_2} \in \mathcal{B}$ since $\mathcal{B}$ is a π -system on X . This further implies that $A_1 \cup A_2 \in \mathcal{B}$. Hence, $\mathcal{B}$ is an algebra on X . $\forall (E_i)_{i=1}^\infty \subseteq \mathcal{B}$, $\forall n \in \mathbb{N}$, let $A_n := \bigcup_{i=1}^n E_i \in \mathcal{B}$, since $\mathcal{B}$ is an algebra. Then, $A_n \subseteq A_{n+1}$, $\forall n \in \mathbb{N}$. By $\mathcal{B}$ being a monotone class on X , we have $\bigcup_{n=1}^\infty A_n = \bigcup_{i=1}^\infty E_i \in \mathcal{B}$. This shows that $\mathcal{B}$ is a σ -algebra.

This completes the proof of the proposition. $\square$

Lemma 12.19 (Monotone Class Lemma) Let X be a set, $\mathcal{A} \subseteq {}^X 2$ be a π -system on X , $\mathcal{B}$ be the σ -algebra on X generated by $\mathcal{A}$, and $\mathcal{D}$ be the smallest monotone class on X that contains $\mathcal{A}$. Then, $\mathcal{B} = \mathcal{D}$.

Proof By Proposition 12.18, $\mathcal{B}$ is a monotone class on X . Hence, $\mathcal{D} \subseteq \mathcal{B}$.

Define $\mathcal{D}_1 := \{E \in \mathcal{D} \mid E \cap A \in \mathcal{D}, \forall A \in \mathcal{A}\}$. Since $\mathcal{A}$ is a π -system on X , then $\mathcal{A} \subseteq \mathcal{D}_1$. Clearly $\mathcal{D}_1 \subseteq \mathcal{D}$. We will show that $\mathcal{D}_1$ is a monotone class on X . Then, $\mathcal{D} \subseteq \mathcal{D}_1$ and $\mathcal{D} = \mathcal{D}_1$.

Clearly, $\emptyset \in \mathcal{A} \subseteq \mathcal{D}_1$ and $X \in \mathcal{A} \subseteq \mathcal{D}_1$. $\forall E_1, E_2 \in \mathcal{D}_1$ with $E_1 \subseteq E_2$, we have $E_1, E_2 \in \mathcal{D}$ and $E_2 \setminus E_1 \in \mathcal{D}$. $\forall A \in \mathcal{A}$, we have $A \cap E_1 \in \mathcal{D}$ and $A \cap E_2 \in \mathcal{D}$. Clearly, $A \cap E_1 \subseteq A \cap E_2$. Since $\mathcal{D}$ is a monotone class on X , we have $\mathcal{D} \ni (A \cap E_2) \setminus (A \cap E_1) = A \cap (E_2 \setminus E_1)$. By the arbitrariness of A , we have $E_2 \setminus E_1 \in \mathcal{D}_1$. $\forall (E_i)_{i=1}^\infty \subseteq \mathcal{D}_1$ with $E_i \subseteq E_{i+1}$, $\forall i \in \mathbb{N}$, we have $\bigcup_{i=1}^\infty E_i \in \mathcal{D}$ since $\mathcal{D}$ is a monotone class on X . $\forall A \in \mathcal{A}$, $\forall i \in \mathbb{N}$, $E_i \in \mathcal{D}_1$ implies that $A \cap E_i \in \mathcal{D}$. Clearly, $A \cap E_i \subseteq A \cap E_{i+1}$, $\forall i \in \mathbb{N}$. Then, $\mathcal{D} \ni \bigcup_{i=1}^\infty (A \cap E_i) = A \cap (\bigcup_{i=1}^\infty E_i)$, since $\mathcal{D}$ is a monotone class on X . By the arbitrariness of A , we have $\bigcup_{i=1}^\infty E_i \in \mathcal{D}_1$. This shows that $\mathcal{D}_1$ is a monotone class on X .

Define $\mathcal{D}_2 := \{E \in \mathcal{D} \mid E \cap B \in \mathcal{D}, \forall B \in \mathcal{D}\}$. Since $\mathcal{D}_1 = \mathcal{D}$, then $\mathcal{A} \subseteq \mathcal{D}_2$. Clearly, $\mathcal{D}_2 \subseteq \mathcal{D}$. We will show that $\mathcal{D}_2$ is a monotone class on X . Then, $\mathcal{D} \subseteq \mathcal{D}_2$ and $\mathcal{D} = \mathcal{D}_2$.

Clearly, $\emptyset \in \mathcal{A} \subseteq \mathcal{D}_2$ and $X \in \mathcal{A} \subseteq \mathcal{D}_2$. $\forall E_1, E_2 \in \mathcal{D}_2$ with $E_1 \subseteq E_2$, we have $E_2 \setminus E_1 \in \mathcal{D}$ since $\mathcal{D}$ is a monotone class on X . $\forall B \in \mathcal{D}$, we have $E_1 \cap B \in \mathcal{D}$ and $E_2 \cap B \in \mathcal{D}$ since $E_1, E_2 \in \mathcal{D}_2$. Clearly, $E_1 \cap B \subseteq E_2 \cap B$. Then, $\mathcal{D} \ni (E_2 \cap B) \setminus (E_1 \cap B) = (E_2 \setminus E_1) \cap B$, since $\mathcal{D}$ is a monotone class on X . By the arbitrariness of B , we have $E_2 \setminus E_1 \in \mathcal{D}_2$. $\forall (E_i)_{i=1}^\infty \subseteq \mathcal{D}_2$ with $E_i \subseteq E_{i+1}$, $\forall i \in \mathbb{N}$, we have $\bigcup_{i=1}^\infty E_i \in \mathcal{D}$ since $\mathcal{D}$ is a monotone class on X . $\forall B \in \mathcal{D}$, $\forall i \in \mathbb{N}$, $E_i \in \mathcal{D}_2$ implies that $B \cap E_i \in \mathcal{D}$. Clearly, $B \cap E_i \subseteq B \cap E_{i+1}$, $\forall i \in \mathbb{N}$. Then, $\mathcal{D} \ni \bigcup_{i=1}^\infty (B \cap E_i) = B \cap (\bigcup_{i=1}^\infty E_i)$, since $\mathcal{D}$ is a monotone class on X . By the arbitrariness of B , we have $\bigcup_{i=1}^\infty E_i \in \mathcal{D}_2$. This shows that $\mathcal{D}_2$ is a monotone class on X .

Therefore, $\mathcal{D}_2 = \mathcal{D}$. This implies, by the definition of $\mathcal{D}_2$, that $\mathcal{D}$ is a π -system on X . By Proposition 12.18, we have $\mathcal{D}$ is σ -algebra. Then, $\mathcal{B} \subseteq \mathcal{D}$. Hence, $\mathcal{B} = \mathcal{D}$. This completes the proof of the lemma. $\square$

Proposition 12.20 *Let X be a set, $\mathcal{C}$ be a semialgebra on X , $\mathcal{Y}$ be a normed linear space, $\mathcal{A}$ be the algebra on X generated by $\mathcal{C}$, $\nu : \mathcal{C} \rightarrow [0, \infty] \subset \mathbb{R}_e$, and μ be a function from $\mathcal{C}$ to $\mathcal{Y}$. Assume that*

- (i) $\mu(\emptyset) = \vartheta_{\mathcal{Y}}$ and $\nu(\emptyset) = 0$;
- (ii) $\forall C \in \mathcal{C}$, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(C_i)_{i=1}^n \subseteq \mathcal{C}$ with $C = \bigcup_{i=1}^n C_i$, we have $\nu(C) = \sum_{i=1}^n \nu(C_i)$;
- (iii) $\forall C \in \mathcal{C}$, $\forall$ pairwise disjoint $(C_i)_{i=1}^\infty \subseteq \mathcal{C}$ with $C = \bigcup_{i=1}^\infty C_i$, we have $\nu(C) \leq \sum_{i=1}^\infty \nu(C_i)$;
- (iv) $\forall C \in \mathcal{C}$ with $\nu(C) = \infty$, $\mu(C)$ is undefined;
- (v) $\forall C \in \mathcal{C}$ with $\nu(C) < \infty$, $\mu(C) \in \mathcal{Y}$ and $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(C_i)_{i=1}^n \subseteq \mathcal{C}$ with $C = \bigcup_{i=1}^n C_i$, we have $\mu(C) = \sum_{i=1}^n \mu(C_i)$.
- (vi) $\forall C \in \mathcal{C}$ with $\nu(C) < \infty$, we have

$$\nu(C) = \sup_{n \in \mathbb{Z}_+, (C_i)_{i=1}^n \subseteq \mathcal{C}, C = \bigcup_{i=1}^n C_i, C_i \cap C_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(C_i)\|$$

Then, ν admits a unique extension to a measure $\bar{\nu}$ on the algebra $\mathcal{A}$; and μ admits a unique extension to a $\mathcal{Y}$ -valued measure $\bar{\mu}$ on the algebra $\mathcal{A}$ with $(\mathcal{P} \circ \bar{\mu})|_{\mathcal{C}} = \nu$. Furthermore, $\mathcal{P} \circ \bar{\mu} = \bar{\nu}$.

Proof By Proposition 11.31, $\mathcal{A} = \{A \subseteq X \mid A \text{ is the finite disjoint union of sets in } \mathcal{C}\}$. Define $\bar{\nu} : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $\bar{\nu}(A) = \sum_{i=1}^n \nu(C_i)$, $\forall A = \bigcup_{i=1}^n C_i \in \mathcal{A}$, where $n \in \mathbb{Z}_+$ and $(C_i)_{i=1}^n \subseteq \mathcal{C}$ is pairwise disjoint. By (ii), (iii), and Proposition 11.32 and its proof, $\bar{\nu}$ is the unique measure on the algebra $\mathcal{A}$ which is an extension of ν . Hence, $\bar{\nu}$ is the measure on the algebra $\mathcal{A}$ that we seek. Define $\bar{\mu}$ to be a function from $\mathcal{A}$ to $\mathcal{Y}$ by: $\bar{\mu}(A)$ is undefined if $\bar{\nu}(A) = \infty$; and $\bar{\mu}(A) = \sum_{i=1}^n \mu(C_i) \in \mathcal{Y}$ if $\bar{\nu}(A) < \infty$, $A = \bigcup_{i=1}^n C_i$, $n \in \mathbb{Z}_+$, and $(C_i)_{i=1}^n \subseteq \mathcal{C}$ is pairwise disjoint. By (i) and (v), $\bar{\mu}$ is well-defined. Clearly, $\bar{\mu}|_{\mathcal{C}} = \mu$ and hence $\bar{\mu}$ is an extension of μ .

We will show that $\bar{\mu}$ is a $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}$ with $\mathcal{P} \circ \bar{\mu} = \bar{\nu}$. Then, $\bar{\mu}$ is the $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}$ that we seek. By (i), $\bar{\mu}(\emptyset) = \vartheta_{\mathcal{Y}}$. $\forall A \in \mathcal{A}$ with $\bar{\nu}(A) = \infty$, $\bar{\mu}(A)$ is undefined. $\forall A \in \mathcal{A}$ with $\bar{\nu}(A) < \infty$, $\forall$ pairwise disjoint $(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ with $A = \bigcup_{i=1}^{\infty} A_i$, we have $\bar{\mu}(A) \in \mathcal{Y}$, $\bar{\nu}(A_i) < \infty$, and $\bar{\mu}(A_i) \in \mathcal{Y}$, $\forall i \in \mathbb{N}$. Since $A \in \mathcal{A}$, then $A = \bigcup_{i=1}^{n_0} C_i$, where $n_0 \in \mathbb{Z}_+$ and $(C_i)_{i=1}^{n_0} \subseteq \mathcal{C}$ is pairwise disjoint. Fix any $i \in \mathbb{N}$, $A_i \in \mathcal{A}$ implies that $A_i = \bigcup_{j=1}^{n_i} C_{i,j}$, where $n_i \in \mathbb{Z}_+$ and $(C_{i,j})_{j=1}^{n_i} \subseteq \mathcal{C}$ is pairwise disjoint. This leads to $\sum_{i=1}^{\infty} \|\bar{\mu}(A_i)\| = \sum_{i=1}^{\infty} \left\| \sum_{j=1}^{n_i} \mu(C_{i,j}) \right\| \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \|\mu(C_{i,j})\| \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \nu(C_{i,j}) = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \bar{\nu}(C_{i,j}) = \bar{\nu}(\bigcup_{i=1}^{\infty} \bigcup_{j=1}^{n_i} C_{i,j}) = \bar{\nu}(A) < \infty$, where the first equality follows from the definition of $\bar{\mu}$; the second inequality follows from (vi); the second equality follows from $\bar{\nu}|_{\mathcal{C}} = \nu$; and the third equality follows from the fact that $\bar{\nu}$ is a measure on the algebra $\mathcal{A}$. This implies that $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $\sum_{i=N+1}^{\infty} \sum_{j=1}^{n_i} \bar{\nu}(C_{i,j}) < \epsilon$. $\forall n \in \mathbb{N}$ with $N \leq n$, let $\bar{A} := A \setminus (\bigcup_{i=1}^n \bigcup_{j=1}^{n_i} C_{i,j}) \in \mathcal{A}$. Clearly, $\bar{A} = \bigcup_{i=n+1}^{\infty} \bigcup_{j=1}^{n_i} C_{i,j}$ and $\bar{\nu}(\bar{A}) = \sum_{i=n+1}^{\infty} \sum_{j=1}^{n_i} \bar{\nu}(C_{i,j}) < \epsilon$. By $\bar{A} \in \mathcal{A}$, we have $\bar{A} = \bigcup_{i=1}^{\bar{n}} \bar{C}_i$, where $\bar{n} \in \mathbb{Z}_+$ and $(\bar{C}_i)_{i=1}^{\bar{n}} \subseteq \mathcal{C}$ is pairwise disjoint. Then, $\|\bar{\mu}(A) - \sum_{i=1}^n \bar{\mu}(A_i)\| = \left\| \sum_{k=1}^{n_0} \mu(C_k) - \sum_{i=1}^n \sum_{j=1}^{n_i} \mu(C_{i,j}) \right\| = \left\| \sum_{k=1}^{n_0} \mu(C_k \cap ((\bigcup_{i=1}^n \bigcup_{j=1}^{n_i} C_{i,j}) \cup \bar{A})) - \sum_{i=1}^n \sum_{j=1}^{n_i} \mu(C_{i,j} \cap (\bigcup_{k=1}^{n_0} C_k)) \right\| = \left\| \sum_{k=1}^{n_0} \mu((\bigcup_{i=1}^n \bigcup_{j=1}^{n_i} (C_k \cap C_{i,j})) \cup (\bigcup_{i=1}^{\bar{n}} (C_k \cap \bar{C}_i))) - \sum_{i=1}^n \sum_{j=1}^{n_i} \mu(\bigcup_{k=1}^{n_0} (C_{i,j} \cap C_k)) \right\| = \left\| \sum_{k=1}^{n_0} \sum_{i=1}^n \sum_{j=1}^{n_i} \mu(C_k \cap C_{i,j}) + \sum_{k=1}^{n_0} \sum_{i=1}^{\bar{n}} \mu(C_k \cap \bar{C}_i) - \sum_{i=1}^n \sum_{j=1}^{n_i} \sum_{k=1}^{n_0} \mu(C_{i,j} \cap C_k) \right\| = \left\| \sum_{i=1}^{\bar{n}} \sum_{k=1}^{n_0} \mu(C_k \cap \bar{C}_i) \right\| = \left\| \sum_{i=1}^{\bar{n}} \mu(\bar{C}_i) \right\| \leq \sum_{i=1}^{\bar{n}} \|\mu(\bar{C}_i)\| \leq \sum_{i=1}^{\bar{n}} \nu(\bar{C}_i) = \sum_{i=1}^{\bar{n}} \bar{\nu}(\bar{C}_i) = \bar{\nu}(\bigcup_{i=1}^{\bar{n}} \bar{C}_i) = \bar{\nu}(\bar{A}) < \epsilon$, where the first equality follows from the definition of $\bar{\mu}$; the fourth equality follows from (v); the sixth equality follows from (v); the second inequality follows from (vi); the seventh equality follows from $\bar{\nu}|_{\mathcal{C}} = \nu$; the eighth equality follows from the fact that $\bar{\nu}$ is a measure on the algebra $\mathcal{A}$. Hence, $\bar{\mu}(A) = \sum_{i=1}^{\infty} \bar{\mu}(A_i)$.

$\forall A \in \mathcal{A}$ with $\bar{\nu}(A) < \infty$, then $A = \bigcup_{i=1}^{n_0} C_i$, where $n_0 \in \mathbb{Z}_+$ and $(C_i)_{i=1}^{n_0} \subseteq \mathcal{C}$ is pairwise disjoint. $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(A_i)_{i=1}^n \subseteq \mathcal{A}$ with $A = \bigcup_{i=1}^n A_i$, $\forall i \in \{1, \dots, n\}$, $A_i \in \mathcal{A}$ implies that $A_i = \bigcup_{j=1}^{n_i} C_{i,j}$, where $n_i \in \mathbb{Z}_+$ and $(C_{i,j})_{j=1}^{n_i} \subseteq \mathcal{C}$ is pairwise disjoint. Then, $\sum_{i=1}^n \|\bar{\mu}(A_i)\| = \sum_{i=1}^n \left\| \sum_{j=1}^{n_i} \mu(C_{i,j}) \right\| \leq \sum_{i=1}^n \sum_{j=1}^{n_i} \|\mu(C_{i,j})\| \leq \sum_{i=1}^n \sum_{j=1}^{n_i} \nu(C_{i,j}) = \sum_{i=1}^n \sum_{j=1}^{n_i} \bar{\nu}(C_{i,j}) = \sum_{i=1}^n \bar{\nu}(A_i) = \bar{\nu}(A)$, where the first equality follows from the definition of $\bar{\mu}$; the second inequality follows from (vi); the second equality follows from $\bar{\nu}|_{\mathcal{C}} = \nu$; and the last two equalities follow from the fact that $\bar{\nu}$ is a measure on the algebra $\mathcal{A}$. Hence, $s_A := \sup_{n \in \mathbb{Z}_+, (A_i)_{i=1}^n \subseteq \mathcal{A}, A = \bigcup_{i=1}^n A_i, A_i \cap A_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\bar{\mu}(A_i)\| \leq \bar{\nu}(A)$. On the other hand, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall i \in \{1, \dots, n_0\}$, by (vi), $\exists n_i \in \mathbb{Z}_+$, $\exists$ pairwise disjoint $(C_{i,j})_{j=1}^{n_i} \subseteq \mathcal{C}$ with $C_i = \bigcup_{j=1}^{n_i} C_{i,j}$ such that $\bar{\nu}(C_i) = \nu(C_i) < \sum_{j=1}^{n_i} \|\mu(C_{i,j})\| + \epsilon/(1+n_0) = \sum_{j=1}^{n_i} \|\bar{\mu}(C_{i,j})\| + \epsilon/(1+n_0)$. Then, $\bar{\nu}(A) = \sum_{i=1}^{n_0} \bar{\nu}(C_i) < \sum_{i=1}^{n_0} \sum_{j=1}^{n_i} \|\bar{\mu}(C_{i,j})\| + \epsilon \leq s_A + \epsilon$. By the arbitrariness of ϵ , we have $\bar{\nu}(A) \leq s_A$. Hence, $s_A = \bar{\nu}(A)$. This shows that $\bar{\mu}$ is a $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}$ with $\mathcal{P} \circ \bar{\mu} = \bar{\nu}$.

To show the uniqueness of $\bar{\mu}$, let $\hat{\mu}$ be another $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}$ with $\hat{\mu}|_{\mathcal{C}} = \mu$ and $(\mathcal{P} \circ \hat{\mu})|_{\mathcal{C}} = \nu$. Since $\mathcal{P} \circ \hat{\mu}$ is a measure on the algebra $\mathcal{A}$ that is an extension of ν , then $\mathcal{P} \circ \hat{\mu} = \bar{\nu} = \mathcal{P} \circ \bar{\mu}$ by Proposition 11.32. $\forall A \in \mathcal{A}$ with $\bar{\nu}(A) = \infty$, then $\hat{\mu}(A)$ and $\bar{\mu}(A)$ are both undefined. $\forall A \in \mathcal{A}$ with $\bar{\nu}(A) < \infty$, then $A = \bigcup_{i=1}^n C_i$, where $n \in \mathbb{Z}_+$ and $(C_i)_{i=1}^n \subseteq \mathcal{C}$ is pairwise disjoint. Then, $\hat{\mu}(A) = \sum_{i=1}^n \hat{\mu}(C_i) = \sum_{i=1}^n \mu(C_i) = \bar{\mu}(A)$, where the first equality follows from the fact that $\hat{\mu}$ is a $\mathcal{Y}$ -valued measure on the algebra $\mathcal{A}$; the second equality follows from $\hat{\mu}|_{\mathcal{C}} = \mu$; and the last equality follows from the definition of $\bar{\mu}$. Hence, $\hat{\mu} = \bar{\mu}$. Therefore, $\bar{\mu}$ is unique.

This completes the proof of the proposition. $\square$

Proposition 12.21 *Let $m \in \mathbb{N}$, $\mathcal{X}_j := (X_j, \mathcal{B}_j, \mu_j)$ be a finite $\mathbb{K}$ -valued measure space, $\forall j \in \{1, \dots, m\}$, $\mathcal{C}_m := \{\prod_{j=1}^m E_j \subseteq \prod_{j=1}^m X_j \mid E_j \in \mathcal{B}_j, \forall j \in \{1, \dots, m\}\}$, which is the set of measurable rectangles in $\prod_{j=1}^m X_j$, $\bar{\mathcal{B}}_m$ be the σ -algebra on $\prod_{j=1}^m X_j$ generated by $\mathcal{C}_m$, $\mu : \mathcal{C}_m \rightarrow \mathbb{K}$ be defined by $\mu(\prod_{j=1}^m E_j) = \prod_{j=1}^m \mu_j(E_j)$, $\forall \prod_{j=1}^m E_j \in \mathcal{C}_m$, and $\nu : \mathcal{C}_m \rightarrow [0, \infty) \subset \mathbb{R}$ be defined by $\nu(\prod_{j=1}^m E_j) = \prod_{j=1}^m \mathcal{P} \circ \mu_j(E_j)$, $\forall \prod_{j=1}^m E_j \in \mathcal{C}_m$. Then, $\mathcal{C}_m$ is a semialgebra on $\prod_{j=1}^m X_j$, μ and ν satisfy the assumptions of Proposition 12.20. Therefore, there exists a unique $\mathbb{K}$ -valued measure $\lambda := \prod_{j=1}^m \mu_j$ on the measurable space $(\prod_{j=1}^m X_j, \bar{\mathcal{B}}_m)$ such that $\lambda(\prod_{j=1}^m E_j) = \prod_{j=1}^m \mu_j(E_j)$ and $\mathcal{P} \circ \lambda(\prod_{j=1}^m E_j) = \prod_{j=1}^m \mathcal{P} \circ \mu_j(E_j)$, $\forall \prod_{j=1}^m E_j \in \mathcal{C}_m$. Furthermore, λ is finite. The $\mathbb{K}$ -valued measure space $\mathcal{X} := (\prod_{j=1}^m X_j, \bar{\mathcal{B}}_m, \lambda)$ is said to be the product $\mathbb{K}$ -valued measure space of $\mathcal{X}_1, \dots, \mathcal{X}_m$, and denoted by $\prod_{j=1}^m \mathcal{X}_j$. λ is said to be the product $\mathbb{K}$ -valued measure of $\mu_1, \dots, \mu_m$ and denoted by $\prod_{j=1}^m \mu_j$.*

Proof We will prove the result using mathematical induction on m .

1° $m = 1$. The result is trivial and $\prod_{j=1}^1 \mathcal{X}_1 = \mathcal{X}_1$ and $\prod_{j=1}^1 \mu_1 = \mu_1$.

2° Assume that the result holds $\forall m \in \{1, \dots, k-1\}$ with $k \in \mathbb{N}$ and $k \geq 2$.

3° Consider the case where $m = k$. Let $\bar{\lambda} := \prod_{j=1}^{k-1} \mu_j$ by the inductive assumption.

$\forall \prod_{j=1}^k E_{j,1}, \prod_{j=1}^k E_{j,2} \in \mathcal{C}_k$, we have $(\prod_{j=1}^k E_{j,1}) \cap (\prod_{j=1}^k E_{j,2}) = \prod_{j=1}^k (E_{j,1} \cap E_{j,2}) \in \mathcal{C}_m$. Note that $(\prod_{j=1}^k E_{j,1})^\sim = ((\prod_{j=1}^{k-1} E_{j,1})^\sim \times E_{k,1}) \cup ((\prod_{j=1}^{k-1} E_{j,1})^\sim \times \widetilde{E_{k,1}}) \cup ((\prod_{j=1}^{k-1} E_{j,1})^\sim \times \widetilde{E_{k,1}})$. By inductive assumption, $\prod_{j=1}^{k-1} E_{j,1} \in \mathcal{C}_{k-1}$ and $\mathcal{C}_{k-1}$ is a semialgebra on $\prod_{j=1}^{k-1} X_j$, then $(\prod_{j=1}^{k-1} E_{j,1})^\sim$ is a finite disjoint union of sets in $\mathcal{C}_{k-1}$. Hence, $(\prod_{j=1}^k E_{j,1})^\sim$ is a finite disjoint union of sets in $\mathcal{C}_k$. Clearly, $\emptyset = \prod_{j=1}^k \emptyset \in \mathcal{C}_k \neq \emptyset$. Hence, $\mathcal{C}_k$ is a semialgebra on $X_1 \times X_2$.

Next, we will show that μ and ν satisfy (i) – (vi) of Proposition 12.20. $\mu(\emptyset) = \mu(\prod_{j=1}^k \emptyset) = \prod_{j=1}^k \mu_j(\emptyset) = 0$ and $\nu(\emptyset) = \nu(\prod_{j=1}^k \emptyset) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(\emptyset) = 0$. Hence, (i) of Proposition 12.20 holds.

(iv) of Proposition 12.20 holds trivially since $\nu(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_j) \leq \prod_{j=1}^k \mathcal{P} \circ \mu_j(X_j) < \infty$, $\forall \prod_{j=1}^k E_j \in \mathcal{C}_k$.

$\forall \prod_{j=1}^k E_j \in \mathcal{C}_k$, $\forall$ pairwise disjoint $(\prod_{j=1}^k E_{j,i})_{i=1}^\infty \subseteq \mathcal{C}_k$ with $\prod_{j=1}^k E_j = \bigcup_{i=1}^\infty (\prod_{j=1}^k E_{j,i})$. Fix any $x_k \in E_k$. $\forall (x_1, \dots, x_{k-1}) \in \prod_{j=1}^{k-1} E_j$, $\exists! i \in \mathbb{N}$ such that $(x_1, \dots, x_k) \in \prod_{j=1}^k E_{j,i}$. Then, $\prod_{j=1}^{k-1} E_j = \bigcup_{i \in \mathbb{N}, x_k \in E_{k,i}} \prod_{j=1}^{k-1} E_{j,i}$ and the sets in the union are pairwise disjoint.

Then, $\sum_{i \in \mathbb{N}} \chi_{E_{k,i}, X_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) = \sum_{i \in \mathbb{N}, x_k \in E_{k,i}} \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) =$

$\sum_{i \in \mathbb{N}, x_k \in E_{k,i}} \mathcal{P} \circ \bar{\lambda}(\prod_{j=1}^{k-1} E_{j,i}) = \mathcal{P} \circ \bar{\lambda}(\bigcup_{i \in \mathbb{N}, x_k \in E_{k,i}} \prod_{j=1}^{k-1} E_{j,i}) = \mathcal{P} \circ \bar{\lambda}(\prod_{j=1}^{k-1} E_j) =$

$\chi_{E_k, X_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_j) \in [0, \infty) \subset \mathbb{R}$, where the last four equalities fol-

low from the inductive assumption. $\forall x_k \in X_k \setminus E_k$, we have $\chi_{E_k, X_k}(x_k) \cdot \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_j) = 0 = \sum_{i \in \mathbb{N}} \chi_{E_{k,i}, X_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i})$. Then, $\forall x_k \in X_k$, $\chi_{E_k, X_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_j) = \sum_{i \in \mathbb{N}} \chi_{E_{k,i}, X_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i})$. By Monotone Convergence Theorem 11.81 and Proposition 11.75, $\nu(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_j) = \int_{X_k} \chi_{E_k, X_k} \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_j) d\mathcal{P} \circ \mu_k = \int_{X_k} \sum_{i \in \mathbb{N}} \chi_{E_{k,i}, X_k} \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) d\mathcal{P} \circ \mu_k = \sum_{i=1}^\infty (\prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i})) \mathcal{P} \circ \mu_k(E_{k,i}) = \sum_{i=1}^\infty \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_{j,i}) = \sum_{i=1}^\infty \nu(\prod_{j=1}^k E_{j,i})$. Hence, (ii) and (iii) of Proposition 12.20 hold.

$\forall \prod_{j=1}^k E_j \in \mathcal{C}_k, \forall n \in \mathbb{Z}_+, \forall$ pairwise disjoint $\left(\prod_{j=1}^k E_{j,i}\right)_{i=1}^n \subseteq \mathcal{C}_k$ with $\prod_{j=1}^k E_j = \bigcup_{i=1}^n \prod_{j=1}^k E_{j,i}$. Fix any $x_k \in E_k$. $\forall (x_1, \dots, x_{k-1}) \in \prod_{j=1}^{k-1} E_j, \exists! i \in N := \{1, \dots, n\}$ such that $(x_1, \dots, x_k) \in \prod_{j=1}^k E_{j,i}$. Then, $\prod_{j=1}^{k-1} E_j = \bigcup_{i \in N, x_k \in E_{k,i}} \prod_{j=1}^{k-1} E_{j,i}$ and the sets in the union are pairwise disjoint. Then, $\sum_{i \in N} \chi_{E_{k,i}, X_k}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = \sum_{i \in N, x_k \in E_{k,i}} \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = \sum_{i \in N, x_k \in E_{k,i}} \bar{\lambda}\left(\prod_{j=1}^{k-1} E_{j,i}\right) = \bar{\lambda}\left(\bigcup_{i \in N, x_k \in E_{k,i}} \prod_{j=1}^{k-1} E_{j,i}\right) = \bar{\lambda}\left(\prod_{j=1}^{k-1} E_j\right) = \chi_{E_k, X_k}(x_k) \prod_{j=1}^{k-1} \mu_j(E_j) \in \mathbb{K}$, where the last four equalities follow from the inductive assumption. $\forall x_k \in X_k \setminus E_k$, we have $\chi_{E_k, X_k}(x_k) \prod_{j=1}^{k-1} \mu_j(E_j) = 0 = \sum_{i \in N} \chi_{E_{k,i}, X_k}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i})$. Then, $\forall x_1 \in X_1, \chi_{E_k, X_k}(x_k) \cdot \prod_{j=1}^{k-1} \mu_j(E_j) = \sum_{i \in N} \chi_{E_{k,i}, X_k}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i})$. By Propositions 11.125 and 11.132, we have $\mu(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mu_j(E_j) = \int_{X_k} \chi_{E_k, X_k} \cdot \prod_{j=1}^{k-1} \mu_j(E_j) d\mu_k = \int_{X_k} (\sum_{i \in N} \chi_{E_{k,i}, X_k} \prod_{j=1}^{k-1} \mu_j(E_{j,i})) d\mu_k = \sum_{i=1}^n \prod_{j=1}^k \mu_j(E_{j,i}) = \sum_{i=1}^n \mu(\prod_{j=1}^k E_{j,i})$. Hence, (v) of Proposition 12.20 holds.

$\forall \prod_{j=1}^k E_j \in \mathcal{C}_k, \forall n \in \mathbb{Z}_+, \forall$ pairwise disjoint $\left(\prod_{j=1}^k E_{j,i}\right)_{i=1}^n \subseteq \mathcal{C}_k$ with $\prod_{j=1}^k E_j = \bigcup_{i=1}^n \prod_{j=1}^k E_{j,i}$, we have $\sum_{i=1}^n \left| \mu(\prod_{j=1}^k E_{j,i}) \right| = \sum_{i=1}^n \prod_{j=1}^k |\mu_j(E_{j,i})| \leq \sum_{i=1}^n \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_{j,i}) = \sum_{i=1}^n \nu(\prod_{j=1}^k E_{j,i}) = \nu(\prod_{j=1}^k E_j)$, where the inequality follows from Definition 11.108; and the last equality follows from (ii). Then, $\nu(\prod_{j=1}^k E_j) \geq s_{\prod_{j=1}^k E_j} := \sup_{n \in \mathbb{Z}_+, (\prod_{j=1}^k E_{j,i})_{i=1}^n \subseteq \mathcal{C}_k, \prod_{j=1}^k E_j = \bigcup_{i=1}^n \prod_{j=1}^k E_{j,i}, \sum_{i=1}^n \left| \mu(\prod_{j=1}^k E_{j,i}) \right|}$. On the other hand, $\forall \epsilon \in (0, 1] \subset \mathbb{R}, \forall j \in \{1, \dots, k\}, \exists n_j \in \mathbb{Z}_+, \exists$ pairwise disjoint $(E_{j,i})_{i=1}^{n_j} \subseteq \mathcal{B}_j$ with $E_j = \bigcup_{i=1}^{n_j} E_{j,i}$ such that $\sum_{i=1}^{n_j} |\mu_j(E_{j,i})| \geq \mathcal{P} \circ \mu_j(E_j) (1 - \frac{\epsilon}{1 + \nu(\prod_{j=1}^k E_j)})^{1/k}$. Then, $\prod_{j=1}^k E_j = \bigcup_{i_1=1}^{n_1} \dots \bigcup_{i_k=1}^{n_k} (\prod_{j=1}^k E_{j,i_j})$ and the sets in the union are pairwise disjoint and in $\mathcal{C}_k$. Then, $\sum_{i_1=1}^{n_1} \dots \sum_{i_k=1}^{n_k} \left| \mu(\prod_{j=1}^k E_{j,i_j}) \right| = \sum_{i_1=1}^{n_1} \dots \sum_{i_k=1}^{n_k} \prod_{j=1}^k |\mu_j(E_{j,i_j})| = \prod_{j=1}^k (\sum_{i_j=1}^{n_j} |\mu_j(E_{j,i_j})|) \geq (1 - \frac{\epsilon}{1 + \nu(\prod_{j=1}^k E_j)}) \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_j) > \nu(\prod_{j=1}^k E_j) - \epsilon$. By the arbitrariness of ϵ , we have $s_{\prod_{j=1}^k E_j} \geq \nu(\prod_{j=1}^k E_j)$. Hence, (vi) of Proposition 12.20 holds.

Let $\mathcal{A}_k$ be the algebra on $\prod_{j=1}^k X_j$ generated by $\mathcal{C}_k$. By Proposition 12.20, ν admits a unique extension to a measure $\bar{\nu}$ on the algebra $\mathcal{A}_k$; and μ admits a unique extension to a $\mathbb{K}$ -valued measure $\bar{\mu}$ on the

algebra $\mathcal{A}_k$ with $(\mathcal{P} \circ \bar{\mu})|_{\mathcal{C}_k} = \nu$. Furthermore, $\mathcal{P} \circ \bar{\mu} = \bar{\nu}$. Clearly, $\bar{\mu}$ is finite. By Carathéodory Extension Theorem 12.4, there is a unique $\mathbb{K}$ -valued measure λ on the measurable space $(\prod_{j=1}^k X_j, \bar{\mathcal{B}}_k)$ such that $\lambda|_{\mathcal{A}} = \bar{\mu}$ and $(\mathcal{P} \circ \lambda)|_{\mathcal{A}} = \mathcal{P} \circ \bar{\mu} = \bar{\nu}$. Then, $\lambda|_{\mathcal{C}_k} = \mu$ and $(\mathcal{P} \circ \lambda)|_{\mathcal{C}_k} = \nu$. This implies that $\mathcal{P} \circ \lambda(\prod_{j=1}^k X_j) = \nu(\prod_{j=1}^k X_j) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(X_j) < \infty$. Hence, λ is finite.

This completes the induction process.

This completes the proof of the proposition. $\square$

Proposition 12.22 *Let $k \in \mathbb{N}$, $\mathcal{X}_j := (X_j, \mathcal{B}_j, \mu_j)$ be a σ -finite $\mathbb{K}$ -valued measure space, $\forall j \in \{1, \dots, k\}$, $\mathcal{C} := \left\{ \prod_{j=1}^k E_j \subseteq \prod_{j=1}^k X_j \mid E_j \in \mathcal{B}_j, \forall j \in \{1, \dots, k\} \right\}$ be the set of measurable rectangles, $\mathcal{B}$ be the σ -algebra on $\prod_{j=1}^k X_j$ generated by $\mathcal{C}$. Then, there is a unique $\mathbb{K}$ -valued measure $\lambda := \prod_{j=1}^k \mu_j$ on the measurable space $(\prod_{j=1}^k X_j, \mathcal{B})$ such that $\lambda(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mu_j(E_j)$ and $\mathcal{P} \circ \lambda(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_j)$, $\forall E_j \in \text{dom}(\mu_j), \forall j \in \{1, \dots, k\}$. Furthermore, λ is σ -finite. The $\mathbb{K}$ -valued measure space $\mathcal{X} := (\prod_{j=1}^k X_j, \mathcal{B}, \lambda)$ is said to be the product $\mathbb{K}$ -valued measure space of $\mathcal{X}_1, \dots, \mathcal{X}_k$ and denoted by $\prod_{j=1}^k \mathcal{X}_j$. λ is said to be the product $\mathbb{K}$ -valued measure of $\mu_1, \dots, \mu_k$ and denoted by $\prod_{j=1}^k \mu_j$.*

Proof Fix any $j \in \{1, \dots, k\}$. Since $\mathcal{X}_j$ is σ -finite, then $\exists (X_{j,n})_{n=1}^\infty \subseteq \mathcal{B}_j$ such that $X_j = \bigcup_{n=1}^\infty X_{j,n}$ and $\mathcal{P} \circ \mu_j(X_{j,n}) < \infty, \forall n \in \mathbb{N}$. Without loss of generality, we may assume that $(X_{j,n})_{n=1}^\infty$ is pairwise disjoint. Let $\mathcal{X}_{j,n} := (X_{j,n}, \mathcal{B}_{j,n}, \mu_{j,n})$ be the finite $\mathbb{K}$ -valued measure subspace of $\mathcal{X}_j, \forall n \in \mathbb{N}$. Fix any $n_1, \dots, n_k \in \mathbb{N}$. Let $\mathcal{C}_{n_1, \dots, n_k} := \left\{ \prod_{j=1}^k E_j \subseteq \prod_{j=1}^k X_{j,n_j} \mid E_j \in \mathcal{B}_{j,n_j}, j = 1, \dots, k \right\}$ and $\bar{\mathcal{B}}_{n_1, \dots, n_k}$ be the σ -algebra on $\prod_{j=1}^k X_{j,n_j}$ generated by $\mathcal{C}_{n_1, \dots, n_k}$. By Proposition 12.21, there is a unique $\mathbb{K}$ -valued measure $\lambda_{n_1, \dots, n_k} = \prod_{j=1}^k \mu_{j,n_j}$ on $(\prod_{j=1}^k X_{j,n_j}, \bar{\mathcal{B}}_{n_1, \dots, n_k})$ such that $\lambda_{n_1, \dots, n_k}(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mu_{j,n_j}(E_j)$ and $\mathcal{P} \circ \lambda_{n_1, \dots, n_k}(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mathcal{P} \circ \mu_{j,n_j}(E_j), \forall \prod_{j=1}^k E_j \in \mathcal{C}_{n_1, \dots, n_k}$. Furthermore, $\lambda_{n_1, \dots, n_k}$ is finite. Denote the finite $\mathbb{K}$ -valued measure space $(\prod_{j=1}^k X_{j,n_j}, \bar{\mathcal{B}}_{n_1, \dots, n_k}, \lambda_{n_1, \dots, n_k}) =: \bar{\mathcal{X}}_{n_1, \dots, n_k}$. By Proposition 11.118, the generation process on $(\bar{\mathcal{X}}_{n_1, \dots, n_k})_{n_j \in \mathbb{N}, j \in \{1, \dots, k\}}$ yields a unique σ -finite $\mathbb{K}$ -valued measure space $\mathcal{X} := (\prod_{j=1}^k X_j, \bar{\mathcal{B}}, \lambda)$ on $\prod_{j=1}^k X_j$ such that $\bar{\mathcal{X}}_{n_1, \dots, n_k}$ is the finite $\mathbb{K}$ -valued measure subspace of $\mathcal{X}, \forall n_1, \dots, n_k \in \mathbb{N}$, where $\bar{\mathcal{B}} := \left\{ E \subseteq \prod_{j=1}^k X_j \mid E \cap (\prod_{j=1}^k X_{j,n_j}) \in \bar{\mathcal{B}}_{n_1, \dots, n_k}, \forall n_j \in \mathbb{N}, \forall j \in \{1, \dots, k\} \right\}$, $\mathcal{P} \circ \lambda(E) = \sum_{n_1=1}^\infty \cdots \sum_{n_k=1}^\infty \mathcal{P} \circ \lambda_{n_1, \dots, n_k}(E \cap (\prod_{j=1}^k X_{j,n_j}))$, $\forall E \in \bar{\mathcal{B}}$, and $\lambda(E) = \sum_{n_1=1}^\infty \cdots \sum_{n_k=1}^\infty \lambda_{n_1, \dots, n_k}(E \cap (\prod_{j=1}^k X_{j,n_j}))$, $\forall E \in \bar{\mathcal{B}}$ with $\mathcal{P} \circ \lambda(E) < \infty$.

Claim 12.22.1 $\mathcal{B} = \bar{\mathcal{B}}$.

Proof of claim: $\forall \prod_{j=1}^k E_j \in \mathcal{C}, \forall n_1, \dots, n_k \in \mathbb{N}, E_j \cap X_{j,n_j} \in \mathcal{B}_{j,n_j}, \forall j \in \{1, \dots, k\}$, by Proposition 11.115. Then, $(\prod_{j=1}^k E_j) \cap (\prod_{j=1}^k X_{j,n_j}) \in \mathcal{C}_{n_1, \dots, n_k} \subseteq \bar{\mathcal{B}}_{n_1, \dots, n_k}$. This implies that $\prod_{j=1}^k E_j \in \bar{\mathcal{B}}$. Then, $\mathcal{C} \subseteq \bar{\mathcal{B}}$. Since $\bar{\mathcal{B}}$ is a σ -algebra, then $\mathcal{B} \subseteq \bar{\mathcal{B}}$.

On the other hand, $\forall n_1, \dots, n_k \in \mathbb{N}, \forall \prod_{j=1}^k E_j \in \mathcal{C}_{n_1, \dots, n_k}$, we have $E_j \in \mathcal{B}_{j,n_j} \subseteq \mathcal{B}_j, \forall j \in \{1, \dots, k\}$. Thus, $\prod_{j=1}^k E_j \in \mathcal{C} \subseteq \mathcal{B}$. Then, $\mathcal{C}_{n_1, \dots, n_k} \subseteq \mathcal{B}$ and $\bar{\mathcal{B}}_{n_1, \dots, n_k} \subseteq \mathcal{B}$. This further implies that $\bar{\mathcal{B}} \subseteq \mathcal{B}$. Hence, $\mathcal{B} = \bar{\mathcal{B}}$. This completes the proof of the claim. $\square$

$\forall j \in \{1, \dots, k\}, \forall E_j \in \text{dom}(\mu_j)$, we have $\infty > \mathcal{P} \circ \mu_j(E_j) = \sum_{n_j=1}^{\infty} \mathcal{P} \circ \mu_j(E_j \cap X_{j,n_j}) = \sum_{n_j=1}^{\infty} \mathcal{P} \circ \mu_{j,n_j}(E_j \cap X_{j,n_j}), \mathbb{K} \ni \mu_j(E_j) = \sum_{n_j=1}^{\infty} \mu_j(E_j \cap X_{j,n_j}) = \sum_{n_j=1}^{\infty} \mu_{j,n_j}(E_j \cap X_{j,n_j})$. Then,

$$\begin{aligned} \mathcal{P} \circ \lambda \left(\prod_{j=1}^k E_j \right) &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \mathcal{P} \circ \lambda_{n_1, \dots, n_k} \left(\left(\prod_{j=1}^k E_j \right) \cap \left(\prod_{j=1}^k X_{j,n_j} \right) \right) \\ &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \mathcal{P} \circ \lambda_{n_1, \dots, n_k} \left(\prod_{j=1}^k (E_j \cap X_{j,n_j}) \right) \\ &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \prod_{j=1}^k \mathcal{P} \circ \mu_{j,n_j}(E_j \cap X_{j,n_j}) \\ &= \prod_{j=1}^k \left(\sum_{n_j=1}^{\infty} \mathcal{P} \circ \mu_{j,n_j}(E_j \cap X_{j,n_j}) \right) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_j) \\ &< \infty \end{aligned}$$

where the first equality follows from Proposition 11.118; and the third equality follows from Proposition 12.21. We also have

$$\begin{aligned} \lambda \left(\prod_{j=1}^k E_j \right) &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \lambda_{n_1, \dots, n_k} \left(\left(\prod_{j=1}^k E_j \right) \cap \left(\prod_{j=1}^k X_{j,n_j} \right) \right) \\ &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \lambda_{n_1, \dots, n_k} \left(\prod_{j=1}^k (E_j \cap X_{j,n_j}) \right) \\ &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \prod_{j=1}^k \mu_{j,n_j}(E_j \cap X_{j,n_j}) \\ &= \prod_{j=1}^k \left(\sum_{n_j=1}^{\infty} \mu_{j,n_j}(E_j \cap X_{j,n_j}) \right) = \prod_{j=1}^k \mu_j(E_j) \in \mathbb{K} \end{aligned}$$

where the first equality follows from Proposition 11.118; and the third equality follows from Proposition 12.21. Hence, $\mathcal{X}$ is the σ -finite $\mathbb{K}$ -valued

measure space that we seek.

This completes the proof of the proposition. $\square$

Proposition 12.23 *Let $m \in \mathbb{N}$, $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a σ -finite $\mathbb{K}$ -valued measure space, $\bar{X}_i \in \mathcal{B}_i$, $\bar{\mathcal{X}}_i := (\bar{X}_i, \bar{\mathcal{B}}_i, \bar{\mu}_i)$ be the σ -finite $\mathbb{K}$ -valued measure subspace of $\mathcal{X}_i$, $i = 1, \dots, m$, and $\mathcal{X} := (\prod_{i=1}^m X_i, \mathcal{B}, \lambda) := \prod_{i=1}^m \mathcal{X}_i$ be the σ -finite product $\mathbb{K}$ -valued measure space of $\mathcal{X}_1, \dots, \mathcal{X}_m$. Then, $\bar{\mathcal{X}} := (\prod_{i=1}^m \bar{X}_i, \bar{\mathcal{B}}, \bar{\lambda})$ is the $\mathbb{K}$ -valued measure subspace of $\mathcal{X}$ if, and only if, $\bar{\mathcal{X}} = \prod_{i=1}^m \bar{\mathcal{X}}_i$.*

Proof Let $\bar{\mathcal{X}} := (\prod_{i=1}^m \bar{X}_i, \bar{\mathcal{B}}, \bar{\lambda})$ be the σ -finite $\mathbb{K}$ -valued measure subspace of $\mathcal{X}$ and $\tilde{\mathcal{X}} := \prod_{i=1}^m \tilde{\mathcal{X}}_i := (\prod_{i=1}^m \bar{X}_i, \tilde{\mathcal{B}}, \tilde{\lambda})$ be the σ -finite product $\mathbb{K}$ -valued measure space of $\bar{\mathcal{X}}_1, \dots, \bar{\mathcal{X}}_m$. All we need to show is that $\bar{\mathcal{X}} = \tilde{\mathcal{X}}$.

Let $\bar{X}_{i,1} := \bar{X}_i$ and $\bar{X}_{i,2} := X_i \setminus \bar{X}_i$, $i = 1, \dots, m$, $\bar{\mathcal{C}}_{n_1, \dots, n_m} := \{\prod_{i=1}^m B_i \subseteq \prod_{i=1}^m \bar{X}_{i,n_i} \mid B_i \in \mathcal{B}_i, \forall i \in \{1, \dots, m\}\}$ be the set of measurable rectangles on $\prod_{i=1}^m \bar{X}_{i,n_i}$, $\tilde{\mathcal{B}}_{n_1, \dots, n_m}$ be the σ -algebra generated by $\bar{\mathcal{C}}_{n_1, \dots, n_m}$ on $\prod_{i=1}^m \bar{X}_{i,n_i}$, $\forall n_1, \dots, n_m \in \{1, 2\}$, and $\mathcal{C} := \{\prod_{i=1}^m B_i \subseteq \prod_{i=1}^m X_i \mid B_i \in \mathcal{B}_i, \forall i \in \{1, \dots, m\}\}$ be the set of measurable rectangles on $\prod_{i=1}^m X_i$. Then, by Proposition 12.22, $\mathcal{B}$ is the σ -algebra generated by $\mathcal{C}$ on $\prod_{i=1}^m X_i$.

Claim 12.23.1 $\mathcal{B} = \left\{ E \subseteq \prod_{i=1}^m X_i \mid E \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) \in \tilde{\mathcal{B}}_{n_1, \dots, n_m}, \forall n_i \in \{1, 2\}, \forall i \in \{1, \dots, m\} \right\} =: \hat{\mathcal{B}}$.

Proof of claim: $\forall \prod_{i=1}^m B_i \in \mathcal{C}$, we have $B_i \in \mathcal{B}_i$, $i = 1, \dots, m$, which implies that $\mathcal{B}_i \ni B_i \cap \bar{X}_{i,n_i} \subseteq \bar{X}_{i,n_i}$, $n_i = 1, 2$, $i = 1, \dots, m$. Hence, $\prod_{i=1}^m B_i \in \hat{\mathcal{B}}$. By the arbitrariness of $\prod_{i=1}^m B_i$, we have $\mathcal{C} \subseteq \hat{\mathcal{B}}$.

Now, we will show that $\hat{\mathcal{B}}$ is a σ -algebra on $\prod_{i=1}^m X_i$. Clearly, $\emptyset = \prod_{i=1}^m \emptyset \in \mathcal{C} \subseteq \hat{\mathcal{B}}$ and $\prod_{i=1}^m X_i \in \mathcal{C} \subseteq \hat{\mathcal{B}}$. $\forall (E_n)_{n=1}^\infty \subseteq \hat{\mathcal{B}}$, $\forall n_i = 1, 2$, $\forall i = 1, \dots, m$, $\forall n \in \mathbb{N}$, by the definition of $\hat{\mathcal{B}}$, $E_n \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) \in \tilde{\mathcal{B}}_{n_1, \dots, n_m}$, which implies $(\bigcup_{n=1}^\infty E_n) \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) = \bigcup_{n=1}^\infty (E_n \cap (\prod_{i=1}^m \bar{X}_{i,n_i})) \in \tilde{\mathcal{B}}_{n_1, \dots, n_m}$, since $\tilde{\mathcal{B}}_{n_1, \dots, n_m}$ is a σ -algebra. Hence, $\bigcup_{n=1}^\infty E_n \in \hat{\mathcal{B}}$. $\forall E \in \hat{\mathcal{B}}$, $\forall n_i = 1, 2$, $\forall i = 1, \dots, m$, $E \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) \in \tilde{\mathcal{B}}_{n_1, \dots, n_m}$, then, $((\prod_{i=1}^m X_i) \setminus E) \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) = (\prod_{i=1}^m \bar{X}_{i,n_i}) \setminus E = (\prod_{i=1}^m \bar{X}_{i,n_i}) \setminus (E \cap (\prod_{i=1}^m \bar{X}_{i,n_i})) \in \tilde{\mathcal{B}}_{n_1, \dots, n_m}$, since $\tilde{\mathcal{B}}_{n_1, \dots, n_m}$ is a σ -algebra on $\prod_{i=1}^m \bar{X}_{i,n_i}$. This implies that $(\prod_{i=1}^m X_i) \setminus E \in \hat{\mathcal{B}}$. Hence, $\hat{\mathcal{B}}$ is a σ -algebra on $\prod_{i=1}^m X_i$.

Since $\mathcal{B}$ is the σ -algebra on $\prod_{i=1}^m X_i$ generated by $\mathcal{C}$ and $\mathcal{C} \subseteq \hat{\mathcal{B}}$, then $\mathcal{B} \subseteq \hat{\mathcal{B}}$.

On the other hand, $\bar{\mathcal{C}}_{n_1, \dots, n_m} \subseteq \mathcal{C}$, $n_i = 1, 2$, $i = 1, \dots, m$, implies that $\tilde{\mathcal{B}}_{n_1, \dots, n_m} \subseteq \mathcal{B}$. $\forall E \in \hat{\mathcal{B}}$, we have $E \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) \in \tilde{\mathcal{B}}_{n_1, \dots, n_m} \subseteq \mathcal{B}$, $n_i = 1, 2$, $i = 1, \dots, m$, which implies that $E = \bigcup_{n_1=1}^2 \cdots \bigcup_{n_m=1}^2 (E \cap (\prod_{i=1}^m \bar{X}_{i,n_i})) \in \mathcal{B}$. Hence, $\hat{\mathcal{B}} \subseteq \mathcal{B}$. Therefore, $\mathcal{B} = \hat{\mathcal{B}}$. This completes the proof of the claim. $\square$

Note that $\tilde{\mathcal{B}} = \tilde{\mathcal{B}}_{1,\dots,1} = \left\{ E \subseteq \prod_{i=1}^m \bar{X}_i \mid E \in \hat{\mathcal{B}} \right\} = \left\{ E \subseteq \prod_{i=1}^m \bar{X}_i \mid E \in \mathcal{B} \right\} = \bar{\mathcal{B}}$, where the first equality follows from Proposition 12.22; the second equality follows from the definition of $\hat{\mathcal{B}}$; the third equality follows from Claim 12.23.1; and the last equality follows from Proposition 11.115.

Finally, we will show that $\bar{\lambda} = \tilde{\lambda}$ by Propositions 12.22 and 11.115. By Proposition 11.115 and Claim 12.23.1, $\bar{\lambda}$ is a $\mathbb{K}$ -valued measure on $(\prod_{i=1}^m \bar{X}_i, \bar{\mathcal{B}})$, $\bar{\lambda} = \lambda|_{\bar{\mathcal{B}}} = \lambda|_{\bar{\mathcal{B}}}$, and $\mathcal{P} \circ \bar{\lambda} = (\mathcal{P} \circ \lambda)|_{\bar{\mathcal{B}}}$. By Proposition 12.22, λ is the unique $\mathbb{K}$ -valued measure on $(\prod_{i=1}^m X_i, \mathcal{B})$ such that $\lambda(\prod_{i=1}^m E_i) = \prod_{i=1}^m \mu_i(E_i)$ and $\mathcal{P} \circ \lambda(\prod_{i=1}^m E_i) = \prod_{i=1}^m \mathcal{P} \circ \mu_i(E_i)$, $\forall E_i \in \text{dom}(\mu_i)$, $\forall i \in \{1, \dots, m\}$. Then, $\forall i \in \{1, \dots, m\}$, $\forall E_i \in \text{dom}(\bar{\mu}_i)$, by Proposition 11.115, $\bar{\mu}_i = \mu_i|_{\bar{\mathcal{B}}_i}$ and $\mathcal{P} \circ \bar{\mu}_i = (\mathcal{P} \circ \mu_i)|_{\bar{\mathcal{B}}_i}$. Hence, $E_i \in \text{dom}(\mu_i)$. This implies that $\bar{\lambda}(\prod_{i=1}^m E_i) = \lambda(\prod_{i=1}^m E_i) = \prod_{i=1}^m \mu_i(E_i) = \prod_{i=1}^m \bar{\mu}_i(E_i)$ and $\mathcal{P} \circ \bar{\lambda}(\prod_{i=1}^m E_i) = \mathcal{P} \circ \lambda(\prod_{i=1}^m E_i) = \prod_{i=1}^m \mathcal{P} \circ \mu_i(E_i) = \prod_{i=1}^m \mathcal{P} \circ \bar{\mu}_i(E_i)$. By Proposition 12.22, $\tilde{\lambda}$ is the unique $\mathbb{K}$ -valued measure on $(\prod_{i=1}^m \bar{X}_i, \bar{\mathcal{B}})$ such that $\tilde{\lambda}(\prod_{i=1}^m E_i) = \prod_{i=1}^m \bar{\mu}_i(E_i)$ and $\mathcal{P} \circ \tilde{\lambda}(\prod_{i=1}^m E_i) = \prod_{i=1}^m \mathcal{P} \circ \bar{\mu}_i(E_i)$, $\forall E_i \in \text{dom}(\bar{\mu}_i)$, $\forall i \in \{1, \dots, m\}$. Hence, we have $\bar{\lambda} = \tilde{\lambda}$.

Hence, $\bar{\mathcal{X}} = \tilde{\mathcal{X}}$. This completes the proof of the proposition. $\square$

Proposition 12.24 *Let Γ be a finite nonempty index set, Λ_β be a finite nonempty index set, $\forall \beta \in \Gamma$, with Λ_β 's being pairwise disjoint, $\Lambda := \bigcup_{\beta \in \Gamma} \Lambda_\beta$, $\mathcal{X}_\alpha := (X_\alpha, \mathcal{B}_\alpha, \mu_\alpha)$ be a σ -finite $\mathbb{K}$ -valued measure space, $\forall \alpha \in \Lambda$, $\hat{\mathcal{X}}_\beta := \prod_{\alpha \in \Lambda_\beta} X_\alpha := (\prod_{\alpha \in \Lambda_\beta} X_\alpha, \hat{\mathcal{B}}_\beta, \hat{\mu}_\beta)$ be the σ -finite product $\mathbb{K}$ -valued measure space, $\forall \beta \in \Gamma$, $\hat{\mathcal{X}} := \prod_{\beta \in \Gamma} \hat{\mathcal{X}}_\beta := (\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, \hat{\mathcal{B}}, \hat{\mu})$ be the σ -finite product $\mathbb{K}$ -valued measure space, and $\bar{\mathcal{X}} := \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha := (\prod_{\alpha \in \Lambda} X_\alpha, \bar{\mathcal{B}}, \bar{\mu})$ be the σ -finite product $\mathbb{K}$ -valued measure space. Then, $\bar{\mathcal{X}}$ and $\hat{\mathcal{X}}$ are isomeasuric.*

Proof Define $\Psi : \hat{\mathcal{X}} \rightarrow \bar{\mathcal{X}}$ by, $\forall x \in \hat{\mathcal{X}}$, $\forall \alpha \in \Lambda$, $\exists ! \beta_\alpha \in \Gamma \ni \alpha \in \Lambda_{\beta_\alpha}$, $\pi_\alpha(\Psi(x)) = \pi_\alpha^{\langle \beta_\alpha \rangle}(\pi_{\beta_\alpha}^{\langle \Gamma \rangle}(x))$. By the proof of Proposition 3.30, Ψ is bijective with inverse $\Psi_{\text{inv}} : \bar{\mathcal{X}} \rightarrow \hat{\mathcal{X}}$.

Let $\hat{\mathcal{C}}_\beta := \{ \prod_{\alpha \in \Lambda_\beta} E_\alpha \subseteq \prod_{\alpha \in \Lambda_\beta} X_\alpha \mid E_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda_\beta \}$, $\forall \beta \in \Gamma$, $\hat{\mathcal{C}} := \left\{ \prod_{\beta \in \Gamma} E_\beta \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \mid E_\beta \in \hat{\mathcal{B}}_\beta, \forall \beta \in \Gamma \right\}$, $\bar{\mathcal{C}} := \{ \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \mid E_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda \}$, and $\bar{\mathcal{C}} := \{ \prod_{\alpha \in \Lambda} E_\alpha \subseteq \prod_{\alpha \in \Lambda} X_\alpha \mid E_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda \}$. By Proposition 12.22, $\hat{\mathcal{B}}_\beta$ is the σ -algebra generated by $\hat{\mathcal{C}}_\beta$ on $\prod_{\alpha \in \Lambda_\beta} X_\alpha$, $\forall \beta \in \Gamma$, $\hat{\mathcal{B}}$ is the σ -algebra generated by $\hat{\mathcal{C}}$ on $\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$, and $\bar{\mathcal{B}}$ is the σ -algebra generated by $\bar{\mathcal{C}}$ on $\prod_{\alpha \in \Lambda} X_\alpha$. Clearly, $\bar{\mathcal{C}} \subseteq \hat{\mathcal{C}}$.

Fix any $\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha \in \bar{\mathcal{C}}$, we will show that $\Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha) = \prod_{\alpha \in \Lambda} E_\alpha$. $\forall x \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha$, $\forall \alpha \in \Lambda$, we have $\pi_\alpha(\Psi(x)) =$

$\pi_\alpha^{(\beta_\alpha)}(\pi_{\beta_\alpha}^{(\Gamma)}(x)) \in E_\alpha$. Hence, $\Psi(x) \in \prod_{\alpha \in \Lambda} E_\alpha$. This implies that $\Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha) \subseteq \prod_{\alpha \in \Lambda} E_\alpha$. On the other hand, $\forall y \in \prod_{\alpha \in \Lambda} E_\alpha, \forall \beta \in \Gamma, \forall \alpha \in \Lambda_\beta$, we have $\pi_\alpha^{(\beta)}(\pi_\beta^{(\Gamma)}(\Psi_{\text{inv}}(y))) = \pi_\alpha(y) \in E_\alpha$. Then, $\Psi_{\text{inv}}(y) \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha$. This leads to $\Psi_{\text{inv}}(\prod_{\alpha \in \Lambda} E_\alpha) \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha$. This is equivalent to $\prod_{\alpha \in \Lambda} E_\alpha \subseteq \Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha)$. Therefore, $\Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha) = \prod_{\alpha \in \Lambda} E_\alpha \in \bar{\mathcal{C}}$. Conversely, $\forall \prod_{\alpha \in \Lambda} E_\alpha \in \bar{\mathcal{C}}$, we have $\Psi_{\text{inv}}(\prod_{\alpha \in \Lambda} E_\alpha) = \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha \in \bar{\mathcal{C}}$.

Let $\tilde{\mathcal{B}}$ be the σ -algebra on $\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$ generated by $\bar{\mathcal{C}}$. We will show that $\tilde{\mathcal{B}} = \{E \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \mid \Psi(E) \in \bar{\mathcal{B}}\} =: \hat{\mathcal{B}}$. Since $\bar{\mathcal{C}} \subseteq \tilde{\mathcal{B}}$, by the previous paragraph, we have $\bar{\mathcal{C}} \subseteq \hat{\mathcal{B}}$. It is easy to show that $\hat{\mathcal{B}}$ is a σ -algebra on $\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$ since $\bar{\mathcal{B}}$ is a σ -algebra. Then, $\tilde{\mathcal{B}} \subseteq \hat{\mathcal{B}}$. On the other hand, let $\tilde{\mathcal{B}} := \{E \subseteq \prod_{\alpha \in \Lambda} X_\alpha \mid \Psi_{\text{inv}}(E) \in \bar{\mathcal{B}}\}$. Since $\tilde{\mathcal{B}}$ is a σ -algebra on $\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$, then $\tilde{\mathcal{B}}$ is a σ -algebra on $\prod_{\alpha \in \Lambda} X_\alpha$. Since $\bar{\mathcal{C}} \subseteq \tilde{\mathcal{B}}$, by the previous paragraph, then $\bar{\mathcal{C}} \subseteq \tilde{\mathcal{B}}$. By the definition of $\bar{\mathcal{B}}$, we have $\tilde{\mathcal{B}} \subseteq \hat{\mathcal{B}}$. Since $\tilde{\mathcal{B}} \subseteq \hat{\mathcal{B}}$, then, by the definitions of $\tilde{\mathcal{B}}$ and $\hat{\mathcal{B}}$, $\tilde{\mathcal{B}} \subseteq \tilde{\mathcal{B}}$. Thus, we have $\tilde{\mathcal{B}} = \bar{\mathcal{B}}$ and $\hat{\mathcal{B}} = \tilde{\mathcal{B}}$. We can now prove the following result.

Claim 12.24.1 $\hat{\mathcal{C}} \subseteq \tilde{\mathcal{B}}$.

Proof of claim: Without loss of generality, assume that $\Gamma = \{\beta_1, \dots, \beta_n\}$ for some $n \in \mathbb{N}$. We will prove the following statement:

- $\forall i \in \{1, \dots, n+1\}, \forall E_{\beta_j} := \prod_{\alpha \in \Lambda_{\beta_j}} E_\alpha \in \hat{\mathcal{C}}_{\beta_j}, j = i, \dots, n, \forall E_{\beta_j} \in \hat{\mathcal{B}}_{\beta_j}, j = 1, \dots, i-1$, we have $\prod_{j=1}^n E_{\beta_j} \in \tilde{\mathcal{B}}$.

by mathematical induction on i .

1° $i = 1$. $\forall E_{\beta_j} := \prod_{\alpha \in \Lambda_{\beta_j}} E_\alpha \in \hat{\mathcal{C}}_{\beta_j}, j = 1, \dots, n$, we have $\prod_{j=1}^n E_{\beta_j} = \prod_{j=1}^n \prod_{\alpha \in \Lambda_{\beta_j}} E_\alpha \in \bar{\mathcal{C}} \subseteq \tilde{\mathcal{B}}$. This case is proved.

2° Assume that the statement holds for $i = k \in \{1, \dots, n\}$.

3° Consider the case $i = k+1$. Fix any $E_{\beta_j} := \prod_{\alpha \in \Lambda_{\beta_j}} E_\alpha \in \hat{\mathcal{C}}_{\beta_j}, j = k+1, \dots, n$, and any $E_{\beta_j} \in \hat{\mathcal{B}}_{\beta_j}, j = 1, \dots, k-1$. Define $\mathcal{F} := \{E_{\beta_k} \in \hat{\mathcal{B}}_{\beta_k} \mid \prod_{j=1}^n E_{\beta_j} \in \tilde{\mathcal{B}}\}$. By the inductive assumption, we have $\hat{\mathcal{C}}_{\beta_k} \subseteq \mathcal{F}$. Then, $\emptyset = \prod_{\alpha \in \Lambda_{\beta_k}} \emptyset \in \hat{\mathcal{C}}_{\beta_k} \subseteq \mathcal{F}$ and $X_{\beta_k} := \prod_{\alpha \in \Lambda_{\beta_k}} X_\alpha \in \hat{\mathcal{C}}_{\beta_k} \subseteq \mathcal{F}$. $\forall E_{\beta_k} \in \mathcal{F}$, we have $E_{\beta_k} \in \hat{\mathcal{B}}_{\beta_k}$ and $\prod_{j=1}^n E_{\beta_j} \in \tilde{\mathcal{B}}$. By $X_{\beta_k} \in \hat{\mathcal{C}}_{\beta_k} \subseteq \mathcal{F}$, we have $(\prod_{j=1}^{k-1} E_{\beta_j}) \times X_{\beta_k} \times (\prod_{j=k+1}^n E_{\beta_j}) \in \tilde{\mathcal{B}}$. This implies that $((\prod_{j=1}^{k-1} E_{\beta_j}) \times X_{\beta_k} \times (\prod_{j=k+1}^n E_{\beta_j})) \setminus (\prod_{j=1}^n E_{\beta_j}) = (\prod_{j=1}^{k-1} E_{\beta_j}) \times (X_{\beta_k} \setminus E_{\beta_k}) \times (\prod_{j=k+1}^n E_{\beta_j}) \in \tilde{\mathcal{B}}$. Then, $X_{\beta_k} \setminus E_{\beta_k} =$

$(\prod_{\alpha \in \Lambda_{\beta_k}} X_\alpha) \setminus E_{\beta_k} \in \mathcal{F}$. $\forall (E_{\beta_k, i})_{i=1}^\infty \subseteq \mathcal{F}$, we have $E_{\beta_k, i} \in \hat{\mathcal{B}}_{\beta_k}$ and $(\prod_{j=1}^{k-1} E_{\beta_j}) \times E_{\beta_k, i} \times (\prod_{j=k+1}^n E_{\beta_j}) \in \tilde{\mathcal{B}}$, $\forall i \in \mathbb{N}$. Then, $\bigcup_{i=1}^\infty E_{\beta_k, i} \in \hat{\mathcal{B}}_{\beta_k}$ and $\bigcup_{i=1}^\infty \left((\prod_{j=1}^{k-1} E_{\beta_j}) \times E_{\beta_k, i} \times (\prod_{j=k+1}^n E_{\beta_j}) \right) = (\prod_{j=1}^{k-1} E_{\beta_j}) \times (\bigcup_{i=1}^\infty E_{\beta_k, i}) \times (\prod_{j=k+1}^n E_{\beta_j}) \in \tilde{\mathcal{B}}$. This leads to $\bigcup_{i=1}^\infty E_{\beta_k, i} \in \mathcal{F}$. Hence, $\mathcal{F}$ is a σ -algebra on $\prod_{\alpha \in \Lambda_{\beta_k}} X_\alpha$. Since $\hat{\mathcal{B}}_{\beta_k} \supseteq \mathcal{F} \supseteq \hat{\mathcal{C}}_{\beta_k}$, then $\mathcal{F} = \hat{\mathcal{B}}_{\beta_k}$. Therefore, the statement holds for $i = k + 1$.

This completes the induction process and the statement holds for $i = 1, \dots, n + 1$.

By the above statement for $i = n + 1$ case, we have $\forall E_{\beta_j} \in \hat{\mathcal{B}}_{\beta_j}$, $j = 1, \dots, n$, $\prod_{j=1}^n E_{\beta_j} \in \tilde{\mathcal{B}}$. Hence, $\hat{\mathcal{C}} \subseteq \tilde{\mathcal{B}}$. This completes the proof of the claim. $\square$

By Claim 12.24.1 and the definition of $\hat{\mathcal{B}}$, we have $\hat{\mathcal{B}} \subseteq \tilde{\mathcal{B}}$. Since $\tilde{\mathcal{C}} \subseteq \hat{\mathcal{C}}$, we have $\tilde{\mathcal{B}} \subseteq \hat{\mathcal{B}}$. Therefore, $\hat{\mathcal{B}} = \tilde{\mathcal{B}}$. This further implies that, $\forall E \in \tilde{\mathcal{B}}$, $\Psi_{\text{inv}}(E) \in \tilde{\mathcal{B}} = \hat{\mathcal{B}} = \tilde{\mathcal{B}}$ and, $\forall E \in \tilde{\mathcal{B}} = \hat{\mathcal{B}}$, $\Psi(E) \in \tilde{\mathcal{B}}$.

By Proposition 12.10, Ψ is an isomeasure between $\hat{\mathcal{X}}$ and $(\prod_{\alpha \in \Lambda} X_\alpha, \tilde{\mathcal{B}}, \Psi_* \hat{\mu})$. $\forall \alpha \in \Lambda$, $\forall E_\alpha \in \text{dom}(\mu_\alpha)$, we have

$$\begin{aligned} \mathcal{P} \circ (\Psi_* \hat{\mu}) \left(\prod_{\alpha \in \Lambda} E_\alpha \right) &= \Psi_* (\mathcal{P} \circ \hat{\mu}) \left(\prod_{\alpha \in \Lambda} E_\alpha \right) = \mathcal{P} \circ \hat{\mu} \left(\Psi_{\text{inv}} \left(\prod_{\alpha \in \Lambda} E_\alpha \right) \right) \\ &= \mathcal{P} \circ \hat{\mu} \left(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha \right) = \prod_{\beta \in \Gamma} \mathcal{P} \circ \hat{\mu}_\beta \left(\prod_{\alpha \in \Lambda_\beta} E_\alpha \right) \\ &= \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} \mathcal{P} \circ \mu_\alpha(E_\alpha) = \prod_{\alpha \in \Lambda} \mathcal{P} \circ \mu_\alpha(E_\alpha) \in [0, \infty) \subset \mathbb{R} \end{aligned}$$

and

$$\begin{aligned} \Psi_* \hat{\mu} \left(\prod_{\alpha \in \Lambda} E_\alpha \right) &= \hat{\mu} \left(\Psi_{\text{inv}} \left(\prod_{\alpha \in \Lambda} E_\alpha \right) \right) = \hat{\mu} \left(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha \right) \\ &= \prod_{\beta \in \Gamma} \hat{\mu}_\beta \left(\prod_{\alpha \in \Lambda_\beta} E_\alpha \right) = \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} \mu_\alpha(E_\alpha) = \prod_{\alpha \in \Lambda} \mu_\alpha(E_\alpha) \in \mathbb{K} \end{aligned}$$

By Proposition 12.22, $\bar{\mu}$ is the unique $\mathbb{K}$ -valued measure on $(\prod_{\alpha \in \Lambda} X_\alpha, \tilde{\mathcal{B}})$ such that $\bar{\mu}(\prod_{\alpha \in \Lambda} E_\alpha) = \prod_{\alpha \in \Lambda} \mu_\alpha(E_\alpha)$ and $\mathcal{P} \circ \bar{\mu}(\prod_{\alpha \in \Lambda} E_\alpha) = \prod_{\alpha \in \Lambda} \mathcal{P} \circ \mu_\alpha(E_\alpha)$, $\forall E_\alpha \in \text{dom}(\mu_\alpha)$, $\forall \alpha \in \Lambda$. Therefore, $\bar{\mu} = \Psi_* \hat{\mu}$. This implies that Ψ is an isomeasure between $\hat{\mathcal{X}}$ and $\tilde{\mathcal{X}}$. This completes the proof of the proposition. $\square$

Proposition 12.25 *Let $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a σ -finite measure space, $i = 1, 2$, and $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := \mathcal{X}_1 \times \mathcal{X}_2$ be the product measure space. Then, $\forall E \in \mathcal{B}$, $\forall x_1 \in X_1$, $E_{x_1} := \{\bar{x}_2 \in X_2 \mid (x_1, \bar{x}_2) \in E\} \in \mathcal{B}_2$; and, $\forall x_2 \in X_2$, $E_{\langle x_2 \rangle} := \{\bar{x}_1 \in X_1 \mid (\bar{x}_1, x_2) \in E\} \in \mathcal{B}_1$.*

Proof Let $\mathcal{E} := \{E \in \mathcal{B} \mid E_{x_1} \in \mathcal{B}_2, \forall x_1 \in X_1\} \subseteq \mathcal{B}$. $\forall E \in \mathcal{C} := \{E_1 \times E_2 \subseteq X_1 \times X_2 \mid E_1 \in \mathcal{B}_1, E_2 \in \mathcal{B}_2\}$, we have $E \in \mathcal{E}$. Then, $\mathcal{C} \subseteq \mathcal{E}$. We will show that $\mathcal{E}$ is a σ -algebra on $X_1 \times X_2$. Then, by Proposition 12.22, we have $\mathcal{B} \subseteq \mathcal{E}$ and $\mathcal{B} = \mathcal{E}$.

Clearly, $\emptyset \in \mathcal{C} \subseteq \mathcal{E}$ and $X_1 \times X_2 \in \mathcal{C} \subseteq \mathcal{E}$. $\forall E \in \mathcal{E}, \forall x_1 \in X_1$, we have $((X_1 \times X_2) \setminus E)_{x_1} = X_2 \setminus E_{x_1} \in \mathcal{B}_2$ since $\mathcal{B}_2$ is a σ -algebra. By the arbitrariness of x_1 , we have $(X_1 \times X_2) \setminus E \in \mathcal{E}$. $\forall (E_i)_{i=1}^\infty \subseteq \mathcal{E}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$. $\forall x_1 \in X_1$, we have $E_{x_1} = \bigcup_{i=1}^\infty E_{i x_1} \in \mathcal{B}_2$ since $\mathcal{B}_2$ is a σ -algebra. Then, $E \in \mathcal{E}$. Therefore, $\mathcal{E}$ is a σ -algebra and $\mathcal{B} = \mathcal{E}$. By the definition of $\mathcal{E}$, we have $\forall E \in \mathcal{B} = \mathcal{E}, \forall x_1 \in X_1, E_{x_1} \in \mathcal{B}_2$. By symmetry, $\forall E \in \mathcal{B}, \forall x_2 \in X_2, E_{\langle x_2 \rangle} \in \mathcal{B}_1$. This completes the proof of the proposition. $\square$

Proposition 12.26 Let $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a σ -finite measure space, $i = 1, 2$, and $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := \mathcal{X}_1 \times \mathcal{X}_2$ be the product measure space. Then, $\forall E \in \mathcal{B}$, the function $g_E : X_1 \rightarrow [0, \infty] \subseteq \mathbb{R}_e$, defined by $g_E(x_1) = \mu_2(E_{x_1})$, $\forall x_1 \in X_1$, is $\mathcal{B}_1$ -measurable, where E_{x_1} is as defined in Proposition 12.25. By symmetry, $\forall E \in \mathcal{B}$, the function $h_E : X_2 \rightarrow [0, \infty] \subseteq \mathbb{R}_e$, defined by $h_E(x_2) = \mu_1(E_{\langle x_2 \rangle})$, $\forall x_2 \in X_2$, is $\mathcal{B}_2$ -measurable, where $E_{\langle x_2 \rangle}$ is as defined in Proposition 12.25.

Proof By Proposition 12.25, $\forall x_1 \in X_1, E_{x_1} \in \mathcal{B}_2$. Then, $g_E(x_1) \in [0, \infty] \subset \mathbb{R}_e$. Hence, the function g_E is well-defined.

First, consider the special case when $\mathcal{X}_1$ and $\mathcal{X}_2$ are finite. Let $\mathcal{E} := \{E \in \mathcal{B} \mid g_E \text{ is } \mathcal{B}_1\text{-measurable}\}$, and $\mathcal{C}$ be the set of measurable rectangles in $X_1 \times X_2$. $\forall E := E_1 \times E_2 \in \mathcal{C}$, $g_E(x_1) = \mu_2(E_2)\chi_{E_1, X_1}(x_1)$, $\forall x_1 \in X_1$. Clearly, g_E is $\mathcal{B}_1$ -measurable. Then, $E \in \mathcal{E}$ and $\mathcal{C} \subseteq \mathcal{E}$.

Clearly, $\mathcal{C}$ is a π -system on $X_1 \times X_2$. We will show that $\mathcal{E}$ is a monotone class on $X_1 \times X_2$. Then, by Monotone Class Lemma 12.19, we have $\mathcal{E} = \mathcal{B}$.

Clearly, $\emptyset = \emptyset \times \emptyset \in \mathcal{C} \subseteq \mathcal{E}$ and $X_1 \times X_2 \in \mathcal{C} \subseteq \mathcal{E}$. $\forall E_1, E_2 \in \mathcal{E}$ with $E_1 \subseteq E_2$, then $g_{E_i} : X_1 \rightarrow [0, \mu_2(X_2)] \subset \mathbb{R}$ is $\mathcal{B}_1$ -measurable, $i = 1, 2$. Let $E := E_2 \setminus E_1$. $\forall x_1 \in X_1$, we have $g_E(x_1) = \mu_2(E_{2 x_1} \setminus E_{1 x_1}) = \mu_2(E_{2 x_1}) - \mu_2(E_{1 x_1}) = g_{E_2}(x_1) - g_{E_1}(x_1) \in [0, \mu_2(X_2)] \subset \mathbb{R}$. By Propositions 7.23, 11.39, and 11.38, g_E is $\mathcal{B}_1$ -measurable. Then, $E = E_2 \setminus E_1 \in \mathcal{E}$. $\forall (E_i)_{i=1}^\infty \subseteq \mathcal{E}$ with $E_i \subseteq E_{i+1}, \forall i \in \mathbb{N}$, then $g_{E_i} : X_1 \rightarrow [0, \mu_2(X_2)] \subset \mathbb{R}$ is $\mathcal{B}_1$ -measurable, $\forall i \in \mathbb{N}$. Let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$. $\forall x_1 \in X_1$, $g_E(x_1) = \mu_2(\bigcup_{i=1}^\infty E_{i x_1}) = \lim_{i \in \mathbb{N}} \mu_2(E_{i x_1}) = \lim_{i \in \mathbb{N}} g_{E_i}(x_1)$, where the first equality follows from the definition of g_E ; the second equality follows from Proposition 11.7; and the third equality follows from the definition of g_{E_i} 's. Hence, by Proposition 11.48, g_E is $\mathcal{B}_1$ -measurable. Then, $E = \bigcup_{i=1}^\infty E_i \in \mathcal{E}$. This shows that $\mathcal{E}$ is a monotone class on $X_1 \times X_2$.

Then, $\mathcal{E} = \mathcal{B}$ and, $\forall E \in \mathcal{B}, g_E$ is $\mathcal{B}_1$ -measurable. By symmetry, $\forall E \in \mathcal{B}, h_E$ is $\mathcal{B}_2$ -measurable. The proposition holds for the special case.

Next, consider the general case when $\mathcal{X}_1$ and $\mathcal{X}_2$ are σ -finite. Since $\mathcal{X}_1$ is σ -finite, then $\exists (X_{1,n})_{n=1}^\infty \subseteq \mathcal{B}_1$ such that $X_1 = \bigcup_{n=1}^\infty X_{1,n}$ and

$\mu_1(X_{1,n}) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $X_{1,n} \subseteq X_{1,n+1}$, $\forall n \in \mathbb{N}$. Since $\mathcal{X}_2$ is σ -finite, then $\exists (X_{2,m})_{m=1}^\infty \subseteq \mathcal{B}_2$ such that $X_2 = \bigcup_{m=1}^\infty X_{2,m}$ and $\mu_2(X_{2,m}) < \infty$, $\forall m \in \mathbb{N}$. Without loss of generality, we may assume that $X_{2,m} \subseteq X_{2,m+1}$, $\forall m \in \mathbb{N}$. $\forall n, m \in \mathbb{N}$, let $\mathcal{X}_{1,n} := (X_{1,n}, \mathcal{B}_{1,n}, \mu_{1,n})$ be the finite measure subspace of $\mathcal{X}_1$, $\mathcal{X}_{2,m} := (X_{2,m}, \mathcal{B}_{2,m}, \mu_{2,m})$ be the finite measure subspace of $\mathcal{X}_2$, $(X_{1,n} \times X_{2,m}, \mathcal{B}_{n,m}, \lambda_{n,m}) := \mathcal{X}_{1,n} \times \mathcal{X}_{2,m}$, which is the finite measure subspace of $\mathcal{X}$ by Proposition 12.22 and its proof. $\forall E \in \mathcal{B}$, $\forall n, m \in \mathbb{N}$, let $E_{n,m} := E \cap (X_{1,n} \times X_{2,m}) \in \mathcal{B}_{n,m}$. By the special case, the function $g_{n,m} : X_{1,n} \rightarrow [0, \infty) \subset \mathbb{R}$, defined by $g_{n,m}(x_1) = \mu_{2,m}((E_{n,m})_{x_1} \cap X_{2,m})$, $\forall x_1 \in X_{1,n}$, is $\mathcal{B}_{1,n}$ -measurable. Define $\bar{g}_{n,m} : X_1 \rightarrow [0, \infty) \subset \mathbb{R}$ by $\bar{g}_{n,m}(x_1) = \mu_2((E_{n,m})_{x_1})$, $\forall x_1 \in X_1$. Then, $\bar{g}_{n,m}(x_1) = \begin{cases} g_{n,m}(x_1) & \forall x_1 \in X_{1,n} \\ 0 & \forall x_1 \in X_1 \setminus X_{1,n} \end{cases}$, $\forall x_1 \in X_1$. By Proposition 11.41, $\bar{g}_{n,m}$ is $\mathcal{B}_1$ -measurable. Note that $g_E(x_1) = \mu_2(E_{x_1}) = \mu_2(\bigcup_{n=1}^\infty (E_{n,n})_{x_1}) = \lim_{n \in \mathbb{N}} \mu_2((E_{n,n})_{x_1}) = \lim_{n \in \mathbb{N}} \bar{g}_{n,n}(x_1)$, $\forall x_1 \in X_1$, where the second equality follows from the fact that $E = \bigcup_{n=1}^\infty E_{n,n}$; and the third equality follows from Proposition 11.7. By Proposition 11.48 and the fact that $[0, \infty] \subset \mathbb{R}_e$ is metrizable, g_E is $\mathcal{B}_1$ -measurable. By symmetry, $\forall E \in \mathcal{B}$, h_E is $\mathcal{B}_2$ -measurable.

This completes the proof of the proposition. $\square$

Lemma 12.27 *Let $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a σ -finite measure space, $i = 1, 2$, and $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := \mathcal{X}_1 \times \mathcal{X}_2$ be the product measure space. Then, $\forall E \in \mathcal{B}$, we have $\mu(E) = \int_{X_1} g_E d\mu_1 = \int_{X_2} h_E d\mu_2$, where the function $g_E : X_1 \rightarrow [0, \infty] \subseteq \mathbb{R}_e$, as defined in Proposition 12.26, is $\mathcal{B}_1$ -measurable; and the function $h_E : X_2 \rightarrow [0, \infty] \subseteq \mathbb{R}_e$, as defined in Proposition 12.26, is $\mathcal{B}_2$ -measurable.*

Proof First, consider the special case when $\mathcal{X}_1$ and $\mathcal{X}_2$ are finite. Then $\mathcal{X}$ is finite. $\forall E \in \mathcal{B}$, $0 \leq \mu(E) \leq \mu(X_1 \times X_2) = \mu_1(X_1)\mu_2(X_2) < \infty$. By Proposition 12.26, the function $g_E : X_1 \rightarrow [0, \mu_2(X_2)] \subset \mathbb{R}$ is $\mathcal{B}_1$ -measurable. Let $\mathcal{E} := \{E \in \mathcal{B} \mid \mu(E) = \int_{X_1} g_E d\mu_1\} \subseteq \mathcal{B}$. Let $\mathcal{C}$ be the set of measurable rectangles in $X_1 \times X_2$, which is clearly a π -system. $\forall E := E_1 \times E_2 \in \mathcal{C}$, $g_E = \mu_2(E_2)\chi_{E_1, X_1}$. Then, by Proposition 11.75, $\mu(E) = \mu_1(E_1)\mu_2(E_2) = \int_{X_1} g_E d\mu_1$. This shows that $E \in \mathcal{E}$. Then $\mathcal{C} \subseteq \mathcal{E}$. We will show that $\mathcal{E}$ is a monotone class on $X_1 \times X_2$. By Monotone Class Lemma 12.19, we have $\mathcal{E} = \mathcal{B}$.

Clearly, $\emptyset \in \mathcal{C} \subseteq \mathcal{E}$ and $X_1 \times X_2 \in \mathcal{C} \subseteq \mathcal{E}$. $\forall E_1, E_2 \in \mathcal{E}$ with $E_1 \subseteq E_2$, let $E := E_2 \setminus E_1 \in \mathcal{B}$. By Proposition 12.26, g_{E_1} , g_{E_2} and g_E are $\mathcal{B}_1$ -measurable. As shown in the proof of Proposition 12.26, $g_E = g_{E_2} - g_{E_1}$. Then, $\mu(E) = \mu(E_2) - \mu(E_1) = \int_{X_1} g_{E_2} d\mu_1 - \int_{X_1} g_{E_1} d\mu_1 = \int_{X_1} g_E d\mu_1$, where the first equality follows from the fact that μ is a measure; the second equality follows from the definition of $\mathcal{E}$; the last equality follows from Proposition 11.83; and all equality follows from the fact all quantity involved in the equalities are finite. This shows that $E = E_2 \setminus E_1 \in \mathcal{E}$.

$\forall (E_i)_{i=1}^\infty \subseteq \mathcal{E}$ with $E_i \subseteq E_{i+1}$, $\forall i \in \mathbb{N}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$. $\forall i \in \mathbb{N}$, by Proposition 12.26, g_{E_i} is $\mathcal{B}_1$ -measurable. $\forall x_1 \in X_1$, we have $g_{E_i}(x_1) \leq g_{E_{i+1}}(x_1)$ since $E_i \subseteq E_{i+1}$. By Proposition 12.26, g_E is $\mathcal{B}_1$ -measurable. As shown in the proof of Proposition 12.26, $g_E(x_1) = \lim_{i \in \mathbb{N}} g_{E_i}(x_1)$, $\forall x_1 \in X_1$. This implies that $\mu(E) = \lim_{i \in \mathbb{N}} \mu(E_i) = \lim_{i \in \mathbb{N}} \int_{X_1} g_{E_i} d\mu_1 = \int_{X_1} g_E d\mu_1$, where the first equality follows from Proposition 11.7; the second equality follows from the definition of $\mathcal{E}$; and the last equality follows from Monotone Convergence Theorem 11.81. Hence, $E = \bigcup_{i=1}^\infty E_i \in \mathcal{E}$.

Hence, $\mathcal{E}$ is a monotone class on $X_1 \times X_2$ and $\mathcal{E} = \mathcal{B}$. Therefore, $\mu(E) = \int_{X_1} g_E d\mu_1$, $\forall E \in \mathcal{B}$. By symmetry, $\mu(E) = \int_{X_2} h_E d\mu_2$, $\forall E \in \mathcal{B}$. This completes the proof for the special case.

Next, consider the general case when $\mathcal{X}_1$ and $\mathcal{X}_2$ are σ -finite. Since $\mathcal{X}_1$ is σ -finite, then $\exists (X_{1,n})_{n=1}^\infty \subseteq \mathcal{B}_1$ such that $X_1 = \bigcup_{n=1}^\infty X_{1,n}$ and $\mu_1(X_{1,n}) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $X_{1,n} \subseteq X_{1,n+1}$, $\forall n \in \mathbb{N}$. Since $\mathcal{X}_2$ is σ -finite, then $\exists (X_{2,m})_{m=1}^\infty \subseteq \mathcal{B}_2$ such that $X_2 = \bigcup_{m=1}^\infty X_{2,m}$ and $\mu_2(X_{2,m}) < \infty$, $\forall m \in \mathbb{N}$. Without loss of generality, we may assume that $X_{2,m} \subseteq X_{2,m+1}$, $\forall m \in \mathbb{N}$. $\forall n, m \in \mathbb{N}$, let $\mathcal{X}_{1,n} := (X_{1,n}, \mathcal{B}_{1,n}, \mu_{1,n})$ be the finite measure subspace of $\mathcal{X}_1$, $\mathcal{X}_{2,m} := (X_{2,m}, \mathcal{B}_{2,m}, \mu_{2,m})$ be the finite measure subspace of $\mathcal{X}_2$, $(X_{1,n} \times X_{2,m}, \bar{\mathcal{B}}_{n,m}, \lambda_{n,m}) := \mathcal{X}_{1,n} \times \mathcal{X}_{2,m}$, which is the finite measure subspace of $\mathcal{X}$ by Proposition 12.22. $\forall E \in \mathcal{B}$, $\forall n, m \in \mathbb{N}$, let $E_{n,m} := E \cap (X_{1,n} \times X_{2,m}) \in \bar{\mathcal{B}}_{n,m}$. By the special case, the function $g_{n,m} : X_{1,n} \rightarrow [0, \infty) \subset \mathbb{R}$, defined by $g_{n,m}(x_1) = \mu_{2,m}((E_{n,m})_{x_1} \cap X_{2,m})$, $\forall x_1 \in X_{1,n}$, is $\mathcal{B}_{1,n}$ -measurable and $\mu(E_{n,m}) = \lambda_{n,m}(E_{n,m}) = \int_{X_{1,n}} g_{n,m} d\mu_{1,n}$. Define $\bar{g}_{n,m} : X_1 \rightarrow [0, \infty) \subset \mathbb{R}$ by $\bar{g}_{n,m}(x_1) = \mu_2((E_{n,m})_{x_1})$, $\forall x_1 \in X_1$. Then, $\bar{g}_{n,m}(x_1) = \begin{cases} g_{n,m}(x_1) & \forall x_1 \in X_{1,n} \\ 0 & \forall x_1 \in X_1 \setminus X_{1,n} \end{cases}$, $\forall x_1 \in X_1$. By Proposition 11.41, $\bar{g}_{n,m}$ is $\mathcal{B}_1$ -measurable. By Definition 11.79, $\mu(E_{n,m}) = \int_{X_1} \bar{g}_{n,m} d\mu_1$. By Proposition 12.26 and its proof, we have $g_E(x_1) = \lim_{n \in \mathbb{N}} \bar{g}_{n,n}(x_1)$, $\forall x_1 \in X_1$, and g_E is $\mathcal{B}_1$ -measurable. Clearly, $\bar{g}_{n,n}(x_1) \leq \bar{g}_{n+1,n+1}(x_1)$, $\forall x_1 \in X_1$, $\forall n \in \mathbb{N}$, since $E_{n,n} \subseteq E_{n+1,n+1}$. Then, $\mu(E) = \lim_{n \in \mathbb{N}} \mu(E_{n,n}) = \lim_{n \in \mathbb{N}} \int_{X_1} \bar{g}_{n,n} d\mu_1 = \int_{X_1} g_E d\mu_1$, where the first equality follows from Proposition 11.7; and the last equality follows from Monotone Convergence Theorem 11.81.

By symmetry, $\mu(E) = \int_{X_2} h_E d\mu_2$, $\forall E \in \mathcal{B}$.

This completes the proof of the proposition. $\square$

Proposition 12.28 *Let $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a σ -finite measure space, $i = 1, 2$, $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := \mathcal{X}_1 \times \mathcal{X}_2$ be the product measure space, $\mathcal{Y} := (Y, \mathcal{O})$ be a topological space, and $f : X_1 \times X_2 \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. Then,*

- (i) $\forall x_1 \in X_1$, $f_{x_1} : X_2 \rightarrow \mathcal{Y}$, defined by $f_{x_1}(x_2) = f(x_1, x_2)$, $\forall x_2 \in X_2$, is $\mathcal{B}_2$ -measurable;
- (ii) $\forall x_2 \in X_2$, $f_{\langle x_2 \rangle} : X_1 \rightarrow \mathcal{Y}$, defined by $f_{\langle x_2 \rangle}(x_1) = f(x_1, x_2)$, $\forall x_1 \in X_1$, is $\mathcal{B}_1$ -measurable.

Proof (i) $\forall$ open set $O \in \mathcal{O}$, $f_{\text{inv}}(O) \in \mathcal{B}$ since f is $\mathcal{B}$ -measurable. $\forall x_1 \in X_1$, by Proposition 12.25, $(f_{\text{inv}}(O))_{x_1} := \{x_2 \in X_2 \mid (x_1, x_2) \in f_{\text{inv}}(O)\} \in \mathcal{B}_2$. Clearly, $f_{x_1 \text{ inv}}(O) = (f_{\text{inv}}(O))_{x_1} \in \mathcal{B}_2$. Then, f_{x_1} is $\mathcal{B}_2$ -measurable.

(ii) By symmetry, (ii) holds. This completes the proof of the proposition. $\square$

Theorem 12.29 (Tonelli) Let $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a σ -finite measure space, $i = 1, 2$, $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := \mathcal{X}_1 \times \mathcal{X}_2$ be the product measure space, and $f : X_1 \times X_2 \rightarrow [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$ -measurable. Define $f_{x_1} : X_2 \rightarrow \mathbb{Y}$ and $f_{\langle x_2 \rangle} : X_1 \rightarrow \mathbb{Y}$ as in Proposition 12.28. Then, the following statements hold.

(i) $\int_{X_1 \times X_2} f \, d\mu = \int_{X_1} p \, d\mu_1$, where $p : X_1 \rightarrow [0, \infty] \subset \mathbb{R}_e$, defined by $p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2$, $\forall x_1 \in X_1$, is $\mathcal{B}_1$ -measurable.

(ii) $\int_{X_1 \times X_2} f \, d\mu = \int_{X_2} q \, d\mu_2$, where $q : X_2 \rightarrow [0, \infty] \subset \mathbb{R}_e$, defined by $q(x_2) = \int_{X_1} f_{\langle x_2 \rangle} \, d\mu_1$, $\forall x_2 \in X_2$, is $\mathcal{B}_2$ -measurable.

Proof (i) By Proposition 11.67, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : \mathcal{X} \rightarrow [0, \infty] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, such that $0 \leq \varphi_n(x_1, x_2) \leq \varphi_{n+1}(x_1, x_2) \leq f(x_1, x_2)$, $\forall (x_1, x_2) \in X_1 \times X_2$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n(x_1, x_2) = f(x_1, x_2)$, $\forall (x_1, x_2) \in X_1 \times X_2$. By Proposition 12.28 and Definition 11.79, p is well-defined. Define $p_n : X_1 \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $p_n(x_1) = \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2)$, $\forall x_1 \in X_1$, $\forall n \in \mathbb{N}$. By Proposition 12.28, p_n 's are well-defined. By Proposition 11.83, $p_n(x_1) \leq p_{n+1}(x_1) \leq p(x_1)$, $\forall x_1 \in X_1$, $\forall n \in \mathbb{N}$. By Monotone Convergence Theorem 11.81, $p(x_1) = \lim_{n \in \mathbb{N}} p_n(x_1)$, $\forall x_1 \in X_1$. Fix any $n \in \mathbb{N}$. Let φ_n admit the canonical representation $\varphi_n = \sum_{i=1}^{m_n} a_i \chi_{E_i, X_1 \times X_2}$, where $m_n \in \mathbb{Z}_+$, $a_1, \dots, a_{m_n} \in (0, \infty) \subset \mathbb{R}$ are distinct, $E_1, \dots, E_{m_n} \in \mathcal{B}$ are nonempty, pairwise disjoint, and $\mu(E_i) < \infty$, $\forall i \in \{1, \dots, m_n\}$. Fix any $i \in \{1, \dots, m_n\}$, define $g_i : X_1 \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $g_i(x_1) = \mu_2((E_i)_{x_1})$, $\forall x_1 \in X_1$, where $(E_i)_{x_1} := \{x_2 \in X_2 \mid (x_1, x_2) \in E_i\}$. By Proposition 12.26, g_i is $\mathcal{B}_1$ -measurable. By Lemma 12.27, $\mu(E_i) = \int_{X_1} g_i \, d\mu_1$. By Proposition 11.75, $\int_{X_1 \times X_2} \varphi_n \, d\mu = \sum_{i=1}^{m_n} a_i \mu(E_i) = \sum_{i=1}^{m_n} a_i \int_{X_1} g_i \, d\mu_1 = \int_{X_1} (\sum_{i=1}^{m_n} a_i g_i) \, d\mu_1 =: \int_{X_1} \bar{p}_n \, d\mu_1$, where the third equality follows from Proposition 11.83. By Propositions 3.89, 11.38, and 11.39, $\bar{p}_n : X_1 \rightarrow [0, \infty] \subset \mathbb{R}_e$ is $\mathcal{B}_1$ -measurable. Note that $p_n(x_1) = \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) = \int_{X_2} (\sum_{i=1}^{m_n} a_i \chi_{E_i, X_1 \times X_2}(x_1, x_2)) \, d\mu_2(x_2) = \sum_{i=1}^{m_n} a_i \int_{X_2} \chi_{(E_i)_{x_1}, X_2} \, d\mu_2 = \sum_{i=1}^{m_n} a_i \mu_2(E_{i, x_1}) = \sum_{i=1}^{m_n} a_i g_i(x_1) = \bar{p}_n(x_1)$, $\forall x_1 \in X_1$, where the first equality follows from the definition of p_n ; the second equality follows from the canonical representation of φ_n ; the third equality follows from Proposition 11.83; the fourth equality follows from the fact that $\mathcal{X}_2$ is σ -finite, Propositions 11.75 and 11.7, and the Monotone Convergence Theorem 11.81; the fifth equality follows from the definition

of g_i 's; and the last equality follows from the definition of $\bar{p}_n$. Then, p_n is $\mathcal{B}_1$ -measurable. By Proposition 11.48 and the fact that $[0, \infty] \subset \mathbb{R}_e$ is metrizable, p is $\mathcal{B}_1$ -measurable. Then, $\int_{X_1 \times X_2} \varphi_n d\mu = \int_{X_1} \bar{p}_n d\mu_1 = \int_{X_1} p_n d\mu_1 \in [0, \infty] \subset \mathbb{R}_e$. By Monotone Convergence Theorem 11.81, $\int_{X_1 \times X_2} f d\mu = \lim_{n \in \mathbb{N}} \int_{X_1 \times X_2} \varphi_n d\mu = \lim_{n \in \mathbb{N}} \int_{X_1} p_n d\mu_1 = \int_{X_1} p d\mu_1$.

Hence, (i) holds. By symmetry, (ii) holds. This completes the proof of the theorem. $\square$

Theorem 12.30 (Fubini) *Let $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a σ -finite measure space, $i = 1, 2$, $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := \mathcal{X}_1 \times \mathcal{X}_2$ be the product measure space, $\mathcal{Y}$ be a separable Banach space, and $f : X_1 \times X_2 \rightarrow \mathcal{Y}$ be absolutely integrable over $\mathcal{X}$. Define $f_{x_1} : X_2 \rightarrow \mathcal{Y}$ and $f_{\langle x_2 \rangle} : X_1 \rightarrow \mathcal{Y}$ as in Proposition 12.28. Then, the following statements hold.*

- (i) *There exists $\check{p} : X_1 \rightarrow \mathcal{Y}$ and exists $\check{U} \in \mathcal{B}_1$ such that $\mu_1(X_1 \setminus \check{U}) = 0$; $\forall x_1 \in \check{U}$, f_{x_1} is absolutely integrable over $\mathcal{X}_2$ and $\check{p}(x_1) = \int_{X_2} f_{x_1} d\mu_2 \in \mathcal{Y}$; $\forall x_1 \in X_1 \setminus \check{U}$, $\check{p}(x_1) = \vartheta_{\mathcal{Y}}$; $\check{p}$ is absolutely integrable over $\mathcal{X}_1$; and $\int_{X_1 \times X_2} f d\mu = \int_{X_1} \check{p} d\mu_1 \in \mathcal{Y}$.*
- (ii) *There exists $\check{q} : X_2 \rightarrow \mathcal{Y}$ and exists $\check{V} \in \mathcal{B}_2$ such that $\mu_2(X_2 \setminus \check{V}) = 0$; $\forall x_2 \in \check{V}$, $f_{\langle x_2 \rangle}$ is absolutely integrable over $\mathcal{X}_1$ and $\check{q}(x_2) = \int_{X_1} f_{\langle x_2 \rangle} d\mu_1 \in \mathcal{Y}$; $\forall x_2 \in X_2 \setminus \check{V}$, $\check{q}(x_2) = \vartheta_{\mathcal{Y}}$; $\check{q}$ is absolutely integrable over $\mathcal{X}_2$; and $\int_{X_1 \times X_2} f d\mu = \int_{X_2} \check{q} d\mu_2 \in \mathcal{Y}$.*
- (iii) *$\forall p : X_1 \rightarrow \mathcal{Y}$ satisfying: p is $\mathcal{B}_1$ -measurable; and $\exists U \in \mathcal{B}_1$ with $\mu_1(X_1 \setminus U) = 0$ such that, $\forall x_1 \in U$, f_{x_1} is absolutely integrable over $\mathcal{X}_2$ and $p(x_1) = \int_{X_2} f_{x_1} d\mu_2 \in \mathcal{Y}$, we have p is absolutely integrable over $\mathcal{X}_1$ and $\int_{X_1 \times X_2} f d\mu = \int_{X_1} p d\mu_1 \in \mathcal{Y}$.*
- (iv) *$\forall q : X_2 \rightarrow \mathcal{Y}$ satisfying: q is $\mathcal{B}_2$ -measurable; and $\exists V \in \mathcal{B}_2$ with $\mu_2(X_2 \setminus V) = 0$ such that, $\forall x_2 \in V$, $f_{\langle x_2 \rangle}$ is absolutely integrable over $\mathcal{X}_1$ and $q(x_2) = \int_{X_1} f_{\langle x_2 \rangle} d\mu_1 \in \mathcal{Y}$, we have q is absolutely integrable over $\mathcal{X}_2$ and $\int_{X_1 \times X_2} f d\mu = \int_{X_2} q d\mu_2 \in \mathcal{Y}$.*

Proof (i) Since f is absolutely integrable over $\mathcal{X}$, then $\mathcal{P} \circ f$ is $\mathcal{B}$ -measurable and $\int_{X_1 \times X_2} \mathcal{P} \circ f d\mu < \infty$. By Tonelli Theorem 12.29, $g : X_1 \rightarrow [0, \infty] \subset \mathbb{R}_e$, defined by $g(x_1) = \int_{X_2} \mathcal{P} \circ f(x_1, x_2) d\mu_2(x_2)$, $\forall x_1 \in X_1$, is $\mathcal{B}_1$ -measurable, and $\int_{X_1 \times X_2} \mathcal{P} \circ f d\mu = \int_{X_1} g d\mu_1$. By Proposition 11.82, the function $\tilde{g} : X_1 \rightarrow [0, \infty) \subset \mathbb{R}$, defined by $\tilde{g}(x_1) = \begin{cases} g(x_1) & g(x_1) < \infty \\ 0 & g(x_1) = \infty \end{cases}$, $\forall x_1 \in X_1$, is $\mathcal{B}_1$ -measurable, $g = \tilde{g}$ a.e. in $\mathcal{X}_1$, and $\int_{X_1 \times X_2} \mathcal{P} \circ f d\mu = \int_{X_1} \tilde{g} d\mu_1 < \infty$. Let $U := \{x_1 \in X_1 \mid g(x_1) < \infty\}$. Then, $U \in \mathcal{B}_1$ and $\mu_1(X_1 \setminus U) = 0$. $\forall x_1 \in U$, we have $g(x_1) = \int_{X_2} \mathcal{P} \circ f(x_1, x_2) d\mu_2(x_2) = \int_{X_2} \mathcal{P} \circ f_{x_1} d\mu_2 < \infty$. Then, f_{x_1} is absolutely integrable over $\mathcal{X}_2$. By

Propositions 12.28 and 11.92, $\int_{X_2} f_{x_1} d\mu_2 \in \mathcal{Y}$, $\forall x_1 \in U$. Define $\tilde{p} : X_1 \rightarrow \mathcal{Y}$ by $\tilde{p}(x_1) = \begin{cases} \int_{X_2} f_{x_1} d\mu_2 & x_1 \in U \\ \vartheta_{\mathcal{Y}} & x_1 \in X_1 \setminus U \end{cases}$, $\forall x_1 \in X_1$.

By Proposition 11.66, $\exists$ a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : \mathcal{X} \rightarrow \mathcal{Y}$, $\forall n \in \mathbb{N}$, such that $\|\varphi_n(x_1, x_2)\| \leq \|f(x_1, x_2)\|$, $\forall (x_1, x_2) \in X_1 \times X_2$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$. Let $A := \{(x_1, x_2) \in X_1 \times X_2 \mid (\varphi_n(x_1, x_2))_{n=1}^\infty \text{ does not converge to } f(x_1, x_2)\}$. Then, $A \in \mathcal{B}$ and $\mu(A) = 0$. Define $g_A : X_1 \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $g_A(x_1) = \mu_2(A_{x_1})$, $\forall x_1 \in X_1$, where $A_{x_1} := \{\bar{x}_2 \in X_2 \mid (x_1, \bar{x}_2) \in A\}$ is as defined in Proposition 12.25, and $\hat{g}_A : X_1 \rightarrow [0, \infty) \subset \mathbb{R}$ by $\hat{g}_A(x_1) = \begin{cases} g_A(x_1) & g_A(x_1) < \infty \\ 0 & g_A(x_1) = \infty \end{cases}$, $\forall x_1 \in X_1$. Then, by Lemma 12.27 and Proposition 11.82, g_A and $\hat{g}_A$ are well-defined and $\mathcal{B}_1$ -measurable, $g_A = \hat{g}_A$ a.e. in $\mathcal{X}_1$, and $\mu(A) = \int_{X_1} g_A d\mu_1 = \int_{X_1} \hat{g}_A d\mu_1 = 0$. Since $\hat{g}_A(x_1) \geq 0$, $\forall x_1 \in X_1$, then, by Proposition 11.96, we have $\hat{g}_A = 0$ a.e. in $\mathcal{X}_1$. Then, by Lemma 11.44 and the fact that $[0, \infty] \subset \mathbb{R}_e$ is second countable and metrizable, $g_A = 0$ a.e. in $\mathcal{X}_1$. Let $A_1 := \{x_1 \in X_1 \mid g_A(x_1) \neq 0\}$. Then, $A_1 \in \mathcal{B}_1$ and $\mu_1(A_1) = 0$. Fix any $x_1 \in U \setminus A_1$. $g_A(x_1) = 0$ implies that $\mu_2(A_{x_1}) = 0$, which further implies that $\lim_{n \in \mathbb{N}} \varphi_n(x_1, x_2) = f(x_1, x_2)$ a.e. $x_2 \in X_2$. Then, by Lebesgue Dominated Convergence Theorem 11.91, $\tilde{p}(x_1) = \lim_{n \in \mathbb{N}} \int_{X_2} \varphi_n(x_1, x_2) d\mu_2(x_2) \in \mathcal{Y}$, $\forall x_1 \in U \setminus A_1$.

Fix any $n \in \mathbb{N}$. Let φ_n admit the canonical representation $\varphi_n = \sum_{i=1}^m y_i \chi_{E_i, X_1 \times X_2}$, where $m \in \mathbb{Z}_+$, $y_1, \dots, y_m \in \mathcal{Y}$ are distinct and none equals to $\vartheta_{\mathcal{Y}}$, $E_1, \dots, E_m \in \mathcal{B}$ are nonempty, pairwise disjoint, and $\mu(E_i) < \infty$, $i = 1, \dots, m$. Fix any $i \in \{1, \dots, m\}$. Define $g_i : X_1 \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $g_i(x_1) = \mu_2(E_{i, x_1})$, $\forall x_1 \in X_1$, where $E_{i, x_1} := \{\bar{x}_2 \in X_2 \mid (x_1, \bar{x}_2) \in E_i\}$. By Lemma 12.27, g_i is $\mathcal{B}_1$ -measurable and $\mu(E_i) = \int_{X_1} g_i d\mu_1$. Define

$$\hat{g}_i : X_1 \rightarrow [0, \infty) \subset \mathbb{R} \text{ by } \hat{g}_i(x_1) = \begin{cases} g_i(x_1) & g_i(x_1) < \infty \\ 0 & g_i(x_1) = \infty \end{cases}, \forall x_1 \in X_1.$$

By Proposition 11.82, $\hat{g}_i$ is $\mathcal{B}_1$ -measurable, $g_i = \hat{g}_i$ a.e. in $\mathcal{X}_1$, and $\mu(E_i) = \int_{X_1} \hat{g}_i d\mu_1$. Then, $\int_{X_1 \times X_2} \varphi_n d\mu = \sum_{i=1}^m y_i \mu(E_i) = \sum_{i=1}^m y_i \int_{X_1} \hat{g}_i d\mu_1 = \int_{X_1} (\sum_{i=1}^m y_i \hat{g}_i) d\mu_1 =: \int_{X_1} \bar{p}_n d\mu_1$, where the first equality follows from Proposition 11.75; and the third equality follows from Proposition 11.92. By Propositions 7.23, 11.38, and 11.39, $\bar{p}_n : X_1 \rightarrow \mathcal{Y}$ is $\mathcal{B}_1$ -measurable and, by Proposition 11.92, $\bar{p}_n$ is absolutely integrable over $\mathcal{X}_1$. Let $\bar{A}_n := \bigcup_{i=1}^m \{x_1 \in X_1 \mid g_i(x_1) = \infty\}$. Then, $\bar{A}_n \in \mathcal{B}_1$ and $\mu_1(\bar{A}_n) = 0$. $\forall x_1 \in X_1 \setminus \bar{A}_n$, $\bar{p}_n(x_1) = \sum_{i=1}^m y_i \hat{g}_i(x_1) = \sum_{i=1}^m y_i g_i(x_1) = \sum_{i=1}^m y_i \mu_2(E_{i, x_1}) = \sum_{i=1}^m y_i \int_{X_2} \chi_{E_{i, x_1}, X_2} d\mu_2 = \int_{X_2} (\sum_{i=1}^m y_i \chi_{E_{i, x_1}, X_2}) d\mu_2 = \int_{X_2} \varphi_n(x_1, x_2) d\mu_2(x_2) \in \mathcal{Y}$, where the first equality follows from the definition of $\bar{p}_n$; the second equality follows from the definitions of $\hat{g}_i$'s; the third equality follows from the definitions of g_i 's; the fourth equality follows from Proposition 11.75; and the fifth equality follows from Proposition 11.92. Let $A_2 := \bigcup_{n=1}^\infty \bar{A}_n$. Then, $A_2 \in \mathcal{B}_1$ and $\mu_1(A_2) = 0$. $\forall x_1 \in U \setminus (A_1 \cup A_2) =: \check{U}$, we have $\tilde{p}(x_1) = \lim_{n \in \mathbb{N}} \int_{X_2} \varphi_n(x_1, x_2) d\mu_2(x_2) = \lim_{n \in \mathbb{N}} \bar{p}_n(x_1) \in \mathcal{Y}$.

Clearly, $\tilde{U} \in \mathcal{B}_1$ and $0 \leq \mu_1(X_1 \setminus \tilde{U}) = \mu_1((X_1 \setminus U) \cup A_1 \cup A_2) \leq \mu_1(X_1 \setminus U) + \mu_1(A_1) + \mu_1(A_2) = 0$. Fix any $n \in \mathbb{N}$. Define $\check{p}_n : X_1 \rightarrow \mathcal{Y}$ by
$$\check{p}_n(x_1) = \begin{cases} \bar{p}_n(x_1) & x_1 \in \tilde{U} \\ \vartheta_{\mathcal{Y}} & x_1 \in X_1 \setminus \tilde{U} \end{cases}, \forall x_1 \in X_1.$$
 By Proposition 11.41, $\check{p}_n$ is $\mathcal{B}_1$ -measurable. By Lemma 11.43, $\check{p}_n = \bar{p}_n$ a.e. in $\mathcal{X}_1$. Define $\check{p} : X_1 \rightarrow \mathcal{Y}$ by
$$\check{p}(x_1) = \begin{cases} \tilde{p}(x_1) & x_1 \in \tilde{U} \\ \vartheta_{\mathcal{Y}} & x_1 \in X_1 \setminus \tilde{U} \end{cases}, \forall x_1 \in X_1.$$
 Then, $\check{p}(x_1) = \lim_{n \in \mathbb{N}} \check{p}_n(x_1)$, $\forall x_1 \in X_1$. By Proposition 11.48, $\check{p}$ is $\mathcal{B}_1$ -measurable. $\forall x_1 \in \tilde{U}$, $\check{p}(x_1) = \tilde{p}(x_1) = \int_{X_2} f_{x_1} d\mu_2$ and f_{x_1} is absolutely integrable over $\mathcal{X}_2$. Note that $\int_{X_1 \times X_2} f d\mu = \lim_{n \in \mathbb{N}} \int_{X_1 \times X_2} \varphi_n d\mu = \lim_{n \in \mathbb{N}} \int_{X_1} \bar{p}_n d\mu_1 = \lim_{n \in \mathbb{N}} \int_{X_1} \check{p}_n d\mu_1$, where the first equality follows from Lebesgue Dominated Convergence Theorem 11.91; the second equality follows from the definition of $\bar{p}_n$; and the third equality follows from Proposition 11.92. $\forall n \in \mathbb{N}$, $\forall x_1 \in \tilde{U}$, $\|\check{p}_n(x_1)\| = \|\bar{p}_n(x_1)\| = \left\| \int_{X_2} \varphi_n(x_1, x_2) d\mu_2(x_2) \right\| \leq \int_{X_2} \mathcal{P} \circ \varphi_n(x_1, x_2) d\mu_2(x_2) \leq \int_{X_2} \mathcal{P} \circ f(x_1, x_2) d\mu_2(x_2) = g(x_1) = \tilde{g}(x_1)$. $\forall x_1 \in X_1 \setminus \tilde{U}$, $\|\check{p}_n(x_1)\| = 0 \leq \tilde{g}(x_1)$. Hence, we have $\|\check{p}_n(x_1)\| \leq \tilde{g}(x_1)$, $\forall x_1 \in X_1$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91, we have $\lim_{n \in \mathbb{N}} \int_{X_1} \check{p}_n d\mu_1 = \int_{X_1} \check{p} d\mu_1$ and $\check{p}$ is absolutely integrable over $\mathcal{X}_1$. Then, $\int_{X_1 \times X_2} f d\mu = \int_{X_1} \check{p} d\mu_1$. Hence, (i) holds.

By symmetry, (ii) also holds.

(iii) Fix any $p : X_1 \rightarrow \mathcal{Y}$ satisfying: p is $\mathcal{B}_1$ -measurable; and $\exists U \in \mathcal{B}_1$ with $\mu_1(X_1 \setminus U) = 0$ such that, $\forall x_1 \in U$, f_{x_1} is absolutely integrable over $\mathcal{X}_2$ and $p(x_1) = \int_{X_2} f_{x_1} d\mu_2$. By (i), we have $U_0 := \tilde{U} \cap U \in \mathcal{B}_1$ and $\mu_1(X_1 \setminus U_0) = \mu_1((X_1 \setminus \tilde{U}) \cup (X_1 \setminus U)) = 0$. $\forall x_1 \in U_0$, $p(x_1) = \int_{X_2} f_{x_1} d\mu_2 = \check{p}(x_1)$. By Lemma 11.43, $p = \check{p}$ a.e. in $\mathcal{X}_1$. By Proposition 11.83, we have p is absolutely integrable over $\mathcal{X}_1$ since $\check{p}$ is so. By Proposition 11.92, $\int_{X_1} p d\mu_1 = \int_{X_1} \check{p} d\mu_1 = \int_{X_1 \times X_2} f d\mu \in \mathcal{Y}$. Hence, (iii) holds.

By symmetry, (iv) holds. This completes the proof of the theorem. $\square$

When the range of the function f is a σ -compact subset of a separable Banach space, then the set $\tilde{U}$ and $\tilde{V}$ of the previous theorem takes a simpler form. This is established in the following theorem.

Theorem 12.31 (Fubini) *Let $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a σ -finite measure space, $i = 1, 2$, $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := \mathcal{X}_1 \times \mathcal{X}_2$ be the product measure space, $\mathcal{Y}$ be a separable Banach space, $U \subseteq \mathcal{Y}$ be a σ -compact conic segment, $f : X_1 \times X_2 \rightarrow U$ be absolutely integrable over $\mathcal{X}$. Define $f_{x_1} : X_2 \rightarrow \mathcal{Y}$ and $f_{\langle x_2 \rangle} : X_1 \rightarrow \mathcal{Y}$ as in Proposition 12.28. Then, the following statements hold.*

- (i) Define $U := \{x_1 \in X_1 \mid f_{x_1} \text{ is absolutely integrable over } \mathcal{X}_2\}$, and
$$p : X_1 \rightarrow \mathcal{Y} \text{ by } p(x_1) = \begin{cases} \int_{X_2} f_{x_1} d\mu_2 & x_1 \in U \\ \vartheta_{\mathcal{Y}} & x_1 \in X_1 \setminus U \end{cases}, \forall x_1 \in X_1.$$
 Then, $U \in \mathcal{B}_1$ with $\mu_1(X_1 \setminus U) = 0$, p is $\mathcal{B}_1$ -measurable and is absolutely integrable over $\mathcal{X}_1$, and $\int_{X_1 \times X_2} f d\mu = \int_{X_1} p d\mu_1 \in \mathcal{Y}$.

(ii) Define $V := \{x_2 \in X_2 \mid f_{\langle x_2 \rangle} \text{ is absolutely integrable over } \mathcal{X}_1\}$, and $q : X_2 \rightarrow \mathcal{Y}$ by $q(x_2) = \begin{cases} \int_{X_1} f_{\langle x_2 \rangle} d\mu_1 & x_2 \in V \\ \vartheta_{\mathcal{Y}} & x_2 \in X_2 \setminus V \end{cases}, \forall x_2 \in X_2$. Then, $V \in \mathcal{B}_2$ with $\mu_2(X_2 \setminus V) = 0$, q is $\mathcal{B}_2$ -measurable and is absolutely integrable over $\mathcal{X}_2$, and $\int_{X_1 \times X_2} f d\mu = \int_{X_2} q d\mu_2 \in \mathcal{Y}$.

Proof (i) Since f is absolutely integrable over $\mathcal{X}$, then f and $\mathcal{P} \circ f$ are $\mathcal{B}$ -measurable and $\int_{X_1 \times X_2} \mathcal{P} \circ f d\mu < \infty$. By Tonelli Theorem 12.29, $g : X_1 \rightarrow [0, \infty] \subset \mathbb{R}_e$, defined by $g(x_1) = \int_{X_2} \mathcal{P} \circ f(x_1, x_2) d\mu_2(x_2)$, $\forall x_1 \in X_1$, is $\mathcal{B}_1$ -measurable and $\int_{X_1 \times X_2} \mathcal{P} \circ f d\mu = \int_{X_1} g d\mu_1$. By Proposition 11.82, the function $\tilde{g} : X_1 \rightarrow [0, \infty) \subset \mathbb{R}$, defined by $\tilde{g}(x_1) = \begin{cases} g(x_1) & g(x_1) < \infty \\ 0 & g(x_1) = \infty \end{cases}, \forall x_1 \in X_1$, is $\mathcal{B}_1$ -measurable, $g = \tilde{g}$ a.e. in $\mathcal{X}_1$, and $\int_{X_1 \times X_2} \mathcal{P} \circ f d\mu = \int_{X_1} \tilde{g} d\mu_1$. Let $\bar{U} := \{x_1 \in X_1 \mid g(x_1) < \infty\}$. Then, $\bar{U} \in \mathcal{B}_1$, and $\mu_1(X_1 \setminus \bar{U}) = 0$. $\forall x_1 \in \bar{U}$, we have $g(x_1) = \int_{X_2} \mathcal{P} \circ f(x_1, x_2) d\mu_2(x_2) = \int_{X_2} \mathcal{P} \circ f_{x_1} d\mu_2 < \infty$. Then, f_{x_1} is absolutely integrable over $\mathcal{X}_2$, $\forall x_1 \in \bar{U}$. By Propositions 12.28 and 11.92, $\int_{X_2} f_{x_1} d\mu_2 \in \mathcal{Y}, \forall x_1 \in \bar{U}$. $\forall x_1 \in X_1 \setminus \bar{U}$, $\int_{X_2} \mathcal{P} \circ f_{x_1} d\mu_2 = \infty$ and f_{x_1} is not absolutely integrable over $\mathcal{X}_2$. Hence, $U = \bar{U}$.

By Lemma 11.68, $\exists$ a sequence of simple functions $(\varphi_n)_{n=1}^\infty, \varphi_n : \mathcal{X} \rightarrow U, \forall n \in \mathbb{N}$, such that $\|\varphi_n(x_1, x_2)\| \leq \|f(x_1, x_2)\|, \forall (x_1, x_2) \in X_1 \times X_2, \forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n(x_1, x_2) = f(x_1, x_2), \forall (x_1, x_2) \in X_1 \times X_2$. Fix any $x_1 \in X_1$, we have $\lim_{n \in \mathbb{N}} \varphi_n(x_1, x_2) = f_{x_1}(x_2), \forall x_2 \in X_2$. Then, by Lebesgue Dominated Convergence Theorem 11.91, $p(x_1) = \lim_{n \in \mathbb{N}} \int_{X_2} \varphi_n(x_1, x_2) d\mu_2(x_2), \forall x_1 \in U$.

Fix any $n \in \mathbb{N}$. Let φ_n admit the canonical representation $\varphi_n = \sum_{i=1}^m y_i \chi_{E_i, X_1 \times X_2}$, where $m \in \mathbb{Z}_+, y_1, \dots, y_m \in U$ are distinct and none equals to $\vartheta_{\mathcal{Y}}, E_1, \dots, E_m \in \mathcal{B}$ are nonempty, pairwise disjoint, and $\mu(E_i) < \infty, i = 1, \dots, m$. Fix any $i \in \{1, \dots, m\}$. Define $g_i : X_1 \rightarrow [0, \infty] \subset \mathbb{R}_e$ by $g_i(x_1) = \mu_2(E_{i, x_1}), \forall x_1 \in X_1$, where $E_{i, x_1} := \{\bar{x}_2 \in X_2 \mid (x_1, \bar{x}_2) \in E_i\}$. By Lemma 12.27, g_i is $\mathcal{B}_1$ -measurable and $\mu(E_i) = \int_{X_1} g_i d\mu_1$. Define

$$\hat{g}_i : X_1 \rightarrow [0, \infty) \subset \mathbb{R} \text{ by } \hat{g}_i(x_1) = \begin{cases} g_i(x_1) & g_i(x_1) < \infty \\ 0 & g_i(x_1) = \infty \end{cases}, \forall x_1 \in X_1.$$

By Proposition 11.82, $\hat{g}_i$ is $\mathcal{B}_1$ -measurable, $g_i = \hat{g}_i$ a.e. in $\mathcal{X}_1$, and $\mu(E_i) = \int_{X_1} \hat{g}_i d\mu_1$. Then, $\int_{X_1 \times X_2} \varphi_n d\mu = \sum_{i=1}^m y_i \mu(E_i) = \sum_{i=1}^m y_i \int_{X_1} \hat{g}_i d\mu_1 = \int_{X_1} (\sum_{i=1}^m y_i \hat{g}_i) d\mu_1 =: \int_{X_1} \bar{p}_n d\mu_1$, where the first equality follows from Proposition 11.75; and the third equality follows from Proposition 11.92. By Propositions 7.23, 11.38, and 11.39, $\bar{p}_n : X_1 \rightarrow \mathcal{Y}$ is $\mathcal{B}_1$ -measurable and, by Proposition 11.92, $\bar{p}_n$ is absolutely integrable over $\mathcal{X}_1$. $\forall x_1 \in U$, $g(x_1) < \infty$, then f_{x_1} is absolutely integrable over $\mathcal{X}_2$. Since $\|\varphi_n(x_1, x_2)\| \leq \|f(x_1, x_2)\|, \forall (x_1, x_2) \in \mathcal{X}, \forall n \in \mathbb{N}$, then, $\forall x_1 \in U, \|\varphi_n(x_1, x_2)\|$, as a function of x_2 , is absolutely integrable over $\mathcal{X}_2$. Then, $\hat{g}_i(x_1) = g_i(x_1), i = 1, \dots, m, \forall x_1 \in U$. Hence, $\bar{p}_n(x_1) = \sum_{i=1}^m y_i \hat{g}_i(x_1) = \sum_{i=1}^m y_i g_i(x_1) = \sum_{i=1}^m y_i \mu_2(E_{i, x_1}) = \sum_{i=1}^m y_i \int_{X_2} \chi_{E_{i, x_1}, X_2} d\mu_2 = \int_{X_2} (\sum_{i=1}^m y_i \chi_{E_{i, x_1}, X_2}) d\mu_2$

$= \int_{X_2} \varphi_n(x_1, x_2) d\mu_2(x_2) \in \mathcal{Y}, \forall x_1 \in U$, where the first equality follows from the definition of $\bar{p}_n$; the second equality follows from the preceding discussion; the third equality follows from the definitions of g_i 's; the fourth equality follows from Proposition 11.75; and the fifth equality follows from Proposition 11.92. Define $\hat{p}_n : X_1 \rightarrow \mathcal{Y}$ by $\hat{p}_n(x_1) = \begin{cases} \bar{p}_n(x_1) & x_1 \in U \\ \vartheta_{\mathcal{Y}} & x_1 \in X_1 \setminus U \end{cases}$, $\forall x_1 \in X_1$. By Proposition 11.41, $\hat{p}_n$ is $\mathcal{B}_1$ -measurable. Then, $\bar{p}_n = \hat{p}_n$ a.e. in $\mathcal{X}_1$.

Note that, $\forall x_1 \in U$, we have

$$\begin{aligned} \|\hat{p}_n(x_1)\| &= \|\bar{p}_n(x_1)\| = \left\| \int_{X_2} \varphi_n(x_1, x_2) d\mu_2(x_2) \right\| \\ &\leq \int_{X_2} \mathcal{P} \circ \varphi_n(x_1, x_2) d\mu_2(x_2) \leq \int_{X_2} \mathcal{P} \circ f(x_1, x_2) d\mu_2(x_2) = \tilde{g}(x_1) < \infty \end{aligned}$$

where the second equality follows from the previous paragraph; the first inequality follows from Proposition 11.92; the second inequality follows from the property of φ_n ; and the last equality follows from the definition of U . Note also that, $\forall x_1 \in X_1 \setminus U$, we have $\|\hat{p}_n(x_1)\| = 0 = \tilde{g}(x_1)$. Then, $\|\hat{p}_n(x_1)\| \leq \tilde{g}(x_1), \forall x_1 \in X_1$. Hence, $\hat{p}_n$ is absolutely integrable over $\mathcal{X}_1$. By Proposition 11.92, $\int_{X_1 \times X_2} \varphi_n d\mu = \int_{X_1} \bar{p}_n d\mu_1 = \int_{X_1} \hat{p}_n d\mu_1$, where the first equality follows from the definition of $\bar{p}_n$.

By the result of the third preceding paragraph, $p(x_1) = \lim_{n \in \mathbb{N}} \int_{X_2} \varphi_n(x_1, x_2) d\mu_2(x_2) = \lim_{n \in \mathbb{N}} \bar{p}_n(x_1) = \lim_{n \in \mathbb{N}} \hat{p}_n(x_1), \forall x_1 \in U$, where the second equality follows from the result of the second preceding paragraph; the last equality follows from the definition of $\hat{p}_n$. Note that $p(x_1) = \vartheta_{\mathcal{Y}} = \lim_{n \in \mathbb{N}} \hat{p}_n(x_1), \forall x_1 \in X_1 \setminus U$. Then, $p(x_1) = \lim_{n \in \mathbb{N}} \hat{p}_n(x_1), \forall x_1 \in X_1$. This implies that, by Proposition 11.51, p is $\mathcal{B}_1$ -measurable. Then, we have $\int_{X_1 \times X_2} f d\mu = \lim_{n \in \mathbb{N}} \int_{X_1 \times X_2} \varphi_n d\mu = \lim_{n \in \mathbb{N}} \int_{X_1} \hat{p}_n d\mu_1 = \int_{X_1} p d\mu_1$, where the first equality follows from Lebesgue Dominated Convergence Theorem 11.91; the second equality follows from the result of the previous paragraph; and the last equality follows from the Lebesgue Dominated Convergence Theorem 11.91. Hence, (i) holds.

By symmetry, (ii) also holds. This completes the proof of the theorem. $\square$

Proposition 12.32 *Let $m \in \mathbb{N}$, $\mathcal{X}_i := (X_i, \mathcal{O}_i)$ be a second countable topological space, $\mathbb{X}_i := (\mathcal{X}_i, \mathcal{B}_i, \mu_i)$ be a second countable σ -finite topological measure space, $i = 1, \dots, m$, $\mathcal{X} := (\prod_{i=1}^m X_i, \mathcal{O}) := \prod_{i=1}^m \mathcal{X}_i$ be the product topological space, and $(\prod_{i=1}^m X_i, \mathcal{B}, \mu) := \prod_{i=1}^m (X_i, \mathcal{B}_i, \mu_i)$ be the product measure space. Then, $\mathbb{X} := (\mathcal{X}, \mathcal{B}, \mu)$ is a second countable σ -finite topological measure space, which is said to be the product topological measure space of $\mathbb{X}_1, \dots, \mathbb{X}_m$, denoted by $\mathbb{X} = \prod_{i=1}^m \mathbb{X}_i$.*

Proof By Proposition 12.22, $\mathcal{B}$ is the σ -algebra on $\prod_{i=1}^m X_i$ generated by $\mathcal{C} := \{\prod_{i=1}^m B_i \subseteq \prod_{i=1}^m X_i \mid B_i \in \mathcal{B}_i, \forall i \in \{1, \dots, m\}\}$. By $\mathbb{X}_i$

being a topological measure space, we have $\mathcal{B}_i = \mathcal{B}_B(\mathcal{X}_i)$, $i = 1, \dots, m$. Let $\bar{\mathcal{X}}_1 := \mathcal{X}_1$ and $\bar{\mathcal{X}}_i := \bar{\mathcal{X}}_{i-1} \times \mathcal{X}_i := ((\dots(X_1 \times X_2) \times \dots) \times X_i, \bar{\mathcal{O}}_i)$, $i = 2, \dots, m$, $(X_1, \bar{\mathcal{B}}_1, \bar{\mu}_1) := (X_1, \mathcal{B}_1, \mu_1)$ and $((\dots(X_1 \times X_2) \times \dots) \times X_i, \bar{\mathcal{B}}_i, \bar{\mu}_i) := ((\dots(X_1 \times X_2) \times \dots) \times X_{i-1}, \bar{\mathcal{B}}_{i-1}, \bar{\mu}_{i-1}) \times (X_i, \mathcal{B}_i, \mu_i)$, $i = 2, \dots, m$. By Proposition 3.28, $\bar{\mathcal{X}}_i$, $i = 1, \dots, m$, are second countable topological spaces, and by Proposition 11.24, $\bar{\mathcal{B}}_i = \mathcal{B}_B(\bar{\mathcal{X}}_i)$, $i = 1, \dots, m$. By Propositions 3.30 and 12.24, $\bar{\mathcal{X}}_m$ is homeomorphic to $\mathcal{X}$ and $((\dots(X_1 \times X_2) \times \dots) \times X_m, \bar{\mathcal{B}}_m, \bar{\mu}_m)$ is isomeasuric to $(\prod_{i=1}^m X_i, \mathcal{B}, \mu)$ with the same bijective mapping. Then, we have $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$. By Proposition 3.28, $\mathcal{X}$ is second countable.

First, consider the special case when $\mathbb{X}_i$, $i = 1, \dots, m$, are finite. Let $\mathcal{E} := \left\{ E \in \mathcal{B} \mid \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \text{ (i) } \exists U \in \mathcal{O} \text{ with } E \subseteq U \text{ such that } \mu(U \setminus E) < \epsilon; \text{ (ii) } \exists \bar{F} \in \mathcal{O} \text{ with } F \subseteq E \text{ such that } \mu(E \setminus \bar{F}) < \epsilon \right\}$. We will show that $\mathcal{E}$ is a σ -algebra on $\prod_{i=1}^m X_i$ and contains $\mathcal{C}$. Then, $\mathcal{E} = \mathcal{B}$ and $\mathbb{X}$ is a second countable finite topological measure space.

$\forall E := \prod_{i=1}^m B_i \in \mathcal{C}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, fix any $i \in \{1, \dots, m\}$, then $B_i \in \mathcal{B}_i$. Since $\mathbb{X}_i$ is a topological measure space, then $\exists O_i \in \mathcal{O}_i$ with $B_i \subseteq O_i$ such that $\mu_i(O_i \setminus B_i) < \frac{\epsilon \wedge 1}{1 + m \prod_{j=1, j \neq i}^m (1 + \mu_j(B_j))}$. By Proposition 11.27, $\exists X_i \setminus F_i \in \mathcal{O}_i$ with $F_i \subseteq B_i$ such that $\mu_i(B_i \setminus F_i) < \frac{\epsilon}{1 + m \prod_{j=1, j \neq i}^m \mu_j(B_j)}$. Let $U := \prod_{i=1}^m O_i \in \mathcal{O}$ and $F := \prod_{i=1}^m F_i$. Clearly, $(\prod_{i=1}^m X_i) \setminus F = \bigcup_{i=1}^m ((\prod_{j=1}^{i-1} X_j) \times (X_i \setminus F_i) \times (\prod_{j=i+1}^m X_j)) \in \mathcal{O}$. Then, $F \subseteq E \subseteq U$,

$$\begin{aligned}
 \mu(U \setminus E) &= \mu\left(\left(\prod_{i=1}^m O_i\right) \setminus \left(\prod_{i=1}^m B_i\right)\right) \\
 &= \mu\left(\bigcup_{i=1}^m \left(\left(\prod_{j=1}^{i-1} O_j\right) \times (O_i \setminus B_i) \times \left(\prod_{j=i+1}^m O_j\right)\right)\right) \\
 &\leq \sum_{i=1}^m \mu\left(\left(\prod_{j=1}^{i-1} O_j\right) \times (O_i \setminus B_i) \times \left(\prod_{j=i+1}^m O_j\right)\right) \\
 &= \sum_{i=1}^m \mu_i(O_i \setminus B_i) \prod_{j=1, j \neq i}^m \mu_j(O_j) \\
 &\leq \sum_{i=1}^m \frac{\epsilon \wedge 1}{1 + m \prod_{j=1, j \neq i}^m (1 + \mu_j(B_j))} \prod_{j=1, j \neq i}^m (1 + \mu_j(B_j)) < \epsilon
 \end{aligned}$$

where the first inequality follows from Proposition 11.6; and the third equality follows from Proposition 12.21, and

$$\mu(E \setminus F) = \mu\left(\left(\prod_{i=1}^m B_i\right) \setminus \left(\prod_{i=1}^m F_i\right)\right)$$

$$\begin{aligned}
&= \mu \left(\bigcup_{i=1}^m \left(\left(\prod_{j=1}^{i-1} B_j \right) \times (B_i \setminus F_i) \times \left(\prod_{j=i+1}^m B_j \right) \right) \right) \\
&\leq \sum_{i=1}^m \mu \left(\left(\prod_{j=1}^{i-1} B_j \right) \times (B_i \setminus F_i) \times \left(\prod_{j=i+1}^m B_j \right) \right) \\
&= \sum_{i=1}^m \mu_i(B_i \setminus F_i) \prod_{j=1, j \neq i}^m \mu_j(B_j) \\
&\leq \sum_{i=1}^m \frac{\epsilon}{1 + m \prod_{j=1, j \neq i}^m \mu_j(B_j)} \prod_{j=1, j \neq i}^m \mu_j(B_j) < \epsilon
\end{aligned}$$

where the first inequality follows from Proposition 11.6; and the third equality follows from Proposition 12.21. Hence, $E \in \mathcal{E}$. This shows that $\mathcal{C} \subseteq \mathcal{E}$.

$\forall E \in \mathcal{E}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists U \in \mathcal{O}$ with $E \subseteq U$ such that $\mu(U \setminus E) < \epsilon$, $\exists \tilde{F} \in \mathcal{O}$ with $F \subseteq E$ such that $\mu(E \setminus F) < \epsilon$. Then, $\mu(\tilde{E} \setminus \tilde{U}) = \mu(U \setminus E) < \epsilon$ and $\mu(\tilde{F} \setminus \tilde{E}) = \mu(E \setminus F) < \epsilon$. Hence, $\tilde{E} \in \mathcal{E}$.

$\forall (E_i)_{i=1}^\infty \subseteq \mathcal{E}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall i \in \mathbb{N}$, $\exists U_i \in \mathcal{O}$ with $E_i \subseteq U_i$ such that $\mu(U_i \setminus E_i) < 2^{-i}\epsilon$; $\exists \tilde{F}_i \in \mathcal{O}$ with $F_i \subseteq E_i$ such that $\mu(E_i \setminus F_i) < 2^{-i-1}\epsilon$. Let $U := \bigcup_{i=1}^\infty U_i \in \mathcal{O}$. Then, $E = \bigcup_{i=1}^\infty E_i \subseteq \bigcup_{i=1}^\infty U_i = U$ and $\mu(U \setminus E) = \mu(\bigcup_{i=1}^\infty (U_i \setminus E)) \leq \sum_{i=1}^\infty \mu(U_i \setminus E) \leq \sum_{i=1}^\infty \mu(U_i \setminus E_i) < \epsilon$, where the first inequality follows from Proposition 11.6; and the second inequality follows from Proposition 11.4. Hence, E satisfies (i). By Proposition 11.7, $\mu(E) = \mu(\bigcup_{i=1}^\infty E_i) = \lim_{n \in \mathbb{N}} \mu(\bigcup_{i=1}^n E_i)$. Then, by $\mu(E) \leq \mu(\prod_{i=1}^m X_i) = \prod_{i=1}^m \mu_i(X_i) < \infty$, $\exists n_0 \in \mathbb{N}$ such that $\mu(\bigcup_{i=1}^{n_0} E_i) \leq \mu(E) < \mu(\bigcup_{i=1}^{n_0} E_i) + \epsilon/2$. Let $F := \bigcup_{i=1}^{n_0} F_i$. Clearly, $\tilde{F} = \bigcap_{i=1}^{n_0} \tilde{F}_i \in \mathcal{O}$. Then, $F \subseteq E$ and $\mu(E \setminus F) = \mu(E \setminus (\bigcup_{i=1}^{n_0} E_i)) + \mu((\bigcup_{i=1}^{n_0} E_i) \setminus F) = \mu(E) - \mu(\bigcup_{i=1}^{n_0} E_i) + \mu(\bigcup_{i=1}^{n_0} (E_i \setminus F)) < \epsilon/2 + \sum_{i=1}^{n_0} \mu(E_i \setminus F) \leq \epsilon/2 + \sum_{i=1}^{n_0} \mu(E_i \setminus F_i) < \epsilon$, where the first and the second equalities follow from Fact 11.3; the first inequality follows from Proposition 11.6; and the second inequality follows from Proposition 11.4. Hence, E satisfies (ii). Then, $E = \bigcup_{i=1}^\infty E_i \in \mathcal{E}$.

Clearly, $\emptyset = \prod_{i=1}^m \emptyset \in \mathcal{C} \subseteq \mathcal{E}$ and $\prod_{i=1}^m X_i \in \mathcal{C} \subseteq \mathcal{E}$. This shows that $\mathcal{E}$ is a σ -algebra on $\prod_{i=1}^m X_i$ and contains $\mathcal{C}$. This completes the proof of the special case.

Next, consider the general case when $\mathbb{X}_i, i = 1, \dots, m$, are σ -finite. Fix any $i = 1, \dots, m$. By $\mathbb{X}_i$ being σ -finite, $\exists (X_{i,n})_{n=1}^\infty \subseteq \mathcal{B}_i$ such that $X_i = \bigcup_{n=1}^\infty X_{i,n}$ and $\mu_i(X_{i,n}) < \infty, \forall n \in \mathbb{N}$. By $\mathbb{X}_i$ being a topological measure space, without loss of generality, we may assume that $X_{i,n} \subseteq X_{i,n+1}$ and $X_{i,n} \in \mathcal{O}_i, \forall n \in \mathbb{N}$. $\forall n \in \mathbb{N}$, by Proposition 11.29, let $\mathbb{X}_{i,n} := ((X_{i,n}, \mathcal{O}_{i,n}), \mathcal{B}_{i,n}, \mu_{i,n})$ be the finite topological measure subspace of $\mathbb{X}_i$. $\forall n \in \mathbb{N}$, by the special case, let $\hat{\mathbb{X}}_n := \prod_{i=1}^m \mathbb{X}_{i,n} := ((\prod_{i=1}^m X_{i,n}, \hat{\mathcal{O}}_n), \hat{\mathcal{B}}_n, \hat{\mu}_n)$ be the second countable finite product topological measure space. By Proposition 12.23, $(\prod_{i=1}^m X_{i,n}, \hat{\mathcal{B}}_n, \hat{\mu}_n)$ is the finite

measure subspace of $(\prod_{i=1}^m X_i, \mathcal{B}, \mu)$. $\forall E \in \mathcal{B}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N}$, let $E_n := E \cap (\prod_{i=1}^m X_{i,n}) \in \hat{\mathcal{B}}_n$. Since $\hat{\mathbb{X}}_n$ is a topological measure space, $\exists U_n \in \hat{\mathcal{O}}_n$ with $E_n \subseteq U_n$ such that $\hat{\mu}_n(U_n \setminus E_n) < 2^{-n}\epsilon$. Since $X_{i,n} \in \mathcal{O}_i$, $i = 1, \dots, m$, then $\prod_{i=1}^m X_{i,n} \in \mathcal{O}$, which further implies that $U_n \in \mathcal{O}$. Let $U := \bigcup_{n=1}^\infty U_n \in \mathcal{O}$. Then, $E = \bigcup_{n=1}^\infty E_n \subseteq \bigcup_{n=1}^\infty U_n = U$ and $\mu(U \setminus E) = \mu(\bigcup_{n=1}^\infty (U_n \setminus E)) \leq \sum_{n=1}^\infty \mu(U_n \setminus E) \leq \sum_{n=1}^\infty \mu(U_n \setminus E_n) = \sum_{n=1}^\infty \hat{\mu}_n(U_n \setminus E_n) < \epsilon$, where the first inequality follows from Proposition 11.6; the second inequality follows from Proposition 11.4; and the second equality follows from Proposition 11.13. By the arbitrariness of E , we have $\mathbb{X}$ is a σ -finite topological measure space. It is second countable since $\mathcal{X}$ is second countable. This completes the proof of the proposition. $\square$

Proposition 12.33 *Let $m \in \mathbb{N}$, $\mathcal{X}_i := (X_i, \mathcal{O}_i)$ be a second countable topological space, $\mathbb{X}_i := (\mathcal{X}_i, \mathcal{B}_i, \mu_i)$ be a second countable σ -finite $\mathbb{K}$ -valued topological measure space, $i = 1, \dots, m$, $\mathcal{X} := (\prod_{i=1}^m X_i, \mathcal{O}) := \prod_{i=1}^m \mathcal{X}_i$ be the product topological space, and $(\prod_{i=1}^m X_i, \mathcal{B}, \mu) := \prod_{i=1}^m (X_i, \mathcal{B}_i, \mu_i)$ be the product $\mathbb{K}$ -valued measure space. Then, $\mathbb{X} := (\mathcal{X}, \mathcal{B}, \mu)$ is a second countable σ -finite $\mathbb{K}$ -valued topological measure space, which is said to be the product $\mathbb{K}$ -valued topological measure space of $\mathbb{X}_1, \dots, \mathbb{X}_m$, denoted by $\mathbb{X} = \prod_{i=1}^m \mathbb{X}_i$.*

Proof Let $\hat{\mathcal{X}}_i := (X_i, \mathcal{B}_i, \mathcal{P} \circ \mu_i =: \nu_i)$, $i = 1, \dots, m$, and $\hat{\mathcal{X}} := (\prod_{i=1}^m X_i, \mathcal{B}, \mathcal{P} \circ \mu =: \nu)$. By Proposition 12.22, $\hat{\mathcal{X}}_i$, $i = 1, \dots, m$, and $\hat{\mathcal{X}}$ are σ -finite measure spaces and $\hat{\mathcal{X}} = \prod_{i=1}^m \hat{\mathcal{X}}_i$. Since $\mathbb{X}_i$ is a second countable σ -finite $\mathbb{K}$ -valued topological measure space, then, by Definition 11.191, $\hat{\mathbb{X}}_i := (\mathcal{X}_i, \mathcal{B}_i, \nu_i)$ is a second countable σ -finite topological measure space, $i = 1, \dots, m$. By Proposition 12.32, $\hat{\mathbb{X}} := (\mathcal{X}, \mathcal{B}, \nu)$ is a second countable σ -finite topological measure space. Then, $\mathbb{X}$ is a second countable σ -finite $\mathbb{K}$ -valued topological measure space. This completes the proof of the proposition. $\square$

Proposition 12.34 *Let $m \in \mathbb{N}$, $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a σ -finite measure space, ν_i be a σ -finite $\mathbb{K}$ -valued measure on $(X_i, \mathcal{B}_i)$ with $\frac{d\nu_i}{d\mu_i} = f_i : X_i \rightarrow \mathbb{K}$ a.e. in $\mathcal{X}_i$, $i = 1, \dots, m$, and $f : \prod_{i=1}^m X_i \rightarrow \mathbb{K}$ be defined by $f(x_1, \dots, x_m) = \prod_{i=1}^m f_i(x_i)$, $\forall (x_1, \dots, x_m) \in \prod_{i=1}^m X_i$. Then, $\frac{d\nu}{d\mu} = f$ a.e. in $\mathcal{X}$, where $\nu := \prod_{i=1}^m \nu_i$, $\mu := \prod_{i=1}^m \mu_i$, and $\mathcal{X} := (\prod_{i=1}^m X_i, \mathcal{B}, \mu) := \prod_{i=1}^m \mathcal{X}_i$.*

Proof By Proposition 12.22, $\mathcal{X}$ is a σ -finite measure space. By Propositions 7.23, 11.38, and 11.39, f is $\mathcal{B}$ -measurable. By Proposition 11.116 and Definition 11.166, f is a Radon-Nikodym derivative of a σ -finite $\mathbb{K}$ -valued measure $\bar{\nu}$ on $(\prod_{i=1}^m X_i, \mathcal{B})$ with respect to μ . By Proposition 11.167, $\frac{d\bar{\nu}}{d\mu} = f$ a.e. in $\mathcal{X}$.

First, consider the special case when ν_i , $i = 1, \dots, m$, are finite $\mathbb{K}$ -valued measures. Let $\mathcal{E} := \{E \in \mathcal{B} \mid \nu(E) = \bar{\nu}(E)\}$ and $\mathcal{C} := \{\prod_{i=1}^m B_i \subseteq$

$\prod_{i=1}^m X_i \mid B_i \in \mathcal{B}_i, \forall i \in \{1, \dots, m\}$, where $\mathcal{C}$ is the set of measurable rectangles in $\prod_{i=1}^m X_i$. $\forall E := \prod_{i=1}^m B_i \in \mathcal{C}$, we have $\nu(E) = \prod_{i=1}^m \nu_i(B_i)$ by Proposition 12.21. By repeated application of Tonelli Theorem 12.29 and Proposition 12.24, we have

$$\begin{aligned} \int_{\prod_{i=1}^m X_i} \mathcal{P} \circ f \, d\mu &= \left(\int_{\prod_{i=1}^{m-1} X_i} \prod_{i=1}^{m-1} \mathcal{P} \circ f_i \, d\left(\prod_{i=1}^{m-1} \mu_i\right) \right) \left(\int_{X_m} \mathcal{P} \circ f_m \, d\mu_m \right) \\ &= \left(\int_{\prod_{i=1}^{m-1} X_i} \prod_{i=1}^{m-1} \mathcal{P} \circ f_i \, d\left(\prod_{i=1}^{m-1} \mu_i\right) \right) \mathcal{P} \circ \nu_m(X_m) = \cdots = \prod_{i=1}^m \mathcal{P} \circ \nu_i(X_i) \\ &< \infty \end{aligned}$$

where the first equality follows from Tonelli Theorem 12.29; and the second equality follows from Definition 11.166. Hence, f is absolutely integrable over $\mathcal{X}$ and $\bar{\nu}$ is a finite $\mathbb{K}$ -valued measure by Proposition 11.116. By Proposition 11.83, $f\chi_{\prod_{i=1}^m B_i, \prod_{i=1}^m X_i}$ is absolutely integrable over $\mathcal{X}$. By repeated application of Fubini Theorem 12.30,

$$\begin{aligned} \bar{\nu}(E) &= \int_{\prod_{i=1}^m B_i} f \, d\mu = \int_{\prod_{i=1}^m X_i} f \chi_{\prod_{i=1}^m B_i, \prod_{i=1}^m X_i} \, d\left(\prod_{i=1}^m \mu_i\right) \\ &= \left(\int_{\prod_{i=1}^{m-1} X_i} \left(\prod_{i=1}^{m-1} f_i \right) \chi_{\prod_{i=1}^{m-1} B_i, \prod_{i=1}^{m-1} X_i} \, d\left(\prod_{i=1}^{m-1} \mu_i\right) \right) \left(\int_{B_m} f_m \, d\mu_m \right) \\ &= \left(\int_{\prod_{i=1}^{m-1} X_i} \left(\prod_{i=1}^{m-1} f_i \right) \chi_{\prod_{i=1}^{m-1} B_i, \prod_{i=1}^{m-1} X_i} \, d\left(\prod_{i=1}^{m-1} \mu_i\right) \right) \nu_m(B_m) \\ &= \cdots = \prod_{i=1}^m \nu_i(B_i) = \nu(E) \end{aligned}$$

where the first equality follows from Definition 11.166; the second equality follows from Proposition 11.92; the third equality follows from Fubini Theorem; and the fourth equality follows from Definition 11.166. This implies that $E \in \mathcal{E}$. By the arbitrariness of E , we have $\mathcal{C} \subseteq \mathcal{E}$.

Clearly, $\mathcal{C}$ is a π -system on $\prod_{i=1}^m X_i$. We will show that $\mathcal{E}$ is a monotone class on $\prod_{i=1}^m X_i$. Then, by Monotone Class Lemma 12.19, we have $\mathcal{E} = \mathcal{B}$. This further implies that $\nu = \bar{\nu}$ and $\frac{d\nu}{d\mu} = f$ a.e. in $\mathcal{X}$, which completes the proof of the special case.

Note that $\emptyset = \prod_{i=1}^m \emptyset \in \mathcal{C} \subseteq \mathcal{E}$ and $\prod_{i=1}^m X_i \in \mathcal{C} \subseteq \mathcal{E}$. $\forall E_1, E_2 \in \mathcal{E}$ with $E_1 \subseteq E_2$, we have $\nu(E_2 \setminus E_1) = \nu(E_2) - \nu(E_1) = \bar{\nu}(E_2) - \bar{\nu}(E_1) = \bar{\nu}(E_2 \setminus E_1)$, where the first equality follows from the fact that ν is a finite $\mathbb{K}$ -valued measure; the second equality follows from the fact that $E_1, E_2 \in \mathcal{E}$; and the third equality follows from fact that $\bar{\nu}$ is a finite $\mathbb{K}$ -valued measure. Hence, $E_2 \setminus E_1 \in \mathcal{E}$.

$\forall (E_i)_{i=1}^\infty \subseteq \mathcal{E}$ with $E_i \subseteq E_{i+1}, \forall i \in \mathbb{N}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$. Then, $\nu(E) = \lim_{i \in \mathbb{N}} \nu(E_i) = \lim_{i \in \mathbb{N}} \bar{\nu}(E_i) = \bar{\nu}(E)$, where the first equality

follows from Proposition 11.112 and the fact that ν is finite; the second equality follows from the fact that $E_i \in \mathcal{E}$, $\forall i \in \mathbb{N}$; and the third equality follows from Proposition 11.112 and the fact that $\bar{\nu}$ is finite. Then, $E = \bigcup_{i=1}^{\infty} E_i \in \mathcal{E}$.

Hence, $\mathcal{E}$ is a monotone class on $\prod_{i=1}^m X_i$. This completes the proof of the special case.

Next, consider the general case when ν_i , $i = 1, \dots, m$, are σ -finite. Then, there exists $(X_{i,n})_{n=1}^{\infty} \subseteq \mathcal{B}_i$ such that $X_i = \bigcup_{n=1}^{\infty} X_{i,n}$ and $\mathcal{P} \circ \nu_i(X_{i,n}) < \infty$, $\forall n \in \mathbb{N}$, $i = 1, \dots, m$. Without loss of generality, we may assume that $(X_{i,n})_{n=1}^{\infty}$ is pairwise disjoint, $i = 1, \dots, m$. For any $i = 1, \dots, m$ and $\forall n \in \mathbb{N}$, let $\mathcal{X}_{i,n} := (X_{i,n}, \mathcal{B}_{i,n}, \mu_{i,n})$ be the σ -finite measure subspace of $\mathcal{X}_i$, $\hat{\mathcal{X}}_{i,n} := (X_{i,n}, \mathcal{B}_{i,n}, \nu_{i,n})$ be the finite $\mathbb{K}$ -valued measure subspace of $\hat{\mathcal{X}}_i := (X_i, \mathcal{B}_i, \nu_i)$. Fix any $n_1, \dots, n_m \in \mathbb{N}$. Let $\bar{\mathcal{X}}_{n_1, \dots, n_m} := (\prod_{i=1}^m X_{i,n_i}, \bar{\mathcal{B}}_{n_1, \dots, n_m}, \bar{\mu}_{n_1, \dots, n_m}) := \prod_{i=1}^m \mathcal{X}_{i,n_i}$ be the σ -finite product measure space, and $\tilde{\mathcal{X}}_{n_1, \dots, n_m} := (\prod_{i=1}^m X_{i,n_i}, \bar{\mathcal{B}}_{n_1, \dots, n_m}, \tilde{\nu}_{n_1, \dots, n_m}) := \prod_{i=1}^m \hat{\mathcal{X}}_{i,n_i}$ be the finite product $\mathbb{K}$ -valued measure space. By Proposition 12.23, $\bar{\mathcal{X}}_{n_1, \dots, n_m}$ is the measure subspace of $\mathcal{X}$ and $\tilde{\mathcal{X}}_{n_1, \dots, n_m}$ is the finite $\mathbb{K}$ -valued measure subspace of $\tilde{\mathcal{X}} := (\prod_{i=1}^m X_i, \mathcal{B}, \nu)$. By Definition 11.166, we have $\frac{d\nu_{i,n_i}}{d\mu_{i,n_i}} = f_i|_{X_{i,n_i}}$ a.e. in $\mathcal{X}_{i,n_i}$, $i = 1, \dots, m$. By the special case, we have $\frac{d\tilde{\nu}_{n_1, \dots, n_m}}{d\bar{\mu}_{n_1, \dots, n_m}} = f|_{\prod_{i=1}^m X_{i,n_i}}$ a.e. in $\bar{\mathcal{X}}_{n_1, \dots, n_m}$. $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \nu(E) < \infty$, let $E_{n_1, \dots, n_m} := E \cap (\prod_{i=1}^m X_{i,n_i}) \in \bar{\mathcal{B}}_{n_1, \dots, n_m}$. Then, we have

$$\begin{aligned} \mathcal{P} \circ \nu(E) &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \mathcal{P} \circ \nu(E_{n_1, \dots, n_m}) \\ &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \mathcal{P} \circ \tilde{\nu}_{n_1, \dots, n_m}(E_{n_1, \dots, n_m}) \\ &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \int_{E_{n_1, \dots, n_m}} \mathcal{P} \circ (f|_{\prod_{i=1}^m X_{i,n_i}}) d\bar{\mu}_{n_1, \dots, n_m} \\ &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \int_{E_{n_1, \dots, n_m}} \mathcal{P} \circ f d\mu = \int_E \mathcal{P} \circ f d\mu < \infty \end{aligned}$$

where the first equality follows from the fact that $\mathcal{P} \circ \nu$ is a measure; the second equality follows from Proposition 11.115; the third equality follows from the fact that $\frac{d\tilde{\nu}_{n_1, \dots, n_m}}{d\bar{\mu}_{n_1, \dots, n_m}} = f|_{\prod_{i=1}^m X_{i,n_i}}$ a.e. in $\bar{\mathcal{X}}_{n_1, \dots, n_m}$; and the last two equalities follow from Proposition 11.83. We also have

$$\begin{aligned} \nu(E) &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \nu(E_{n_1, \dots, n_m}) = \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \tilde{\nu}_{n_1, \dots, n_m}(E_{n_1, \dots, n_m}) \\ &= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \int_{E_{n_1, \dots, n_m}} f|_{\prod_{i=1}^m X_{i,n_i}} d\bar{\mu}_{n_1, \dots, n_m} \end{aligned}$$

$$= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \int_{E_{n_1, \dots, n_m}} f \, d\mu = \int_E f \, d\mu = \bar{\nu}(E)$$

where the first equality follows from Definition 11.108; the second equality follows from Proposition 11.115; the third equality follows from the fact that $\frac{d\bar{\nu}_{n_1, \dots, n_m}}{d\mu_{n_1, \dots, n_m}} = f|_{\prod_{i=1}^m X_{i, n_i}}$ a.e. in $\bar{\mathcal{X}}_{n_1, \dots, n_m}$; the fourth equality follows from Proposition 11.92; the fifth equality follows from Proposition 11.92 and the fact that $\int_E \mathcal{P} \circ f \, d\mu < \infty$; and the last equality follows from the fact that $\frac{d\bar{\nu}}{d\mu} = f$ a.e. in $\mathcal{X}$. Therefore, by Proposition 11.137, we have $\bar{\nu} = \nu$ and $\frac{d\nu}{d\mu} = f$ a.e. in $\mathcal{X}$.

This completes the proof of the proposition. $\square$

Theorem 12.35 (Fubini) *Let $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a σ -finite $\mathbb{K}$ -valued measure space, $i = 1, 2$, $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := \mathcal{X}_1 \times \mathcal{X}_2$ be the product $\mathbb{K}$ -valued measure space, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, and $f : X_1 \times X_2 \rightarrow \mathcal{Y}$ be absolutely integrable over $\mathcal{X}$. Define $f_{x_1} : X_2 \rightarrow \mathcal{Y}$ and $f_{\langle x_2 \rangle} : X_1 \rightarrow \mathcal{Y}$ as in Proposition 12.28. Then, the following statements hold.*

- (i) *There exists $\check{p} : X_1 \rightarrow \mathcal{Y}$ and exists $\check{U} \in \mathcal{B}_1$ such that $\mathcal{P} \circ \mu_1(X_1 \setminus \check{U}) = 0$; $\forall x_1 \in \check{U}$, f_{x_1} is absolutely integrable over $\mathcal{X}_2$ and $\check{p}(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \mathcal{Y}$; $\forall x_1 \in X_1 \setminus \check{U}$, $\check{p}(x_1) = \vartheta_{\mathcal{Y}}$; $\check{p}$ is absolutely integrable over $\mathcal{X}_1$; and $\int_{X_1 \times X_2} f \, d\mu = \int_{X_1} \check{p} \, d\mu_1 \in \mathcal{Y}$.*
- (ii) *There exists $\check{q} : X_2 \rightarrow \mathcal{Y}$ and exists $\check{V} \in \mathcal{B}_2$ such that $\mathcal{P} \circ \mu_2(X_2 \setminus \check{V}) = 0$; $\forall x_2 \in \check{V}$, $f_{\langle x_2 \rangle}$ is absolutely integrable over $\mathcal{X}_1$ and $\check{q}(x_2) = \int_{X_1} f_{\langle x_2 \rangle} \, d\mu_1 \in \mathcal{Y}$; $\forall x_2 \in X_2 \setminus \check{V}$, $\check{q}(x_2) = \vartheta_{\mathcal{Y}}$; $\check{q}$ is absolutely integrable over $\mathcal{X}_2$; and $\int_{X_1 \times X_2} f \, d\mu = \int_{X_2} \check{q} \, d\mu_2 \in \mathcal{Y}$.*
- (iii) *$\forall p : X_1 \rightarrow \mathcal{Y}$ satisfying: p is $\mathcal{B}_1$ -measurable; and $\exists U \in \mathcal{B}_1$ with $\mathcal{P} \circ \mu_1(X_1 \setminus U) = 0$ such that, $\forall x_1 \in U$, f_{x_1} is absolutely integrable over $\mathcal{X}_2$ and $p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \mathcal{Y}$, we have p is absolutely integrable over $\mathcal{X}_1$ and $\int_{X_1 \times X_2} f \, d\mu = \int_{X_1} p \, d\mu_1 \in \mathcal{Y}$.*
- (iv) *$\forall q : X_2 \rightarrow \mathcal{Y}$ satisfying: q is $\mathcal{B}_2$ -measurable; and $\exists V \in \mathcal{B}_2$ with $\mathcal{P} \circ \mu_2(X_2 \setminus V) = 0$ such that, $\forall x_2 \in V$, $f_{\langle x_2 \rangle}$ is absolutely integrable over $\mathcal{X}_1$ and $q(x_2) = \int_{X_1} f_{\langle x_2 \rangle} \, d\mu_1 \in \mathcal{Y}$, we have q is absolutely integrable over $\mathcal{X}_2$ and $\int_{X_1 \times X_2} f \, d\mu = \int_{X_2} q \, d\mu_2 \in \mathcal{Y}$.*

Proof (i) Let $\hat{\mathcal{X}}_i := (X_i, \mathcal{B}_i, \mathcal{P} \circ \mu_i =: \nu_i)$, $i = 1, 2$, and $\hat{\mathcal{X}} := (X_1 \times X_2, \mathcal{B}, \mathcal{P} \circ \mu =: \nu)$. By Proposition 12.22, $\hat{\mathcal{X}}_1$, $\hat{\mathcal{X}}_2$, and $\hat{\mathcal{X}}$ are σ -finite measure spaces and $\hat{\mathcal{X}} = \hat{\mathcal{X}}_1 \times \hat{\mathcal{X}}_2$. Fix any $i = 1, 2$. By Radon-Nikodym Theorem 11.171, we may define functions $\frac{d\mu_i}{d\nu_i} =: g_i : X_i \rightarrow \mathbb{K}$. $\forall E \in \mathcal{B}_i$, $\nu_i(E) = \mathcal{P} \circ \mu_i(E) = \int_E \mathcal{P} \circ g_i \, d\nu_i$. By Proposition 11.167, we have $\mathcal{P} \circ g_i = 1$ a.e. in $\hat{\mathcal{X}}_i$. By Proposition 12.34, we have $\frac{d\mu}{d\nu} = g_1 g_2$ a.e. in $\hat{\mathcal{X}}$.

Since f is absolutely integrable over $\mathcal{X}$, then, by Proposition 11.168, $\int_{X_1 \times X_2} f \, d\mu = \int_{X_1 \times X_2} (fg_1g_2) \, d\nu$ and $(\mathcal{P} \circ f)(\mathcal{P} \circ (g_1g_2)) = (\mathcal{P} \circ f)(\mathcal{P} \circ g_1)(\mathcal{P} \circ g_2) = \mathcal{P} \circ (fg_1g_2)$ is integrable over $\hat{\mathcal{X}}$. Define $\bar{f} := fg_1g_2 : X_1 \times X_2 \rightarrow \mathcal{Y}$. By Fubini Theorem 12.30, there exists $\bar{p} : X_1 \rightarrow \mathcal{Y}$ and exists $\bar{U} \in \mathcal{B}_1$ such that $\nu_1(X_1 \setminus \bar{U}) = 0$; $\forall x_1 \in \bar{U}$, $\bar{f}_{x_1}$ is absolutely integrable over $\hat{\mathcal{X}}_2$ and $\bar{p}(x_1) = \int_{X_2} \bar{f}_{x_1} \, d\nu_2$; $\forall x_1 \in X_1 \setminus \bar{U}$, $\bar{p}(x_1) = \vartheta_{\mathcal{Y}}$; $\bar{p}$ is absolutely integrable over $\hat{\mathcal{X}}_1$; and $\int_{X_1 \times X_2} \bar{f} \, d\nu = \int_{X_1} \bar{p} \, d\nu_1$. Let $A_1 := \{x_1 \in X_1 \mid g_1(x_1) = 0\}$. Clearly, $A_1 \in \mathcal{B}_1$ and $\nu_1(A_1) = 0$. Let $\check{U} := \bar{U} \setminus A_1 \in \mathcal{B}_1$. Then, $\mathcal{P} \circ \mu_1(X_1 \setminus \check{U}) = \nu_1(X_1 \setminus \check{U}) = \nu_1((X_1 \setminus \bar{U}) \cup A_1) = 0$. $\forall x_1 \in \check{U}$, $\bar{f}_{x_1}$ is absolutely integrable over $\hat{\mathcal{X}}_2$ implies that $f_{x_1}g_1(x_1)g_2$ is absolutely integrable over $\hat{\mathcal{X}}_2$, which further implies that $(\mathcal{P} \circ f_{x_1})(\mathcal{P} \circ g_2)$ is integrable over $\hat{\mathcal{X}}_2$. By Proposition 11.168, f_{x_1} is absolutely integrable over $\mathcal{X}_2$ and $\int_{X_2} f_{x_1} \, d\mu_2 = \int_{X_2} (f_{x_1}g_2) \, d\nu_2$. Then, $\bar{p}(x_1) = g_1(x_1) \int_{X_2} f_{x_1} \, d\mu_2$. Define $\check{p} : X_1 \rightarrow \mathcal{Y}$ by $\check{p}(x_1) = \begin{cases} \int_{X_2} f_{x_1} \, d\mu_2 & x_1 \in \check{U} \\ \vartheta_{\mathcal{Y}} & x_1 \in X_1 \setminus \check{U} \end{cases}, \forall x_1 \in X_1$. Then, $\check{p}(x_1) = \begin{cases} \bar{p}(x_1)/g_1(x_1) & x_1 \in \check{U} \\ \vartheta_{\mathcal{Y}} & x_1 \in X_1 \setminus \check{U} \end{cases}$. By Propositions 11.41, 11.38, and 11.39, $\check{p}$ is $\mathcal{B}_1$ -measurable. Since $\bar{p}$ and $g_1\check{p}$ are $\mathcal{B}_1$ -measurable, then, by Lemma 11.43, $\bar{p} = g_1\check{p}$ a.e. in $\mathcal{X}_1$. Since $\bar{p}$ is absolutely integrable over $\hat{\mathcal{X}}_1$, then, $\mathcal{P} \circ (g_1\check{p}) = (\mathcal{P} \circ \check{p})(\mathcal{P} \circ g_1)$ is integrable over $\hat{\mathcal{X}}_1$. By Proposition 11.168, $\check{p}$ is absolutely integrable over $\mathcal{X}_1$ and $\int_{X_1} (\check{p}g_1) \, d\nu_1 = \int_{X_1} \check{p} \, d\mu_1$. By Proposition 11.92, we have $\int_{X_1 \times X_2} f \, d\mu = \int_{X_1 \times X_2} \bar{f} \, d\nu = \int_{X_1} \bar{p} \, d\nu_1 = \int_{X_1} (\check{p}g_1) \, d\nu_1 = \int_{X_1} \check{p} \, d\mu_1$. Hence, (i) holds.

By symmetry, (ii) holds.

(iii) Fix any $p : X_1 \rightarrow \mathcal{Y}$ satisfying: p is $\mathcal{B}_1$ -measurable; and $\exists U \in \mathcal{B}_1$ with $\mathcal{P} \circ \mu_1(X_1 \setminus U) = 0$ such that, $\forall x_1 \in U$, f_{x_1} is absolutely integrable over $\mathcal{X}_2$ and $p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \mathcal{Y}$. By (i), we have $U_0 := \check{U} \cap U \in \mathcal{B}_1$ and $\mathcal{P} \circ \mu_1(X_1 \setminus U_0) = \mathcal{P} \circ \mu_1((X_1 \setminus \check{U}) \cup (X_1 \setminus U)) = 0$. $\forall x_1 \in U_0$, $p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 = \check{p}(x_1) \in \mathcal{Y}$. By Lemma 11.43, $p = \check{p}$ a.e. in $\mathcal{X}_1$. By Proposition 11.83, we have p is absolutely integrable over $\mathcal{X}_1$ since $\check{p}$ is so. By Proposition 11.92, $\int_{X_1} p \, d\mu_1 = \int_{X_1} \check{p} \, d\mu_1 = \int_{X_1 \times X_2} f \, d\mu \in \mathcal{Y}$. Hence, (iii) holds.

By symmetry, (iv) holds. This completes the proof of the theorem. $\square$

12.4 Functions of Bounded Variation

Definition 12.36 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $D \subseteq \mathbb{R}^m$, and $F : D \rightarrow \mathcal{Y}$. F is said to be continuous on the right at $x_0 \in D$ if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap D$ with $x_0 \leq x$, we have $\|F(x) - F(x_0)\| < \epsilon$. F is said to be continuous on the right if it is continuous on the right at x , $\forall x \in D$. F is said to be continuous on the right on $E \subseteq D$ if it is continuous on the

right at x , $\forall x \in E$.

Proposition 12.37 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone P , $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $x_0 \in \Omega \subseteq \mathbb{R}^m$, and $F : \Omega \rightarrow \mathcal{Y}$. Then, the following statements are equivalent.*

- (i) F is continuous on the right at x_0 .
- (ii) If x_0 is an accumulation point of $\Omega \cap (x_0 + P)$, then $\lim_{\substack{h \rightarrow \mathbf{0}_m \\ h \geq \mathbf{0}_m}} F(x_0 + h) = F(x_0)$.
- (iii) If x_0 is an accumulation point of $\Omega \cap (x_0 + P)$, then $\forall (h_n)_{n=1}^\infty \subseteq (P \cap (\Omega - x_0)) \setminus \{\mathbf{0}_m\}$ with $\lim_{n \in \mathbb{N}} h_n = \mathbf{0}_m$, we have $\lim_{n \in \mathbb{N}} F(x_0 + h_n) = F(x_0)$.

Proof (i) $\Rightarrow$ (iii). Let F be continuous on the right at x_0 and x_0 be an accumulation point of $\Omega \cap (x_0 + P)$. By (i), $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap \Omega$ with $x_0 \leq x$, we have $\|F(x) - F(x_0)\| < \epsilon$. Then, $\forall (h_n)_{n=1}^\infty \subseteq (P \cap (\Omega - x_0)) \setminus \{\mathbf{0}_m\}$ with $\lim_{n \in \mathbb{N}} h_n = \mathbf{0}_m$, $\exists N \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $N \leq n$, $h_n \in \mathcal{B}_{\mathbb{R}^m}(\mathbf{0}_m, \delta)$. This implies that $x_0 + h_n \in \mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap \Omega$, $x_0 + h_n \geq x_0$, and $\|F(x_0 + h_n) - F(x_0)\| < \epsilon$. Hence, $\lim_{n \in \mathbb{N}} F(x_0 + h_n) = F(x_0)$.

(iii) $\Rightarrow$ (ii). Let $D := P \cap (\Omega - x_0)$. Define $\bar{F} : D \rightarrow \mathcal{Y}$ by $\bar{F}(h) = F(x_0 + h)$, $\forall h \in D$. By (iii), we have $\forall (h_n)_{n=1}^\infty \subseteq D \setminus \{\mathbf{0}_m\}$ with $\lim_{n \in \mathbb{N}} h_n = \mathbf{0}_m$, we have $\lim_{n \in \mathbb{N}} \bar{F}(h_n) = F(x_0)$. By Proposition 4.16, we have $\lim_{h \rightarrow \mathbf{0}_m} \bar{F}(h) = F(x_0)$. Hence, $\lim_{\substack{h \rightarrow \mathbf{0}_m \\ h \geq \mathbf{0}_m}} F(x_0 + h) = F(x_0)$.

(ii) $\Rightarrow$ (i). We will distinguish two exhaustive and mutually exclusive cases: Case 1: x_0 is not an accumulation point of $\Omega \cap (x_0 + P)$; Case 2: x_0 is an accumulation point of $\Omega \cap (x_0 + P)$. Case 1: x_0 is not an accumulation point of $\Omega \cap (x_0 + P)$. $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap ((\Omega \cap (x_0 + P)) \setminus \{x_0\}) = \emptyset$ and $\mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap \Omega \cap (x_0 + P) = \{x_0\}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap \Omega$ with $x_0 \leq x$, we have $x = x_0$ and $\|F(x) - F(x_0)\| = 0 < \epsilon$. Hence, F is continuous on the right at x_0 .

Case 2: x_0 is an accumulation point of $\Omega \cap (x_0 + P)$. By (ii), $\lim_{\substack{h \rightarrow \mathbf{0}_m \\ h \geq \mathbf{0}_m}} F(x_0 + h) = F(x_0)$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall h \in (\mathcal{B}_{\mathbb{R}^m}(\mathbf{0}_m, \delta) \cap P \cap (\Omega - x_0)) \setminus \{\mathbf{0}_m\}$, we have $\|F(x_0 + h) - F(x_0)\| < \epsilon$. Then, $\forall x \in \mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap \Omega \cap (x_0 + P)$, we have $x - x_0 \in \mathcal{B}_{\mathbb{R}^m}(\mathbf{0}_m, \delta) \cap (\Omega - x_0) \cap P$ and $\|F(x) - F(x_0)\| < \epsilon$. Hence, F is continuous on the right at x_0 .

In both cases, F is continuous on the right at x_0 .

This completes the proof of the proposition. $\square$

Definition 12.38 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone. $\forall x_1, x_2 \in \mathbb{R}^m$ with $x_1 \leq x_2$, the semi-open rectangle with corners x_1 and x_2 is $r_{x_1, x_2} := \left\{ x \in \mathbb{R}^m \mid x_1 \leq x \leq x_2 \right\}$. The closed rectangle*

with corners x_1 and x_2 is $\bar{r}_{x_1, x_2} := \left\{ x \in \mathbb{R}^m \mid x_1 \leq x \leq x_2 \right\}$. The open rectangle with corners x_1 and x_2 is $r_{x_1, x_2}^\circ := \left\{ x \in \mathbb{R}^m \mid x_1 < x < x_2 \right\}$. $\Omega \subseteq \mathbb{R}^m$ is said to be a region if $\exists N \subseteq \mathbb{N}$ and $(\bar{x}_i)_{i \in N}, (\tilde{x}_i)_{i \in N} \subseteq \Omega$ with $\bar{x}_i \leq \tilde{x}_i$ and $\bar{r}_{\bar{x}_i, \tilde{x}_i} \subseteq \Omega$, $\forall i \in N$, such that $\Omega = \bigcup_{i \in N} \bar{r}_{\bar{x}_i, \tilde{x}_i}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i \in N}$ is pairwise disjoint, $P(\Omega) := \bigcup_{i \in N} r_{\bar{x}_i, \tilde{x}_i}$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, we have $r_{x_1, x_2} \subseteq P(\Omega)$. $P(\Omega)$ is uniquely defined for each region Ω (see below) and is called the principal of Ω .

Clearly, any region $\Omega \subseteq \mathbb{R}^m$ we have $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ and $P(\Omega) \in \mathcal{B}_B(\mathbb{R}^m)$.

Proposition 12.39 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone. Then, the following holds.*

- (i) *If $\Omega \subseteq \mathbb{R}^m$ is a region, then, $P(\Omega) \subseteq \Omega$ is uniquely defined, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$, and $P(\Omega) \in \mathcal{B}_B(\mathbb{R}^m)$.*
- (ii) *If $\Omega_1, \Omega_2 \subseteq \mathbb{R}^m$ are regions with $\Omega_1 \subseteq \Omega_2$, then $P(\Omega_1) \subseteq P(\Omega_2)$.*
- (iii) *If $\Omega_1, \Omega_2, \Omega_1 \cup \Omega_2 \subseteq \mathbb{R}^m$ are regions, then $P(\Omega_1) \cup P(\Omega_2) \subseteq P(\Omega_1 \cup \Omega_2)$.*
- (iv) *If $\Omega_1, \Omega_2, \Omega_1 \cap \Omega_2 \subseteq \mathbb{R}^m$ are regions, then $P(\Omega_1) \cap P(\Omega_2) \supseteq P(\Omega_1 \cap \Omega_2)$.*
- (v) *Let $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle, that is $\exists$ intervals $I_1, \dots, I_m$ such that $\Omega = \prod_{i=1}^m I_i$. $\forall i \in \{1, \dots, m\} =: \bar{J}$, if $I_i = [a_i, b_i] \subset \mathbb{R}$ with $a_i, b_i \in \mathbb{R}$ and $a_i < b_i$, let $\bar{I}_i := (a_i, b_i] \subset \mathbb{R}$; if $I_i = [a_i, a_i] \subset \mathbb{R}$ with $a_i \in \mathbb{R}$, let $\bar{I}_i := \emptyset$; if $I_i = [a_i, b_i] \subset \mathbb{R}$ with $a_i, b_i \in \mathbb{R}$ and $a_i < b_i$, let $\bar{I}_i := (a_i, b_i) \subset \mathbb{R}$; if $I_i = [a_i, \infty) \subset \mathbb{R}$ with $a_i \in \mathbb{R}$, let $\bar{I}_i := (a_i, \infty) \subset \mathbb{R}$; in all other cases, let $\bar{I}_i := I_i$. Then, Ω is a region and $P(\Omega) = \prod_{i=1}^m \bar{I}_i$ is a rectangle.*
- (vi) *$\forall x_1, x_2 \in \mathbb{R}^m$ with $x_1 \leq x_2$, we have $\bar{r}_{x_1, x_2}$ and r_{x_1, x_2} are regions with $P(\bar{r}_{x_1, x_2}) = P(r_{x_1, x_2}) = r_{x_1, x_2}$.*
- (vii) *Let $M \subseteq \mathbb{N}$ and $(\check{x}_j)_{j \in M}, (\hat{x}_j)_{j \in M} \subseteq \mathbb{R}^m$ with $\check{x}_j \leq \hat{x}_j$, $\forall j \in M$. Then, $\Omega := \bigcup_{j \in M} r_{\check{x}_j, \hat{x}_j} \subseteq \mathbb{R}^m$ is a region and $P(\Omega) = \Omega$.*
- (viii) *If $\Omega_1 \subseteq \mathbb{R}^l$ and $\Omega_2 \subseteq \mathbb{R}^{m-l}$ are regions, where $l \in \{1, \dots, m-1\}$, then $\Omega_1 \times \Omega_2 \subseteq \mathbb{R}^m$ is a region with $P(\Omega_1 \times \Omega_2) = P(\Omega_1) \times P(\Omega_2)$.*

Proof (i) Let $j = 1, 2$, and $\exists N_j \subseteq \mathbb{N}$ and $(\bar{x}_{j,i})_{i \in N_j}, (\tilde{x}_{j,i})_{i \in N_j} \subseteq \Omega$ with $\bar{x}_{j,i} \leq \tilde{x}_{j,i}$ and $\bar{r}_{\bar{x}_{j,i}, \tilde{x}_{j,i}} \subseteq \Omega$, $\forall i \in N_j$, such that $\Omega = \bigcup_{i \in N_j} \bar{r}_{\bar{x}_{j,i}, \tilde{x}_{j,i}}$, $(r_{\bar{x}_{j,i}, \tilde{x}_{j,i}})_{i \in N_j}$ is pairwise disjoint, $\Omega_j := \bigcup_{i \in N_j} r_{\bar{x}_{j,i}, \tilde{x}_{j,i}}$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, we have $r_{x_1, x_2} \subseteq \Omega_j$. We will show that $\Omega_1 = \Omega_2$. Then, the result follows.

$\forall i \in N_1$, we have $\bar{x}_{1,i} \leq \tilde{x}_{1,i}$ and $\bar{r}_{\bar{x}_{1,i}, \bar{x}_{1,i}} \subseteq \Omega$. This implies that $r_{\bar{x}_{1,i}, \bar{x}_{1,i}} \subseteq \Omega_2$. Hence, $\Omega_1 = \bigcup_{i \in N_1} r_{\bar{x}_{1,i}, \bar{x}_{1,i}} \subseteq \Omega_2$. By symmetry, we have $\Omega_2 \subseteq \Omega_1$.

(ii) By Ω_1 being a region, $\exists N_1 \subseteq \mathbb{N}$ and $(\bar{x}_{1,i})_{i \in N_1}, (\tilde{x}_{1,i})_{i \in N_1} \subseteq \Omega_1 \subseteq \Omega_2$ with $\bar{x}_{1,i} \leq \tilde{x}_{1,i}$ and $\bar{r}_{\bar{x}_{1,i}, \bar{x}_{1,i}} \subseteq \Omega_1 \subseteq \Omega_2, \forall i \in N_1$, such that $\Omega_1 = \bigcup_{i \in N_1} \bar{r}_{\bar{x}_{1,i}, \bar{x}_{1,i}}$ and $P(\Omega_1) = \bigcup_{i \in N_1} r_{\bar{x}_{1,i}, \bar{x}_{1,i}}$. By Ω_2 being a region, we have $r_{\bar{x}_{1,i}, \bar{x}_{1,i}} \subseteq P(\Omega_2), \forall i \in N_1$. Hence, $P(\Omega_1) \subseteq P(\Omega_2)$.

(iii) By (ii), $P(\Omega_1 \cup \Omega_2) \supseteq P(\Omega_i), i = 1, 2$. Then, the result follows.

(iv) By (ii), $P(\Omega_1 \cap \Omega_2) \subseteq P(\Omega_i), i = 1, 2$. Then, the result follows.

(v) This is straightforward, and is therefore omitted.

(vi) This follows immediately from (v).

(vii) We will first prove the result under the additional assumption that $(r_{\bar{x}_j, \bar{x}_j})_{j \in M} =: (\Omega_j)_{j \in M}$ is pairwise disjoint. $\forall j \in M$, by (vi), Ω_j is a region and $P(\Omega_j) = \Omega_j$. Then, $\exists N_j \subseteq \mathbb{N}$ and $(\bar{x}_{j,i})_{i \in N_j}, (\tilde{x}_{j,i})_{i \in N_j} \subseteq \Omega_j$ with $\bar{x}_{j,i} \leq \tilde{x}_{j,i}, \forall i \in N_j$, such that $\Omega_j = \bigcup_{i \in N_j} \bar{r}_{\bar{x}_{j,i}, \bar{x}_{j,i}}, (r_{\bar{x}_{j,i}, \bar{x}_{j,i}})_{i \in N_j}$ is pairwise disjoint, $P(\Omega_j) = \Omega_j = \bigcup_{i \in N_j} r_{\bar{x}_{j,i}, \bar{x}_{j,i}}$, and $\forall x_1, x_2 \in \Omega_j$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega_j$, we have $r_{x_1, x_2} \subseteq P(\Omega_j) = \Omega_j$. For $\Omega := \bigcup_{j \in M} \Omega_j$, consider $(\bar{x}_{j,i})_{j \in M, i \in N_j}, (\tilde{x}_{j,i})_{j \in M, i \in N_j} \subseteq \Omega$. We have $\bar{x}_{j,i} \leq \tilde{x}_{j,i}, \forall j \in M, \forall i \in N_j, (r_{\bar{x}_{j,i}, \bar{x}_{j,i}})_{j \in M, i \in N_j}$ is pairwise disjoint, $\Omega = \bigcup_{j \in M} \bigcup_{i \in N_j} \bar{r}_{\bar{x}_{j,i}, \bar{x}_{j,i}}, P(\Omega) = \bigcup_{j \in M} \bigcup_{i \in N_j} r_{\bar{x}_{j,i}, \bar{x}_{j,i}} = \bigcup_{j \in M} \Omega_j = \Omega$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, we have $r_{x_1, x_2} \subseteq \Omega = P(\Omega)$. Hence, Ω is a region with $P(\Omega) = \Omega$.

Next, we prove the general case when $(\Omega_j)_{j \in M}$ is not necessarily pairwise disjoint. Without loss of generality, take $M = \mathbb{N}$. Then, $\Omega = \bigcup_{n=1}^{\infty} \left(\Omega_n \setminus \left(\bigcup_{i=1}^{n-1} \Omega_i \right) \right)$. Since for semi-open rectangles A and B , $A \setminus B$ can be expressed as a finite union of pairwise disjoint semi-open rectangles. Then, $\Omega = \bigcup_{n=1}^{\infty} \bar{\Omega}_n$, where $\bar{\Omega}_n = r_{\bar{x}_n, \hat{x}_n}$ for some $\bar{x}_n, \hat{x}_n \in \mathbb{R}^m$ with $\bar{x}_n \leq \hat{x}_n, \forall n \in \mathbb{N}$, and the sets in the union are pairwise disjoint. By the special case, we have Ω is a region and $P(\Omega) = \Omega$.

(viii) By Ω_j being a region, $\exists N_j \subseteq \mathbb{N}$ and $(\bar{x}_{j,i})_{i \in N_j}, (\tilde{x}_{j,i})_{i \in N_j} \subseteq \Omega_j$ with $\bar{x}_{j,i} \leq \tilde{x}_{j,i}$ and $\bar{r}_{\bar{x}_{j,i}, \bar{x}_{j,i}} \subseteq \Omega_j, \forall i \in N_j$, such that $\Omega_j = \bigcup_{i \in N_j} \bar{r}_{\bar{x}_{j,i}, \bar{x}_{j,i}}, (r_{\bar{x}_{j,i}, \bar{x}_{j,i}})_{i \in N_j}$ is pairwise disjoint, $P(\Omega_j) := \bigcup_{i \in N_j} r_{\bar{x}_{j,i}, \bar{x}_{j,i}}$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega_j$, we have $r_{x_1, x_2} \subseteq P(\Omega_j), j = 1, 2$.

Consider $((\bar{x}_{1,i}, \bar{x}_{2,j}))_{i \in N_1, j \in N_2}, ((\tilde{x}_{1,i}, \tilde{x}_{2,j}))_{i \in N_1, j \in N_2} \subseteq \Omega_1 \times \Omega_2$. We have $(\bar{x}_{1,i}, \bar{x}_{2,j}) \leq (\tilde{x}_{1,i}, \tilde{x}_{2,j})$ and $\bar{r}_{(\bar{x}_{1,i}, \bar{x}_{2,j}), (\bar{x}_{1,i}, \bar{x}_{2,j})} = \bar{r}_{\bar{x}_{1,i}, \bar{x}_{1,i}} \times \bar{r}_{\bar{x}_{2,j}, \bar{x}_{2,j}} \subseteq \Omega_1 \times \Omega_2, \forall i \in N_1, \forall j \in N_2$. Furthermore, $\Omega_1 \times \Omega_2 = \bigcup_{i \in N_1} \bigcup_{j \in N_2} \bar{r}_{\bar{x}_{1,i}, \bar{x}_{1,i}} \times \bar{r}_{\bar{x}_{2,j}, \bar{x}_{2,j}} = \bigcup_{i \in N_1} \bigcup_{j \in N_2} \bar{r}_{(\bar{x}_{1,i}, \bar{x}_{2,j}), (\bar{x}_{1,i}, \bar{x}_{2,j})}, (r_{(\bar{x}_{1,i}, \bar{x}_{2,j}), (\bar{x}_{1,i}, \bar{x}_{2,j})})_{\substack{i \in N_1 \\ j \in N_2}} = (r_{\bar{x}_{1,i}, \bar{x}_{1,i}} \times r_{\bar{x}_{2,j}, \bar{x}_{2,j}})_{\substack{i \in N_1 \\ j \in N_2}}$ are pairwise disjoint. $\bigcup_{i \in N_1} \bigcup_{j \in N_2} r_{\bar{x}_{1,i}, \bar{x}_{1,i}} \times r_{\bar{x}_{2,j}, \bar{x}_{2,j}} = P(\Omega_1) \times P(\Omega_2) = \bigcup_{i \in N_1} \bigcup_{j \in N_2} r_{(\bar{x}_{1,i}, \bar{x}_{2,j}), (\bar{x}_{1,i}, \bar{x}_{2,j})}. \quad \forall (\bar{x}, \bar{y}),$

$(\tilde{x}, \tilde{y}) \in \Omega_1 \times \Omega_2$ with $(\bar{x}, \bar{y}) \leq (\tilde{x}, \tilde{y})$ and $\bar{\Gamma}_{(\bar{x}, \bar{y}), (\tilde{x}, \tilde{y})} \subseteq \Omega_1 \times \Omega_2$, we have $\bar{x} \leq \tilde{x}$ and $\bar{y} \leq \tilde{y}$, $\bar{\Gamma}_{(\bar{x}, \bar{y}), (\tilde{x}, \tilde{y})} = \bar{\Gamma}_{\bar{x}, \bar{x}} \times \bar{\Gamma}_{\bar{y}, \tilde{y}} \subseteq \Omega_1 \times \Omega_2$, $\bar{\Gamma}_{\bar{x}, \bar{x}} \subseteq \Omega_1$, $\bar{\Gamma}_{\bar{y}, \tilde{y}} \subseteq \Omega_2$, which further implies that $r_{\bar{x}, \bar{x}} \subseteq P(\Omega_1)$, $r_{\bar{y}, \tilde{y}} \subseteq P(\Omega_2)$, and $r_{(\bar{x}, \bar{y}), (\tilde{x}, \tilde{y})} = r_{\bar{x}, \bar{x}} \times r_{\bar{y}, \tilde{y}} \subseteq P(\Omega_1) \times P(\Omega_2)$. Hence, $\Omega_1 \times \Omega_2$ is a region with $P(\Omega_1 \times \Omega_2) = P(\Omega_1) \times P(\Omega_2)$.

This completes the proof of the proposition. $\square$

Proposition 12.40 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{O}_{\mathbb{R}^m}$ with the subset topology $\mathcal{O}$. Then, Ω is a region with $P(\Omega) = \Omega$.*

Proof By $\mathbb{R}$ being second countable and Proposition 3.28, $\mathbb{R}^m$ is second countable. By Proposition 4.4, $\mathbb{R}^m$ is separable, which implies that, by Proposition 4.38, $(\Omega, \mathcal{O})$ is separable and second countable. Let $D \subseteq \Omega$ be a countable dense subset and define $\mathcal{B} := \{r_{x-\delta \mathbf{1}_m, x+\delta \mathbf{1}_m}^\circ \subseteq \Omega \mid x \in D, \delta \in \mathbb{Q}, \delta > 0, \bar{\Gamma}_{x-\delta \mathbf{1}_m, x+\delta \mathbf{1}_m} \subseteq \Omega\}$. It is easy to show, by Definition 3.17, that $\mathcal{B}$ is a countable basis for $\mathcal{O}$. Then, $\Omega = \bigcup_{B \in \mathcal{B}} B =: \bigcup_{i \in N} B_i = \bigcup_{i \in N} \bar{\Gamma}_{x_i - \delta_i \mathbf{1}_m, x_i + \delta_i \mathbf{1}_m} = \bigcup_{i \in N} r_{x_i - \delta_i \mathbf{1}_m, x_i + \delta_i \mathbf{1}_m}$, where $N \subseteq \mathbb{N}$. By (vii) of Proposition 12.39, Ω is a region with $P(\Omega) = \Omega$. This completes the proof of the proposition. $\square$

Definition 12.41 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a region, and $F : \Omega \rightarrow \mathcal{Y}$. $\forall x_1, x_2 \in \mathbb{R}^m$ with $x_1 \leq x_2$, $\forall J \subseteq \bar{J} := \{1, \dots, m\}$, the vertex of r_{x_1, x_2} with any of J coordinates equal to that of x_1 and the rest of coordinates equal to that of x_2 is $x_J \in \bar{\Gamma}_{x_1, x_2} \subseteq \mathbb{R}^m$ such that $\pi_i(x_J) = \begin{cases} \pi_i(x_1) & \forall i \in J \\ \pi_i(x_2) & \forall i \in \bar{J} \setminus J \end{cases}$. We define the set of vertexes of r_{x_1, x_2} with i coordinates equal to that of x_1 to be $\text{VRect}_{i, x_1, x_2} := \{x_J \in \mathbb{R}^m \mid J \subseteq \bar{J}, \text{card}(J) = i\}$, if $x_1 < x_2$. When $\bar{\Gamma}_{x_1, x_2} \subseteq \Omega$, the increment of F on r_{x_1, x_2} is $\Delta_F(r_{x_1, x_2}) := \begin{cases} \sum_{i=0}^m \sum_{\bar{x} \in \text{VRect}_{i, x_1, x_2}} (-1)^i F(\bar{x}) & x_1 < x_2 \\ \vartheta_{\mathcal{Y}} & x_1 \not\leq x_2 \end{cases} = \sum_{J \subseteq \bar{J}} (-1)^{\text{card}(J)} F(x_J)$. The total variation of F on the semi-open rectangle r_{x_1, x_2} is*

$$T_F(r_{x_1, x_2}) := \sup_{\substack{n \in \mathbb{Z}_+, \hat{x}_1, \hat{x}_1, \dots, \hat{x}_n, \hat{x}_n \in \bar{\Gamma}_{x_1, x_2}, \hat{x}_i \leq \hat{x}_{i+1}, \forall i \in \{1, \dots, n\} \\ \bigcup_{i=1}^n r_{\hat{x}_i, \hat{x}_{i+1}} = r_{x_1, x_2}, r_{\hat{x}_i, \hat{x}_i} \cap r_{\hat{x}_j, \hat{x}_j} = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n \|\Delta_F(r_{\hat{x}_i, \hat{x}_{i+1}})\|$$

where $T_F(r_{x_1, x_2}) \in [0, \infty] \subset \mathbb{R}_e$. F is said to be of locally bounded variation if

(i) F is continuous on the right;

(ii) $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{\Gamma}_{x_1, x_2} \subseteq \Omega$, we have $T_F(r_{x_1, x_2}) < \infty$;

(iii) $\forall x_1, x_2 \in \Omega$ with $x_1 < x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, $\bar{G} : \bar{r}_{x_1, x_2} \rightarrow [0, \infty) \subset \mathbb{R}$ defined by $\bar{G}(x) = T_F(r_{x_1, x})$, $\forall x \in \bar{r}_{x_1, x_2}$, is continuous on the right.

The total variation of F is defined by $T_F := \sum_{i \in N} T_F(r_{\bar{x}_i, \bar{x}_i})$ if the sum is independent of the choice of $(\bar{x}_i)_{i \in N}$ and $(\bar{x}_i)_{i \in N}$ as prescribed in Definition 12.38. F is said to be of bounded variation if F is of locally bounded variation and $T_F < \infty$.

Definition 12.42 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\mathcal{Y}$ be a normed linear space, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a region with the subset topology, μ be a σ -finite $\mathcal{Y}$ -valued measure on $(P(\Omega), \mathcal{B}_B(P(\Omega)))$, and $F : \Omega \rightarrow \mathcal{Y}$. F is said to be a cumulative distribution function of μ if $\forall x_1, x_2 \in \Omega$ with $\bar{r}_{x_1, x_2} \subseteq \Omega$ and $x_1 \leq x_2$, we have $\mu(r_{x_1, x_2}) = \Delta_F(r_{x_1, x_2})$. μ is said to be locally finite if $\mathcal{P} \circ \mu(K) < \infty$, $\forall$ compact $K \subseteq P(\Omega)$.

Note that, for σ -finite $\mathcal{Y}$ -valued measures, the cumulative distribution function is not unique and may not exist. For finite $\mathcal{Y}$ -valued measures, the function $F : \Omega \rightarrow \mathcal{Y}$ defined by $F(x) = \mu(P(\Omega) \cap \{\bar{x} \in \mathbb{R}^m \mid \bar{x} \leq x\})$, $\forall x \in \Omega$, is a cumulative distribution function of μ .

Proposition 12.43 Let $m \in \mathbb{N}$, $\forall i \in \{1, \dots, m\}$, X_i be a set, $\mathcal{C}_i$ be a semi-algebra on X_i , $\mathcal{B}_i$ be the σ -algebra on X_i generated by $\mathcal{C}_i$, $\mathcal{C} := \{\prod_{i=1}^m C_i \subseteq \prod_{i=1}^m X_i \mid C_i \in \mathcal{C}_i, i = 1, \dots, m\}$, $\mathcal{B}$ be the σ -algebra on $X := \prod_{i=1}^m X_i$ generated by $\mathcal{C}$, and $\mathcal{M} := \{\prod_{i=1}^m B_i \subseteq X \mid B_i \in \mathcal{B}_i, i = 1, \dots, m\}$. Then, $\mathcal{M} \subseteq \mathcal{B}$ and $\mathcal{B}$ is the σ -algebra on X generated by $\mathcal{M}$.

Proof First, we will prove the following intermediate result.

Claim 12.43.1 $\forall C_i \in \mathcal{C}_i, i = 1, \dots, m-1, \forall B_m \in \mathcal{B}_m$, we have $(\prod_{i=1}^{m-1} C_i) \times B_m \in \mathcal{B}$.

Proof of claim: Fix any $C_i \in \mathcal{C}_i, i = 1, \dots, m-1$. Define $\mathcal{F} := \{E \subseteq X_m \mid (\prod_{i=1}^{m-1} C_i) \times E \in \mathcal{B}\}$. By the definition of $\mathcal{B}$, we have $\mathcal{C}_m \subseteq \mathcal{F}$. Clearly, $\emptyset \in \mathcal{C}_m \subseteq \mathcal{F}$. Then, by $\mathcal{C}_m$ being a semialgebra on X_m , we have $X_m = X_m \setminus \emptyset = \bigcup_{i=1}^n C_{m,i}$, where $n \in \mathbb{Z}_+$ and $(C_{m,i})_{i=1}^n \subseteq \mathcal{C}_m$ is pairwise disjoint. Then, $(\prod_{i=1}^{m-1} C_i) \times C_{m,i} \in \mathcal{B}, i = 1, \dots, n$, and hence $(\prod_{i=1}^{m-1} C_i) \times X_m = \bigcup_{i=1}^n ((\prod_{i=1}^{m-1} C_i) \times C_{m,i}) \in \mathcal{B}$. This implies that $X_m \in \mathcal{F}$. $\forall E \in \mathcal{F}$, we have $(\prod_{i=1}^{m-1} C_i) \times E \in \mathcal{B}$. Then, by $\mathcal{B}$ being a σ -algebra and $(\prod_{i=1}^{m-1} C_i) \times X_m \in \mathcal{B}$, we have $(\prod_{i=1}^{m-1} C_i) \times (X_m \setminus E) = ((\prod_{i=1}^{m-1} C_i) \times X_m) \setminus ((\prod_{i=1}^{m-1} C_i) \times E) \in \mathcal{B}$. Thus, $X_m \setminus E \in \mathcal{F}$. $\forall (E_j)_{j=1}^\infty \in \mathcal{F}$, we have $(\prod_{i=1}^{m-1} C_i) \times E_j \in \mathcal{B}, \forall j \in \mathbb{N}$. Then, by $\mathcal{B}$ being a σ -algebra, we have $(\prod_{i=1}^{m-1} C_i) \times (\bigcup_{j=1}^\infty E_j) = \bigcup_{j=1}^\infty ((\prod_{i=1}^{m-1} C_i) \times E_j) \in \mathcal{B}$. Thus, $\bigcup_{j=1}^\infty E_j \in \mathcal{F}$. This shows that $\mathcal{F}$ is a σ -algebra. Hence, $\mathcal{B}_m \subseteq \mathcal{F}$ since $\mathcal{B}_m$ is the σ -algebra generated by $\mathcal{C}_m$. This completes the proof of the claim. $\square$

Fix any $B_i \in \mathcal{B}_i, i = 1, \dots, m$. Clearly, $\emptyset \in \mathcal{C}_i, i = 1, \dots, m-1$. Fix any $i = 1, \dots, m-1$. Since $\mathcal{C}_i$ is a semialgebra on X_i , we have

$X_i = X_i \setminus \emptyset = \bigcup_{j=1}^{n_i} C_{i,j}$, where $n_i \in \mathbb{Z}_+$ and $(C_{i,j})_{j=1}^{n_i} \subseteq \mathcal{C}_i$ is pairwise disjoint. By Claim 12.43.1, $(\prod_{i=1}^{m-1} C_{i,j_i}) \times B_m \in \mathcal{B}$, $j_i = 1, \dots, n_i$, $i = 1, \dots, m-1$. Then, $(\prod_{i=1}^{m-1} X_i) \times B_m = (\bigcup_{j_1=1}^{n_1} \dots \bigcup_{j_{m-1}=1}^{n_{m-1}} \prod_{i=1}^{m-1} C_{i,j_i}) \times B_m = \bigcup_{j_1=1}^{n_1} \dots \bigcup_{j_{m-1}=1}^{n_{m-1}} ((\prod_{i=1}^{m-1} C_{i,j_i}) \times B_m) \in \mathcal{B}$. By symmetry, we have $(\prod_{i=1}^{k-1} X_i) \times B_k \times (\prod_{i=k+1}^m X_i) \in \mathcal{B}$, $k = 1, \dots, m$. Then, $\prod_{i=1}^m B_i = \bigcap_{k=1}^m ((\prod_{i=1}^{k-1} X_i) \times B_k \times (\prod_{i=k+1}^m X_i)) \in \mathcal{B}$. By the arbitrariness of B_i 's, we have $\mathcal{C} \subseteq \mathcal{M} \subseteq \mathcal{B}$. Since $\mathcal{B}$ is the σ -algebra on X generated by $\mathcal{C}$, then $\mathcal{B}$ is the σ -algebra on X generated by $\mathcal{M}$. $\square$

Proposition 12.44 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\bar{x}, \tilde{x} \in \mathbb{R}^m$ with $\bar{x} \leq \tilde{x}$, $X := r_{\bar{x}, \tilde{x}} \subseteq \mathbb{R}^m$ with subset topology $\mathcal{O}$, $\mathcal{X} := (X, \mathcal{O})$ be the topological space, $\mathcal{C} := \{r_{x_1, x_2} \mid \bar{x} \leq x_1 \leq x_2 \leq \tilde{x}\}$, and $\mathcal{B}$ be the σ -algebra on X generated by $\mathcal{C}$. Then, $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$.*

Proof Let $\bar{x} = (\bar{x}_1, \dots, \bar{x}_m)$ and $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_m)$. $\forall i \in \{1, \dots, m\}$, let $I_i := (\bar{x}_i, \tilde{x}_i] \subset \mathbb{R}$ with the subset topology $\mathcal{O}_i$, and $\mathcal{C}_i := \{(a, b] \subset \mathbb{R} \mid \bar{x}_i \leq a \leq b \leq \tilde{x}_i\}$. Clearly, $\mathcal{C}_i$ is a semialgebra on I_i and the σ -algebra $\mathcal{B}_i$ on I_i generated by $\mathcal{C}_i$ is equal to $\mathcal{B}_B(I_i)$. Then, $\mathcal{C} = \{\prod_{i=1}^m C_i \mid C_i \in \mathcal{C}_i, i = 1, \dots, m\}$. Then, by Proposition 12.43, $\mathcal{B}$ equals to the σ -algebra generated by $\mathcal{M} := \{\prod_{i=1}^m B_i \mid B_i \in \mathcal{B}_i, i = 1, \dots, m\}$. By Propositions 12.22 and 12.32, we have $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$. This completes the proof of the proposition. $\square$

Proposition 12.45 *Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $(X_i)_{i=1}^\infty \subseteq \mathcal{B}_B(\mathcal{X})$ be pairwise disjoint with $X = \bigcup_{i=1}^\infty X_i$, $\mathcal{X}_i := (X_i, \mathcal{O}_i)$ be the topological subspace of $\mathcal{X}$, $\forall i \in \mathbb{N}$, and $\mathcal{B} := \{E \subseteq X \mid E \cap X_i \in \mathcal{B}_B(\mathcal{X}_i), \forall i \in \mathbb{N}\}$. Then, $\mathcal{B} = \mathcal{B}_B(\mathcal{X})$.*

Proof $\forall E \in \mathcal{B}$, $\forall i \in \mathbb{N}$, $E \cap X_i \in \mathcal{B}_B(\mathcal{X}_i)$. By Proposition 11.25, $\mathcal{B}_B(\mathcal{X}_i) = \{C \subseteq X_i \mid C \in \mathcal{B}_B(\mathcal{X})\}$, $\forall i \in \mathbb{N}$. Then, $E \cap X_i \in \mathcal{B}_B(\mathcal{X})$, $\forall i \in \mathbb{N}$. Then, $E = \bigcup_{i=1}^\infty (E \cap X_i) \in \mathcal{B}_B(\mathcal{X})$. Hence, $\mathcal{B} \subseteq \mathcal{B}_B(\mathcal{X})$. On the other hand, $\forall E \in \mathcal{B}_B(\mathcal{X})$, $E \cap X_i \in \mathcal{B}_B(\mathcal{X})$, $\forall i \in \mathbb{N}$. By Proposition 11.25, $E \cap X_i \in \mathcal{B}_B(\mathcal{X}_i)$, $\forall i \in \mathbb{N}$. Hence, $E \in \mathcal{B}$. This implies that $\mathcal{B}_B(\mathcal{X}) \subseteq \mathcal{B}$. This completes the proof of the proposition. $\square$

Proposition 12.46 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \subseteq \mathbb{R}^m$, $\mathcal{Y}$ be a normed linear space, and $F : \Omega \rightarrow \mathcal{Y}$. Then, $\forall \bar{x}, \tilde{x} \in \Omega$ with $\bar{x} \leq \tilde{x}$ and $r_{\bar{x}, \tilde{x}} \subseteq \Omega$, $\forall n \in \mathbb{Z}_+$, $\forall (\tilde{x}_i)_{i=1}^n, (\hat{x}_i)_{i=1}^n \subseteq \Omega$ with $\tilde{x}_i \leq \hat{x}_i$, $r_{\tilde{x}_i, \hat{x}_i} \subseteq \Omega$, $\forall i \in \{1, \dots, n\}$, $r_{\bar{x}, \tilde{x}} = \bigcup_{i=1}^n r_{\tilde{x}_i, \hat{x}_i}$, and the sets in the union being pairwise disjoint, we have $\Delta_F(r_{\bar{x}, \tilde{x}}) = \sum_{i=1}^n \Delta_F(r_{\tilde{x}_i, \hat{x}_i})$.*

Proof We will prove the result by mathematical induction on n . 1° $n = 0$. Clearly, $r_{\bar{x}, \tilde{x}} = \emptyset$ and $\bar{x} \not\leq \tilde{x}$. Then, $\Delta_F(r_{\bar{x}, \tilde{x}}) = \emptyset_{\mathcal{Y}} = \sum_{i=1}^0 \Delta_F(r_{\tilde{x}_i, \hat{x}_i})$. This case is proved.

$n = 1$. This case is trivial and proved.

$n = 2$. Without loss of generality, assume that $\bar{x} \leq \tilde{x}$ and $\tilde{x}_i \leq \hat{x}_i$, $i = 1, 2$. Otherwise, the result follows from $n = 1$ or $n = 0$ cases, since $\Delta_F(r_{x_1, x_2}) = \vartheta_y$ if $x_1 \leq x_2$ and $x_1 \not\leq x_2$. Then, we have $(\tilde{x}_i)_{i=1}^n, (\hat{x}_i)_{i=1}^n \subseteq \bar{r}_{\bar{x}, \tilde{x}}$. Without loss of generality, assume $\tilde{x} \in r_{\tilde{x}_2, \hat{x}_2}$. Then, $\tilde{x} = \hat{x}_2$ and $\bar{x} \leq \tilde{x}_2 \leq \hat{x}_2 = \tilde{x}$. Clearly, $\bar{x} \neq \tilde{x}_2$ since $r_{\bar{x}, \tilde{x}} = \bigcup_{i=1}^2 r_{\tilde{x}_i, \hat{x}_i}$ and the sets in the union are nonempty and pairwise disjoint. Then, $\exists j_0 \in \{1, \dots, m\}$ such that $\pi_{j_0}(\tilde{x}) > \pi_{j_0}(\tilde{x}_2) > \pi_{j_0}(\bar{x})$. Without loss of generality, assume that $j_0 = m$. Define $x_C \in \bar{r}_{\bar{x}, \tilde{x}}$ be such that $\pi_j(x_C) = \begin{cases} \pi_j(\bar{x}) & j = 1, \dots, m-1 \\ \pi_m(\tilde{x}_2) & j = m \end{cases}$ and $x_D \in \bar{r}_{\bar{x}, \tilde{x}}$ be such that $\pi_j(x_D) = \begin{cases} \pi_j(\tilde{x}) & j = 1, \dots, m-1 \\ \pi_m(\tilde{x}_2) & j = m \end{cases}$. Clearly, we have $x_D \in r_{\tilde{x}_1, \hat{x}_1} \setminus r_{\tilde{x}_2, \hat{x}_2}$. Then, $\tilde{x}_1 = \bar{x}$, $\hat{x}_1 = x_D$, $\tilde{x}_2 = x_C$, and $\hat{x}_2 = \tilde{x}$. By Definition 12.41, we have

$$\begin{aligned}
\sum_{i=1}^2 \Delta_F(r_{\tilde{x}_i, \hat{x}_i}) &= \sum_{i=0}^m \sum_{s \in \text{VRect}_{i, \bar{x}, x_D}} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \text{VRect}_{i, x_C, \tilde{x}}} (-1)^i F(s) \\
&= \sum_{i=0}^m \sum_{\substack{s \in \text{VRect}_{i, \bar{x}, x_D} \\ \pi_m(s) = \pi_m(\bar{x})}} (-1)^i F(s) + \sum_{i=0}^m \sum_{\substack{s \in \text{VRect}_{i, \bar{x}, x_D} \\ \pi_m(s) = \pi_m(x_D)}} (-1)^i F(s) \\
&\quad + \sum_{i=0}^m \sum_{\substack{s \in \text{VRect}_{i, x_C, \tilde{x}} \\ \pi_m(s) = \pi_m(\tilde{x})}} (-1)^i F(s) + \sum_{i=0}^m \sum_{\substack{s \in \text{VRect}_{i, x_C, \tilde{x}} \\ \pi_m(s) = \pi_m(x_C)}} (-1)^i F(s) \\
&= \sum_{i=0}^m \sum_{\substack{s \in \text{VRect}_{i, \bar{x}, \tilde{x}} \\ \pi_m(s) = \pi_m(\bar{x})}} (-1)^i F(s) + \sum_{i=0}^m \sum_{\substack{s \in \text{VRect}_{i+1, x_C, \tilde{x}} \\ \pi_m(s) = \pi_m(x_C)}} (-1)^i F(s) \\
&\quad + \sum_{i=0}^m \sum_{\substack{s \in \text{VRect}_{i, \bar{x}, \tilde{x}} \\ \pi_m(s) = \pi_m(\tilde{x})}} (-1)^i F(s) + \sum_{i=0}^m \sum_{\substack{s \in \text{VRect}_{i, x_C, \tilde{x}} \\ \pi_m(s) = \pi_m(x_C)}} (-1)^i F(s) \\
&= \sum_{i=0}^m \sum_{s \in \text{VRect}_{i, \bar{x}, \tilde{x}}} (-1)^i F(s) = \Delta_F(r_{\bar{x}, \tilde{x}})
\end{aligned}$$

where the first two equalities follow from Definition 12.41; the third equality follows from the fact that, $\forall i \in \{0, \dots, m\}$, $\forall s \in \text{VRect}_{i, \bar{x}, x_D}$ with $\pi_m(s) = \pi_m(x_D)$ if, and only if, $s \in \text{VRect}_{i+1, x_C, \tilde{x}}$ with $\pi_m(s) = \pi_m(x_C)$, $\forall s \in \text{VRect}_{i, \bar{x}, x_D}$ with $\pi_m(s) = \pi_m(\bar{x})$ if, and only if, $s \in \text{VRect}_{i, \bar{x}, \tilde{x}}$ with $\pi_m(s) = \pi_m(\tilde{x})$, and $\forall s \in \text{VRect}_{i, x_C, \tilde{x}}$ with $\pi_m(s) = \pi_m(\tilde{x})$ if, and only if, $s \in \text{VRect}_{i, \bar{x}, \tilde{x}}$ with $\pi_m(s) = \pi_m(\tilde{x})$; and the fourth equality follows from the fact that the second and fourth summations cancel each other out. This case is proved.

2° Assume that the result holds for $n \leq k-1$, where $k \in \mathbb{N}$ and $k \geq 3$.

3° Consider the case when $n = k \in \{3, 4, \dots\}$. Without loss of generality, assume that $\bar{x} \leq \tilde{x}$ and $\tilde{x}_i \leq \hat{x}_i$, $i = 1, \dots, n$. Otherwise, the result follows from 1° or 2°, since $\Delta_F(r_{x_1, x_2}) = \vartheta_{\mathcal{Y}}$ if $x_1 \leq x_2$ and $x_1 \not\leq x_2$. Then, we have $(\tilde{x}_i)_{i=1}^n, (\hat{x}_i)_{i=1}^n \subseteq \bar{r}_{\bar{x}, \tilde{x}}$. Without loss of generality, assume $\tilde{x} \in r_{\tilde{x}_n, \hat{x}_n}$. Then, $\tilde{x} = \hat{x}_n$ and $\bar{x} \leq \tilde{x}_n \leq \hat{x}_n = \tilde{x}$. Clearly, $\bar{x} \neq \tilde{x}_n$ since $r_{\bar{x}, \tilde{x}} = \bigcup_{i=1}^n r_{\tilde{x}_i, \hat{x}_i}$ and the sets in the union are nonempty and pairwise disjoint. Then, $\exists j_0 \in \{1, \dots, m\}$ such that $\pi_{j_0}(\tilde{x}) > \pi_{j_0}(\tilde{x}_n) > \pi_{j_0}(\bar{x})$. Without loss of generality, assume that $j_0 = m$. Define $x_C \in \bar{r}_{\bar{x}, \tilde{x}}$ be such that $\pi_j(x_C) = \begin{cases} \pi_j(\bar{x}) & j = 1, \dots, m-1 \\ \pi_m(\tilde{x}_n) & j = m \end{cases}$ and $x_D \in \bar{r}_{\bar{x}, \tilde{x}}$ be such that $\pi_j(x_D) = \begin{cases} \pi_j(\tilde{x}) & j = 1, \dots, m-1 \\ \pi_m(\tilde{x}_n) & j = m \end{cases}$. Then, $r_{\bar{x}, \tilde{x}} = r_{\bar{x}, x_D} \cup r_{x_C, \tilde{x}} = (\bigcup_{i=1}^n (r_{\bar{x}, x_D} \cap r_{\tilde{x}_i, \hat{x}_i})) \cup (\bigcup_{i=1}^n (r_{x_C, \tilde{x}} \cap r_{\tilde{x}_i, \hat{x}_i}))$. Clearly, $r_{\bar{x}, x_D} \cap r_{\tilde{x}_n, \hat{x}_n} = \emptyset$. Note that $x_D \in r_{\bar{x}, \tilde{x}} \setminus r_{\tilde{x}_n, \hat{x}_n}$. Without loss of generality, assume that $x_D \in r_{\tilde{x}_{n-1}, \hat{x}_{n-1}}$. Then, $\bar{x} \leq \tilde{x}_{n-1} \leq x_D \leq \hat{x}_{n-1} \leq \tilde{x}$. Then, we must have $x_D = \hat{x}_{n-1}$ since $r_{\tilde{x}_{n-1}, \hat{x}_{n-1}} \cap r_{\tilde{x}_n, \hat{x}_n} = \emptyset$. Thus, $r_{x_C, \tilde{x}} \cap r_{\tilde{x}_{n-1}, \hat{x}_{n-1}} = \emptyset$. Then, $r_{\bar{x}, x_D} = \bigcup_{i=1}^{n-1} (r_{\bar{x}, x_D} \cap r_{\tilde{x}_i, \hat{x}_i})$ and $r_{x_C, \tilde{x}} = \bigcup_{i=1}^{n-1} (r_{x_C, \tilde{x}} \cap r_{\tilde{x}_i, \hat{x}_i})$. By the inductive assumption, we have $\Delta_F(r_{\bar{x}, \tilde{x}}) = \Delta_F(r_{\bar{x}, x_D}) + \Delta_F(r_{x_C, \tilde{x}}) = \sum_{i=1}^{n-1} \Delta_F(r_{\bar{x}, x_D} \cap r_{\tilde{x}_i, \hat{x}_i}) + \sum_{i=1}^{n-1} \Delta_F(r_{x_C, \tilde{x}} \cap r_{\tilde{x}_i, \hat{x}_i}) = \sum_{i=1}^{n-2} (\Delta_F(r_{\bar{x}, x_D} \cap r_{\tilde{x}_i, \hat{x}_i}) + \Delta_F(r_{x_C, \tilde{x}} \cap r_{\tilde{x}_i, \hat{x}_i})) + \Delta_F(r_{\tilde{x}_{n-1}, \hat{x}_{n-1}}) + \Delta_F(r_{\tilde{x}_n, \hat{x}_n}) = \sum_{i=1}^n \Delta_F(r_{\tilde{x}_i, \hat{x}_i})$. This case is proved.

Therefore, the result holds for any $n \in \mathbb{N}$. This completes the proof of the proposition. $\square$

Proposition 12.47 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\mathcal{Y}$ be a normed linear space, $\Omega \subseteq \mathbb{R}^m$, and $F : \Omega \rightarrow \mathcal{Y}$. Assume that F satisfies (i) and (ii) of Definition 12.41. Then, $\forall \bar{x}, \tilde{x} \in \Omega$ with $\bar{x} \leq \tilde{x}$ and $\bar{r}_{\bar{x}, \tilde{x}} \subseteq \Omega$, $\forall n \in \mathbb{Z}_+$, $\forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{r}_{\bar{x}, \tilde{x}}$ with $\bar{x}_i \leq \tilde{x}_i$, $\forall i \in \{1, \dots, n\}$, $r_{\bar{x}, \tilde{x}} = \bigcup_{i=1}^n r_{\bar{x}_i, \tilde{x}_i}$, and the sets in the union being pairwise disjoint, we have $T_F(r_{\bar{x}, \tilde{x}}) = \sum_{i=1}^n T_F(r_{\bar{x}_i, \tilde{x}_i})$.*

Proof $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall i \in \{1, \dots, n\}$, $\exists n_i \in \mathbb{Z}_+$, $\exists (\tilde{x}_{i,j})_{j=1}^{n_i}, (\hat{x}_{i,j})_{j=1}^{n_i} \subseteq \bar{r}_{\bar{x}_i, \tilde{x}_i}$ with $\tilde{x}_{i,j} \leq \hat{x}_{i,j}$, $\forall j \in \{1, \dots, n_i\}$, $r_{\bar{x}_i, \tilde{x}_i} = \bigcup_{j=1}^{n_i} r_{\tilde{x}_{i,j}, \hat{x}_{i,j}}$, and sets in the union being pairwise disjoint, such that $\sum_{j=1}^{n_i} \|\Delta_F(r_{\tilde{x}_{i,j}, \hat{x}_{i,j}})\| > T_F(r_{\bar{x}_i, \tilde{x}_i}) - \epsilon/(n+1)$. Then, $\sum_{i=1}^n T_F(r_{\bar{x}_i, \tilde{x}_i}) - \epsilon < \sum_{i=1}^n \sum_{j=1}^{n_i} \|\Delta_F(r_{\tilde{x}_{i,j}, \hat{x}_{i,j}})\| \leq T_F(r_{\bar{x}, \tilde{x}})$. By the arbitrariness of ϵ , we have $\sum_{i=1}^n T_F(r_{\bar{x}_i, \tilde{x}_i}) \leq T_F(r_{\bar{x}, \tilde{x}})$.

On the other hand, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \bar{n} \in \mathbb{Z}_+$, $\exists (\tilde{x}_i)_{i=1}^{\bar{n}}, (\hat{x}_i)_{i=1}^{\bar{n}} \subseteq \bar{r}_{\bar{x}, \tilde{x}}$ with $\tilde{x}_i \leq \hat{x}_i$, $\forall i \in \{1, \dots, \bar{n}\}$, $r_{\bar{x}, \tilde{x}} = \bigcup_{i=1}^{\bar{n}} r_{\tilde{x}_i, \hat{x}_i}$, and the sets in the union being pairwise disjoint, such that $\sum_{i=1}^{\bar{n}} \|\Delta_F(r_{\tilde{x}_i, \hat{x}_i})\| > T_F(r_{\bar{x}, \tilde{x}}) - \epsilon$. Then, $T_F(r_{\bar{x}, \tilde{x}}) - \epsilon < \sum_{i=1}^{\bar{n}} \|\Delta_F(r_{\tilde{x}_i, \hat{x}_i})\| = \sum_{i=1}^{\bar{n}} \left\| \Delta_F\left(\bigcup_{j=1}^n (r_{\tilde{x}_i, \hat{x}_i} \cap r_{\tilde{x}_j, \hat{x}_j})\right) \right\| = \sum_{i=1}^{\bar{n}} \left\| \sum_{j=1}^n \Delta_F(r_{\tilde{x}_i, \hat{x}_i} \cap r_{\tilde{x}_j, \hat{x}_j}) \right\| \leq \sum_{i=1}^{\bar{n}} \sum_{j=1}^n \|\Delta_F(r_{\tilde{x}_i, \hat{x}_i} \cap r_{\tilde{x}_j, \hat{x}_j})\| =$

$\sum_{j=1}^n \sum_{i=1}^{\bar{n}} \|\Delta_F(r_{\tilde{x}_i, \tilde{x}_i} \cap r_{\tilde{x}_j, \tilde{x}_j})\| \leq \sum_{j=1}^n T_F(r_{\tilde{x}_j, \tilde{x}_j})$, where the second equality follows from Proposition 12.46. By the arbitrariness of ϵ , we have $T_F(r_{\tilde{x}, \tilde{x}}) \leq \sum_{i=1}^n T_F(r_{\tilde{x}_i, \tilde{x}_i})$. This completes the proof of the proposition. $\square$

Proposition 12.48 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a region with the subset topology $\mathcal{O}$, and $F : \Omega \rightarrow \mathbb{R}$ be of locally bounded variation. Assume that $\Delta_F(r_{x_1, x_2}) \geq 0$, $\forall x_1, x_2 \in \Omega$ with $\bar{r}_{x_1, x_2} \subseteq \Omega$ and $x_1 \leq x_2$. Then, there exists a unique σ -finite measure $\hat{\mu}$ on the measurable space $(P(\Omega), \mathcal{B}_B(P(\Omega)))$ such that F is a cumulative distribution function of $\hat{\mu}$. Furthermore, $T_F(r_{x_1, x_2}) = \Delta_F(r_{x_1, x_2}) = \hat{\mu}(r_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, and $T_F = \hat{\mu}(P(\Omega)) \in [0, \infty] \subset \mathbb{R}_e$.*

Proof Since Ω is a region, then $\exists N \subseteq \mathbb{N}$, $\exists (\bar{x}_i)_{i \in N}, (\tilde{x}_i)_{i \in N} \subseteq \Omega$ with $\bar{x}_i \leq \tilde{x}_i$ and $\bar{r}_{\bar{x}_i, \tilde{x}_i} \subseteq \Omega$, $\forall i \in N$, such that $\Omega = \bigcup_{i \in N} \bar{r}_{\bar{x}_i, \tilde{x}_i}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i \in N}$ is pairwise disjoint, $P(\Omega) = \bigcup_{i \in N} r_{\bar{x}_i, \tilde{x}_i}$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, we have $r_{x_1, x_2} \subseteq P(\Omega)$.

Claim 12.48.1 $\forall \bar{x}, \tilde{x} \in \Omega$ with $\bar{r}_{\bar{x}, \tilde{x}} \subseteq \Omega$ and $\bar{x} \leq \tilde{x}$, let $\mathcal{C} := \left\{ r_{x_1, x_2} \mid \bar{x} \leq x_1 \leq x_2 \leq \tilde{x} \right\}$, which is the collection of semi-open rectangles contained in $r_{\bar{x}, \tilde{x}}$. Clearly, $\mathcal{C}$ is semialgebra on $r_{\bar{x}, \tilde{x}}$. Define $\mu : \mathcal{C} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\mu(r_{x_1, x_2}) = \Delta_F(r_{x_1, x_2})$, $\forall r_{x_1, x_2} \in \mathcal{C}$. Then, μ admits a unique extension to a measure $\bar{\mu}$ on the algebra on $r_{\bar{x}, \tilde{x}}$, $\mathcal{A}$, generated by $\mathcal{C}$.

Proof of claim: We will prove the claim by Proposition 11.32. By Proposition 12.46, (i) of Proposition 11.32 is satisfied. $\forall C := r_{\bar{x}, \tilde{x}} \in \mathcal{C}$, $\forall$ pairwise disjoint $(C_j := r_{\bar{x}_j, \tilde{x}_j})_{j=1}^\infty \subseteq \mathcal{C}$ with $C = \bigcup_{j=1}^\infty C_j$, we will show that $\mu(C) \leq \sum_{j=1}^\infty \mu(C_j)$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $C = \emptyset$; Case 2: $C \neq \emptyset$.

Case 1: $C = \emptyset$. Then, $\mu(C) = 0$. Clearly $C_j = \emptyset$, $\forall j \in \mathbb{N}$. Then, $\mu(C) = 0 = \sum_{j=1}^\infty \mu(C_j)$. This case is proved.

Case 2: $C \neq \emptyset$. Then, $\tilde{x} < \hat{x}$. By (i) of Definition 12.41, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \hat{y} \in r_{\bar{x}, \hat{x}}$ such that $\Delta_F(r_{\hat{y}, \hat{x}}) > \Delta_F(r_{\bar{x}, \hat{x}}) - \epsilon/2$. Since $r_{\hat{y}, \hat{x}} \subseteq r_{\bar{x}, \hat{x}}$, by the assumption and (i) of Proposition 11.32, we have $\Delta_F(r_{\hat{y}, \hat{x}}) \leq \Delta_F(r_{\bar{x}, \hat{x}})$.

Fix any $j \in \mathbb{N}$. $C_j = r_{\bar{x}_j, \tilde{x}_j} \subseteq \bar{C}_j \subseteq \bar{C} \subseteq \Omega$. By (i) of Definition 12.41, $\exists \hat{y}_j \in \mathbb{R}^m$ with $\hat{x}_j < \hat{y}_j$ such that $\Delta_F(r_{\bar{x}_j, \hat{y}_j} \cap r_{\bar{x}, \hat{x}}) < \Delta_F(r_{\bar{x}_j, \tilde{x}_j}) + 2^{-j-1}\epsilon$. Since $C_j \subseteq r_{\bar{x}_j, \hat{y}_j} \cap r_{\bar{x}, \hat{x}}$, by the assumption and (i) of Proposition 11.32, we have $\Delta_F(r_{\bar{x}_j, \tilde{x}_j}) \leq \Delta_F(r_{\bar{x}_j, \hat{y}_j} \cap r_{\bar{x}, \hat{x}})$.

Obviously, $\bar{r}_{\hat{y}, \hat{x}} \subseteq r_{\bar{x}, \hat{x}} = C = \bigcup_{j=1}^\infty C_j = \bigcup_{j=1}^\infty r_{\bar{x}_j, \tilde{x}_j} \subseteq \bigcup_{j=1}^\infty (r_{\bar{x}_j, \hat{y}_j}^\circ \cap r_{\bar{x}, \hat{x}})$. Since $\bar{r}_{\hat{y}, \hat{x}}$ is compact, then $\exists \bar{N} \in \mathbb{N}$ such that $r_{\hat{y}, \hat{x}} \subseteq \bar{r}_{\hat{y}, \hat{x}} \subseteq \bigcup_{j=1}^{\bar{N}} (r_{\bar{x}_j, \hat{y}_j}^\circ \cap r_{\bar{x}, \hat{x}}) \subseteq \bigcup_{j=1}^{\bar{N}} (r_{\bar{x}_j, \hat{y}_j} \cap r_{\bar{x}, \hat{x}})$. By (i) of Proposition 11.32, $\mu(C) = \Delta_F(r_{\bar{x}, \hat{x}}) < \Delta_F(r_{\hat{y}, \hat{x}}) + \epsilon/2 < \sum_{j=1}^{\bar{N}} \Delta_F(r_{\bar{x}_j, \hat{y}_j} \cap r_{\bar{x}, \hat{x}}) + \epsilon/2 < \sum_{j=1}^{\bar{N}} (\Delta_F(r_{\bar{x}_j, \tilde{x}_j}) + 2^{-j-1}\epsilon) + \epsilon/2 < \sum_{j=1}^{\bar{N}} \mu(C_j) + \epsilon \leq \sum_{j=1}^\infty \mu(C_j) + \epsilon$.

By the arbitrariness of ϵ , we have $\mu(C) \leq \sum_{j=1}^{\infty} \mu(C_j)$. This case is also proved.

In both cases, we have $\mu(C) \leq \sum_{j=1}^{\infty} \mu(C_j)$. Then, (ii) of Proposition 11.32 holds. Then, by Proposition 11.32, μ admits a unique extension to a measure $\bar{\mu}$ on the algebra $\mathcal{A}$. This completes the proof of the claim. $\square$

Fix any $i \in N$. By Claim 12.48.1 and Carathéodory Extension Theorem 11.19, there exists a unique finite measure $\hat{\mu}_i$ on the measurable space $(r_{\bar{x}_i, \tilde{x}_i}, \mathcal{B}_i)$ that is an extension of $\bar{\mu}$, where $\mathcal{B}_i$ is the σ -algebra on $r_{\bar{x}_i, \tilde{x}_i}$ generated by $\mathcal{C}_i := \left\{ r_{x_1, x_2} \mid \bar{x}_i \leq x_1 \leq x_2 \leq \tilde{x}_i \right\}$. By Proposition 12.44, $\mathcal{B}_i = \mathcal{B}_B(r_{\bar{x}_i, \tilde{x}_i})$. Let $\mathcal{X}_i := (r_{\bar{x}_i, \tilde{x}_i}, \mathcal{B}_i, \hat{\mu}_i)$ be the finite measure space. By Proposition 11.118, the generation process on $(\mathcal{X}_i)_{i=1}^{\infty}$ yields a unique σ -finite measure space $\mathcal{X} := (P(\Omega), \mathcal{B}, \hat{\mu})$ such that $\mathcal{X}_i$ is the finite measure subspace of $\mathcal{X}$, $\forall i \in N$.

By Proposition 12.45, $\mathcal{B} = \mathcal{B}_B(P(\Omega))$. Then, $\mathcal{X} = (P(\Omega), \mathcal{B}_B(P(\Omega)), \hat{\mu})$. $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, $\hat{\mu}(r_{x_1, x_2}) = \hat{\mu}(\bigcup_{i \in N} (r_{x_1, x_2} \cap r_{\bar{x}_i, \tilde{x}_i})) = \sum_{i \in N} \hat{\mu}(r_{x_1, x_2} \cap r_{\bar{x}_i, \tilde{x}_i}) = \sum_{i \in N} \hat{\mu}_i(r_{x_1, x_2} \cap r_{\bar{x}_i, \tilde{x}_i}) = \sum_{i \in N} \Delta_F(r_{x_1, x_2} \cap r_{\bar{x}_i, \tilde{x}_i}) = \Delta_F(\bigcup_{i \in N} (r_{x_1, x_2} \cap r_{\bar{x}_i, \tilde{x}_i})) = \Delta_F(r_{x_1, x_2}) = T_F(r_{x_1, x_2})$, where the second equality follows from the fact that $\hat{\mu}$ is a measure; the third equality follows from Proposition 11.13; the fourth and fifth equalities follow from Claim 12.48.1; and the last equality follows from Definition 12.41 and the assumption. Hence, F is a cumulative distribution function of $\hat{\mu}$.

Finally, we will show that $T_F = \hat{\mu}(P(\Omega)) \in [0, \infty] \subset \mathbb{R}_e$. Note that $T_F = \sum_{i \in N} T_F(r_{\bar{x}_i, \tilde{x}_i}) = \sum_{i \in N} \Delta_F(r_{\bar{x}_i, \tilde{x}_i}) = \sum_{i \in N} \hat{\mu}(r_{\bar{x}_i, \tilde{x}_i}) = \hat{\mu}(P(\Omega)) \in [0, \infty] \subset \mathbb{R}_e$, where the first equality follows from Definition 12.41; the second equality follows from the assumption of the proposition and Definition 12.41; the third equality follows from the fact that F is a cumulative distribution function of $\hat{\mu}$; and the fourth equality follows from the fact that $\hat{\mu}$ is a measure. Clearly, T_F is well-defined since it is independent of $(\bar{x}_i)_{i \in N}$ and $(\tilde{x}_i)_{i \in N}$, by the uniqueness of $\hat{\mu}$. This completes the proof of the proposition. $\square$

Proposition 12.49 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a region with the subset topology $\mathcal{O}$, and $F : \Omega \rightarrow \mathbb{R}$ satisfy*

(i) *F is continuous on the right on Ω ;*

(ii) *$\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, we have $\Delta_F(r_{x_1, x_2}) \geq 0$.*

Then,

1. *F is of locally bounded variation;*

2. *there exists a unique σ -finite measure $\hat{\mu}$ on the measurable space $(P(\Omega), \mathcal{B}_B(P(\Omega)))$ such that F is a cumulative distribution function of $\hat{\mu}$;*

3. $T_F(r_{x_1, x_2}) = \Delta_F(r_{x_1, x_2}) = \hat{\mu}(r_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, and $T_F = \hat{\mu}(P(\Omega)) \in [0, \infty] \subset \mathbb{R}_e$.

Proof By (i), (i) of Definition 12.41 is satisfied. $\forall \bar{x}, \tilde{x} \in \Omega$ with $\bar{x} \leq \tilde{x}$ and $\bar{r}_{\bar{x}, \tilde{x}} \subseteq \Omega$, $\forall n \in \mathbb{Z}_+$, $\forall \check{x}_1, \hat{x}_1, \dots, \check{x}_n, \hat{x}_n \in \bar{r}_{\bar{x}, \tilde{x}}$ with $\check{x}_i \leq \hat{x}_i$, $\forall i \in \{1, \dots, n\}$, $r_{\bar{x}, \tilde{x}} = \bigcup_{i=1}^n r_{\check{x}_i, \hat{x}_i}$, and the sets in the union being pairwise disjoint, we have $\sum_{i=1}^n |\Delta_F(r_{\check{x}_i, \hat{x}_i})| = \sum_{i=1}^n \Delta_F(r_{\check{x}_i, \hat{x}_i}) = \Delta_F(r_{\bar{x}, \tilde{x}}) \in [0, \infty) \subset \mathbb{R}$, where the first equality follows from (ii); and the second equality follows from Proposition 12.46. Then, $T_F(r_{\bar{x}, \tilde{x}}) = \Delta_F(r_{\bar{x}, \tilde{x}}) \in [0, \infty) \subset \mathbb{R}$. Thus, (ii) of Definition 12.41 is satisfied.

$\forall x_1, x_2 \in \Omega$ with $x_1 < x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, define $\bar{G} : \bar{r}_{x_1, x_2} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\bar{G}(x) = T_F(r_{x_1, x})$, $\forall x \in \bar{r}_{x_1, x_2}$. By the previous paragraph, $\bar{G}(x) = \Delta_F(r_{x_1, x})$, $\forall x \in \bar{r}_{x_1, x_2}$. Then, $\forall x \in \bar{r}_{x_1, x_2}$ with $x \neq x_2$, $\bar{G}(x) = \Delta_F(r_{x_1, x}) = \lim_{\substack{h \rightarrow 0_m \\ h \geq 0_m}} \Delta_F(r_{x_1, x+h}) = \lim_{\substack{h \rightarrow 0_m \\ h \geq 0_m}} \bar{G}(x+h)$, where the second equality follows from Definition 12.41 and (i). Hence, by Proposition 12.37, (iii) of Definition 12.41 is satisfied.

This shows that F is of locally bounded variation. Hence, the first statement holds. The second and third statements of the proposition follows directly from Proposition 12.48. This completes the proof of the proposition. $\square$

Theorem 12.50 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a region with the subset topology $\mathcal{O}$, $\mathcal{Y}$ be a Banach space, $F : \Omega \rightarrow \mathcal{Y}$ be of locally bounded variation. Then, there exists a unique σ -finite $\mathcal{Y}$ -valued measure $\hat{\mu}$ on the measurable space $(P(\Omega), \mathcal{B}_B(P(\Omega)))$ such that F is a cumulative distribution function of $\hat{\mu}$. Furthermore, $\mathcal{P} \circ \hat{\mu}(r_{x_1, x_2}) = T_F(r_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, and $\mathcal{P} \circ \hat{\mu}(P(\Omega)) = T_F$.

Proof Since $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ is a region, then $\exists N \subseteq \mathbb{N}$ and $(\bar{x}_i)_{i \in N}$, $(\tilde{x}_i)_{i \in N} \subseteq \Omega$ with $\bar{x}_i \leq \tilde{x}_i$ and $\bar{r}_{\bar{x}_i, \tilde{x}_i} \subseteq \Omega$, $\forall i \in N$, such that $\Omega = \bigcup_{i \in N} \bar{r}_{\bar{x}_i, \tilde{x}_i}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i \in N}$ is pairwise disjoint, $P(\Omega) := \bigcup_{i \in N} r_{\bar{x}_i, \tilde{x}_i}$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, we have $r_{x_1, x_2} \subseteq P(\Omega)$. We need the following result.

Claim 12.50.1 Fix any $\bar{x}, \tilde{x} \in \Omega$ with $\bar{x} \leq \tilde{x}$ and $\bar{r}_{\bar{x}, \tilde{x}} \subseteq \Omega$, let $\hat{\Omega} := r_{\bar{x}, \tilde{x}} \subseteq \Omega$ be endowed with subset topology, then, there exists a unique finite $\mathcal{Y}$ -valued measure $\check{\mu}$ on the measurable space $(\hat{\Omega}, \mathcal{B}_B(\hat{\Omega}))$ such that $\check{\mu}(r_{x_1, x_2}) = \Delta_F(r_{x_1, x_2})$ and $\mathcal{P} \circ \check{\mu}(r_{x_1, x_2}) = T_F(r_{x_1, x_2})$, $\forall \bar{x} \leq x_1 \leq x_2 \leq \tilde{x}$. Furthermore, $\check{\mu} \in \mathcal{M}_{ft}(\hat{\Omega}, \mathcal{Y})$.

Proof of claim: Let $\bar{\Omega} := \bar{r}_{\bar{x}, \tilde{x}}$, $\bar{F} := F|_{\bar{\Omega}} : \bar{\Omega} \rightarrow \mathcal{Y}$, and $\bar{G} : \bar{\Omega} \rightarrow [0, \infty) \subset \mathbb{R}$ be defined by $\bar{G}(x) = T_F(r_{\bar{x}, x})$, $\forall x \in \bar{\Omega}$. Note that $\bar{G}$ is real-valued since F is of locally bounded variation. By (iii)

of Definition 12.41, $\bar{G}$ is continuous on the right. By Proposition 12.47, $\Delta_{\bar{G}}(r_{x_1, x_2}) = T_F(r_{x_1, x_2}) \geq 0$, $\forall x_1, x_2 \in \bar{\Omega}$ with $x_1 \leq x_2$. By Proposition 12.49, $\bar{G}$ is of locally bounded variation, there exists a unique σ -finite measure $\hat{\nu}$ on the measurable space $(\hat{\Omega}, \mathcal{B}_B(\hat{\Omega}))$ such that $\bar{G}$ is a cumulative distribution function of $\hat{\nu}$, and $\hat{\nu}(\hat{\Omega}) = T_{\bar{G}}(\hat{\Omega}) < \infty$. Hence, $\hat{\nu}$ is finite.

Let $\mathcal{C} := \left\{ r_{x_1, x_2} \mid \bar{x} \leq x_1 \leq x_2 \leq \tilde{x} \right\}$ be the collection of semi-open rectangles contained in $\hat{\Omega}$. Clearly, $\mathcal{C}$ is a semialgebra on $\hat{\Omega}$. Define $\nu : \mathcal{C} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\nu := \hat{\nu}|_{\mathcal{C}}$ and $\mu : \mathcal{Y} \rightarrow \mathcal{Y}$ by $\mu(r_{x_1, x_2}) = \Delta_F(r_{x_1, x_2}) \in \mathcal{Y}$, $\forall r_{x_1, x_2} \in \mathcal{C}$. It is easy to show that (i) – (v) of Proposition 12.20 are satisfied by μ and ν . $\forall r_{x_1, x_2} \in \mathcal{C}$, we have

$$\begin{aligned} \nu(r_{x_1, x_2}) &= \hat{\nu}(r_{x_1, x_2}) = T_F(r_{x_1, x_2}) \\ &= \sup_{\substack{n \in \mathbb{Z}_+, \bar{x}_1, \hat{x}_1, \dots, \bar{x}_n, \hat{x}_n \in \mathbb{R}_{x_1, x_2}, \bar{x}_i \leq \hat{x}_i, \forall i \in \{1, \dots, n\} \\ \bigcup_{i=1}^n r_{\bar{x}_i, \hat{x}_i} = r_{x_1, x_2}, r_{\bar{x}_i, \hat{x}_i} \cap r_{\bar{x}_j, \hat{x}_j} = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n \|\Delta_F(r_{\bar{x}_i, \hat{x}_i})\| \\ &= \sup_{n \in \mathbb{Z}_+, (C_i)_{i=1}^n \subseteq \mathcal{C}, r_{x_1, x_2} = \bigcup_{i=1}^n C_i, C_i \cap C_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(C_i)\| \end{aligned}$$

Hence, (vi) of Proposition 12.20 is satisfied. By Proposition 12.20, ν admits a unique extension to a measure $\tilde{\nu}$ on the algebra on $\hat{\Omega}$, $\mathcal{A}$, generated by $\mathcal{C}$; and μ admits a unique extension to a $\mathcal{Y}$ -valued measure $\tilde{\mu}$ on the algebra $\mathcal{A}$ with $\mathcal{P} \circ \tilde{\mu} = \tilde{\nu}$ and $\mathcal{P} \circ \tilde{\mu}|_{\mathcal{C}} = \nu$. Clearly, $\tilde{\nu}$ is finite. By Carathéodory Extension Theorem 12.4, there is a unique $\mathcal{Y}$ -valued measure $\check{\mu}$ on $(\hat{\Omega}, \bar{\mathcal{B}})$ such that $\check{\mu}|_{\mathcal{A}} = \tilde{\mu}$ and $(\mathcal{P} \circ \check{\mu})|_{\mathcal{A}} = \tilde{\nu}$, where $\bar{\mathcal{B}}$ is the σ -algebra on $\hat{\Omega}$ generated by $\mathcal{A}$. Then, $\bar{\mathcal{B}}$ equals to the σ -algebra on $\hat{\Omega}$ generated by $\mathcal{C}$. By Proposition 12.44, $\bar{\mathcal{B}} = \mathcal{B}_B(\hat{\Omega})$. By Carathéodory Extension Theorem 11.19, we have $\mathcal{P} \circ \check{\mu} = \hat{\nu}$. Then, $\check{\mu}$ is a finite $\mathcal{Y}$ -valued measure on $(\hat{\Omega}, \mathcal{B}_B(\hat{\Omega}))$. By Theorem 11.198, $\check{\mu} \in \mathcal{M}_{ft}(\hat{\Omega}, \mathcal{Y})$. This completes the proof of the claim. $\square$

$\forall i \in N$, let $\mathbb{X}_i := (r_{\bar{x}_i, \hat{x}_i}, \mathcal{B}_B(r_{\bar{x}_i, \hat{x}_i}), \hat{\mu}_i)$ be the unique finite $\mathcal{Y}$ -valued metric measure space as defined in Claim 12.50.1. By Proposition 11.118, the generation process on $(\mathbb{X}_i)_{i \in N}$ yields a unique σ -finite $\mathcal{Y}$ -valued measure space $\mathcal{X} := (\mathcal{P}(\Omega), \mathcal{B}, \hat{\mu})$ such that $\mathbb{X}_i$ is the finite $\mathcal{Y}$ -valued measure subspace of $\mathcal{X}$, $\forall i \in N$. By Proposition 12.45, we have $\mathcal{B} = \mathcal{B}_B(\mathcal{P}(\Omega))$.

Next, we will show that F is a cumulative distribution function of $\hat{\mu}$. Fix any $x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\mathbb{R}_{x_1, x_2} \subseteq \Omega$. Let $(r_{x_1, x_2}, \mathcal{B}_B(r_{x_1, x_2}), \check{\mu})$ be the finite $\mathcal{Y}$ -valued metric measure space as prescribed in Claim 12.50.1. Note that $\mathcal{P} \circ \hat{\mu}(r_{x_1, x_2}) = \sum_{i \in N} \mathcal{P} \circ \hat{\mu}(r_{x_1, x_2} \cap r_{\bar{x}_i, \hat{x}_i}) = \sum_{i \in N} \mathcal{P} \circ \hat{\mu}_i(r_{x_1, x_2} \cap r_{\bar{x}_i, \hat{x}_i}) = \sum_{i \in N} T_F(r_{x_1, x_2} \cap r_{\bar{x}_i, \hat{x}_i}) = \sum_{i \in N} \mathcal{P} \circ \check{\mu}(r_{x_1, x_2} \cap r_{\bar{x}_i, \hat{x}_i}) = \mathcal{P} \circ \check{\mu}(r_{x_1, x_2}) = T_F(r_{x_1, x_2}) < \infty$, where the first equality follows from the fact $\mathcal{P} \circ \hat{\mu}$ is a measure; the second equality follows from Proposition 11.115; the third equality follows from Claim 12.50.1; the fourth equality follows from Claim 12.50.1; the fifth equality follows from the fact that $\mathcal{P} \circ \check{\mu}$ is a measure; and the sixth equality and the inequality follow from Claim 12.50.1.

This implies that $\hat{\mu}(r_{x_1, x_2}) = \sum_{i \in N} \hat{\mu}(r_{x_1, x_2} \cap r_{\bar{x}_i, \tilde{x}_i}) = \sum_{i \in N} \hat{\mu}_i(r_{x_1, x_2} \cap r_{\bar{x}_i, \tilde{x}_i}) = \sum_{i \in N} \Delta_F(r_{x_1, x_2} \cap r_{\bar{x}_i, \tilde{x}_i}) = \sum_{i \in N} \check{\mu}(r_{x_1, x_2} \cap r_{\bar{x}_i, \tilde{x}_i}) = \check{\mu}(r_{x_1, x_2}) = \Delta_F(r_{x_1, x_2}) \in \mathcal{Y}$, where the first equality follows from the fact that $\hat{\mu}$ is a $\mathcal{Y}$ -valued measure; the second equality follows from Proposition 11.115; the third equality follows from Claim 12.50.1; the fourth equality follows from Claim 12.50.1; the fifth equality follows from the fact that $\check{\mu}$ is a finite $\mathcal{Y}$ -valued measure; and the last equality follows from Claim 12.50.1. By the arbitrariness of x_1 and x_2 , F is a cumulative distribution function of $\hat{\mu}$. Clearly, $T_F = \sum_{i \in N} T_F(r_{\bar{x}_i, \tilde{x}_i}) = \sum_{i \in N} \mathcal{P} \circ \hat{\mu}(r_{\bar{x}_i, \tilde{x}_i}) = \mathcal{P} \circ \hat{\mu}(P(\Omega))$, where the first equality follows from Definition 12.41; the second equality follows from the preceding argument; and the last equality follows from the fact $\mathcal{P} \circ \hat{\mu}$ is a measure. T_F is well-defined since it is independent of $(\bar{x}_i)_{i \in N}$ and $(\tilde{x}_i)_{i \in N}$, by the uniqueness of $\hat{\mu}$.

This completes the proof of the theorem. $\square$

Proposition 12.51 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone P , $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a region with the subset topology $\mathcal{O}$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $\mu \in \mathcal{M}_f(P(\Omega), \mathcal{B}_B(P(\Omega)), \mathcal{Y})$. Define $F : \Omega \rightarrow \mathcal{Y}$ and $G : \Omega \rightarrow [0, \infty) \subset \mathbb{R}$ by $F(x) = \mu(\{\bar{x} \in P(\Omega) \mid \bar{x} \leq x\})$ and $G(x) = \mathcal{P} \circ \mu(\{\bar{x} \in P(\Omega) \mid \bar{x} \leq x\})$, $\forall x \in \Omega$. Then,*

- (i) *F is a cumulative distribution function of μ , F is of locally bounded variation, $T_F(r_{x_1, x_2}) \leq \mathcal{P} \circ \mu(r_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, and $T_F \leq \mathcal{P} \circ \mu(P(\Omega))$ if it is well-defined;*
- (ii) *G is a cumulative distribution function of $\mathcal{P} \circ \mu$ and G is of bounded variation.*
- (iii) *If, in addition, $\mathcal{Y}$ is a Banach space, then, $T_F(r_{x_1, x_2}) = \mathcal{P} \circ \mu(r_{x_1, x_2}) = \Delta_G(r_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, and $T_F = \mathcal{P} \circ \mu(P(\Omega))$. Therefore, F is of bounded variation.*

Proof Since Ω is a region, then $\exists N \subseteq \mathbb{N}$ and $(\bar{x}_i)_{i \in N}, (\tilde{x}_i)_{i \in N} \subseteq \Omega$ with $\bar{x}_i \leq \tilde{x}_i$ and $\bar{r}_{\bar{x}_i, \tilde{x}_i} \subseteq \Omega$, $\forall i \in N$, such that $\Omega = \bigcup_{i \in N} \bar{r}_{\bar{x}_i, \tilde{x}_i}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i \in N}$ is pairwise disjoint, $P(\Omega) := \bigcup_{i \in N} r_{\bar{x}_i, \tilde{x}_i}$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, we have $r_{x_1, x_2} \subseteq P(\Omega)$.

$\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, we have

$$\begin{aligned} \Delta_F(r_{x_1, x_2}) &= \begin{cases} \sum_{i=0}^m \sum_{\hat{x} \in V\text{Rect}_{i, x_1, x_2}} (-1)^i F(\hat{x}) & x_1 < x_2 \\ \vartheta_{\mathcal{Y}} & x_1 \not< x_2 \end{cases} \\ &= \begin{cases} \sum_{i=0}^m \sum_{\hat{x} \in V\text{Rect}_{i, x_1, x_2}} (-1)^i \mu(\{\bar{x} \in P(\Omega) \mid \bar{x} \leq \hat{x}\}) & x_1 < x_2 \\ \vartheta_{\mathcal{Y}} & x_1 \not< x_2 \end{cases} \\ &= \mu(P(\Omega) \cap r_{x_1, x_2}) = \mu(r_{x_1, x_2}) \end{aligned}$$

Hence, F is a cumulative distribution function of μ .

By Proposition 11.4, $G(x_1) \leq G(x_2)$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. By an argument that is similar to the previous paragraph, we have G is a cumulative distribution function of $\mathcal{P} \circ \mu$ and $\Delta_G(\mathbf{r}_{x_1, x_2}) = \mathcal{P} \circ \mu(\mathbf{r}_{x_1, x_2}) \geq 0$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{\mathbf{r}}_{x_1, x_2} \subseteq \Omega$. Since μ is finite, by Proposition 11.5, we have, $\forall x \in \Omega$, $G(x) = \mathcal{P} \circ \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x\}) = \mathcal{P} \circ \mu(\bigcap_{i=1}^{\infty} \{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x + \mathbf{1}_m/i\}) = \lim_{i \in \mathbb{N}} \mathcal{P} \circ \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x + \mathbf{1}_m/i\})$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\mathcal{P} \circ \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x + \mathbf{1}_m/n_0\}) < G(x) + \epsilon$. Let $\delta := 1/n_0 \in (0, \infty) \subset \mathbb{R}$. $\forall \tilde{x} \in \mathcal{B}_{\mathbb{R}^m}(x, \delta) \cap \Omega$ with $x \leq \tilde{x}$, we have $G(x) = \mathcal{P} \circ \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x\}) \leq \mathcal{P} \circ \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq \tilde{x}\}) = G(\tilde{x}) \leq \mathcal{P} \circ \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x + \delta \mathbf{1}_m\}) < G(x) + \epsilon$. Hence, G is continuous on the right. By Proposition 12.49, G is of locally bounded variation. Then, $T_G = \sum_{i \in \mathbb{N}} T_G(\mathbf{r}_{\bar{x}_i, \tilde{x}_i}) = \sum_{i \in \mathbb{N}} \Delta_G(\mathbf{r}_{\bar{x}_i, \tilde{x}_i}) = \sum_{i \in \mathbb{N}} \mathcal{P} \circ \mu(\mathbf{r}_{\bar{x}_i, \tilde{x}_i}) = \mathcal{P} \circ \mu(\mathbf{P}(\Omega)) < \infty$, where the first equality follows from Definition 12.41; the second equality follows from the fact that $\Delta_G(\mathbf{r}_{x_1, x_2}) = \mathcal{P} \circ \mu(\mathbf{r}_{x_1, x_2}) \geq 0$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{\mathbf{r}}_{x_1, x_2} \subseteq \Omega$; the third equality follows from the fact that G is cumulative distribution function of $\mathcal{P} \circ \mu$; and the last two steps follow from the fact that $\mathcal{P} \circ \mu$ is a finite measure. Clearly, T_G is well defined since it is independent of $(\bar{x}_i)_{i \in \mathbb{N}}$ and $(\tilde{x}_i)_{i \in \mathbb{N}}$. Hence, G is of bounded variation. Thus, (ii) holds.

$\forall x \in \Omega$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by G being continuous on the right, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall \mathbf{0}_m \leq h < \delta \mathbf{1}_m$ with $x + h \in \Omega$, we have $G(x) \leq G(x + h) < G(x) + \epsilon$. Then, $\|F(x + h) - F(x)\| = \left\| \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x + h\}) - \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x\}) \right\| = \left\| \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x + h\} \setminus \{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x\}) \right\| \leq \mathcal{P} \circ \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x + h\} \setminus \{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x\}) = \mathcal{P} \circ \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x + h\}) - \mathcal{P} \circ \mu(\{\bar{x} \in \mathbf{P}(\Omega) \mid \bar{x} \leq x\}) = G(x + h) - G(x) < \epsilon$, where the second equality and the first inequality follow from the fact that μ is a finite $\mathcal{Y}$ -valued measure; and the third equality follows from Fact 11.3. Hence, F is continuous on the right. Then, F satisfies (i) of Definition 12.41.

$\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{\mathbf{r}}_{x_1, x_2} \subseteq \Omega$, we have

$$\begin{aligned} T_F(\mathbf{r}_{x_1, x_2}) &= \sup_{\substack{n \in \mathbb{Z}_+, \bar{x}_1, \hat{x}_1, \dots, \bar{x}_n, \hat{x}_n \in \bar{\mathbf{r}}_{x_1, x_2}, \bar{x}_i \leq \hat{x}_i, \forall i \in \{1, \dots, n\} \\ \bigcup_{i=1}^n \mathbf{r}_{\bar{x}_i, \hat{x}_i} = \mathbf{r}_{x_1, x_2}, \mathbf{r}_{\bar{x}_i, \hat{x}_i} \cap \mathbf{r}_{\bar{x}_j, \hat{x}_j} = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n \|\Delta_F(\mathbf{r}_{\bar{x}_i, \hat{x}_i})\| \\ &= \sup_{\substack{n \in \mathbb{Z}_+, \bar{x}_1, \hat{x}_1, \dots, \bar{x}_n, \hat{x}_n \in \bar{\mathbf{r}}_{x_1, x_2}, \bar{x}_i \leq \hat{x}_i, \forall i \in \{1, \dots, n\} \\ \bigcup_{i=1}^n \mathbf{r}_{\bar{x}_i, \hat{x}_i} = \mathbf{r}_{x_1, x_2}, \mathbf{r}_{\bar{x}_i, \hat{x}_i} \cap \mathbf{r}_{\bar{x}_j, \hat{x}_j} = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n \|\mu(\mathbf{r}_{\bar{x}_i, \hat{x}_i})\| \end{aligned}$$

$$\begin{aligned}
&\leq \sup_{\substack{n \in \mathbb{Z}_+, \bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, \bar{x}_n \in \bar{\Gamma}_{x_1, x_2}, \bar{x}_i \leq \bar{x}_i, \forall i \in \{1, \dots, n\} \\ \bigcup_{i=1}^n \bar{\Gamma}_{\bar{x}_i, \bar{x}_i} = \bar{\Gamma}_{x_1, x_2}, \bar{\Gamma}_{\bar{x}_i, \bar{x}_i} \cap \bar{\Gamma}_{\bar{x}_j, \bar{x}_j} = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n \mathcal{P} \circ \mu(\bar{\Gamma}_{\bar{x}_i, \bar{x}_i}) \\
&= \mathcal{P} \circ \mu(\bar{\Gamma}_{x_1, x_2}) < \mathcal{P} \circ \mu(\Omega) < \infty
\end{aligned}$$

where the second equality follows from the fact that F is a cumulative distribution function of μ ; the first inequality follows from Definition 11.108; the third equality follows from the fact that $\mathcal{P} \circ \mu$ is a measure; the second inequality follows from Proposition 11.4; and the last inequality follows from the fact that μ is finite. Then, F satisfies (ii) of Definition 12.41.

$\forall x_1, x_2 \in \Omega$ with $x_1 < x_2$ and $\bar{\Gamma}_{x_1, x_2} \subseteq \Omega$, define $\bar{G} : \bar{\Gamma}_{x_1, x_2} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\bar{G}(x) = T_F(\bar{\Gamma}_{x_1, x})$, $\forall x \in \bar{\Gamma}_{x_1, x_2}$. $\forall x \in \bar{\Gamma}_{x_1, x_2}$, by $\bar{G}$ being continuous on the right, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall \mathbf{0}_m \leq h < \delta \mathbf{1}_m$ with $x + h \in \Omega$, we have $\bar{G}(x) \leq \bar{G}(x + h) < \bar{G}(x) + \epsilon$. Then, $\forall \mathbf{0}_m \leq h < \delta \mathbf{1}_m$ with $x + h \in \bar{\Gamma}_{x_1, x_2}$, $0 \leq \bar{G}(x + h) - \bar{G}(x) = T_F(\bar{\Gamma}_{x_1, x+h}) - T_F(\bar{\Gamma}_{x_1, x}) = \sum_{(\bar{x}, \hat{x}) \in M} T_F(\bar{\Gamma}_{\bar{x}, \hat{x}}) \leq \sum_{(\bar{x}, \hat{x}) \in M} \mathcal{P} \circ \mu(\bar{\Gamma}_{\bar{x}, \hat{x}}) = \mathcal{P} \circ \mu(\bar{\Gamma}_{x_1, x+h} \setminus \bar{\Gamma}_{x_1, x}) \leq \mathcal{P} \circ \mu(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x + h\} \setminus \{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\}) = \mathcal{P} \circ \mu(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x + h\}) - \mathcal{P} \circ \mu(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\}) = \bar{G}(x + h) - \bar{G}(x) < \epsilon$, where the first inequality follows from the monotonicity of $\bar{G}$; $M := \{(\bar{x}, \hat{x}) \in \bar{\Gamma}_{x_1, x} \times \bar{\Gamma}_{x, x+h} \mid i \in \{0, \dots, m-1\}, J \subseteq \bar{J}_m := \{1, \dots, m\}, \text{card}(J) = i, \pi_k(\bar{x}) = \pi_k(x_1), \forall k \in J, \pi_k(\bar{x}) = \pi_k(x), \forall k \in \bar{J}_m \setminus J, \pi_k(\hat{x}) = \pi_k(x), \forall k \in J, \pi_k(\hat{x}) = \pi_k(x + h), \forall k \in \bar{J}_m \setminus J\}$; the second equality follows from Proposition 12.47; the second inequality follows from the last paragraph; the third equality follows from the fact that $\mathcal{P} \circ \mu$ is a measure; the third inequality follows from Proposition 11.4; and the fourth equality follows from the fact that $\mathcal{P} \circ \mu$ is a finite measure. Hence, $\bar{G}$ is continuous on the right. Thus, F satisfies (iii) of Definition 12.41.

Therefore, F is of locally bounded variation. When T_F is well-defined, we have $T_F = \sum_{i \in N} T_F(\bar{\Gamma}_{\bar{x}_i, \bar{x}_i}) \leq \sum_{i \in N} \mathcal{P} \circ \mu(\bar{\Gamma}_{\bar{x}_i, \bar{x}_i}) = \mathcal{P} \circ \mu(\mathcal{P}(\Omega)) < \infty$, where the first equality follows from Definition 12.41; the first inequality follows from the second to last paragraph; and the second equality and the second inequality follows from the fact that $\mathcal{P} \circ \mu$ is a finite measure. Thus, (i) holds.

(iii) Let $\mathcal{Y}$ be a Banach space. By Theorem 12.50 and (i), μ is the unique σ -finite $\mathcal{Y}$ -valued measure on the measurable space $(\mathcal{P}(\Omega), \mathcal{B}_{\mathcal{B}}(\mathcal{P}(\Omega)))$ such that F is a cumulative distribution function of μ . Theorem 12.50, coupled with (ii), further concludes that $T_F(\bar{\Gamma}_{x_1, x_2}) = \mathcal{P} \circ \mu(\bar{\Gamma}_{x_1, x_2}) = \Delta_G(\bar{\Gamma}_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{\Gamma}_{x_1, x_2} \subseteq \Omega$, and $T_F = \mathcal{P} \circ \mu(\mathcal{P}(\Omega)) < \infty$. Since T_F is well-defined and is finite, then F is of bounded variation.

This completes the proof of the proposition. $\square$

Proposition 12.52 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone P , $\Omega \in \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m)$ be a rectangle with the subset topology $\mathcal{O}$, $\mathcal{Y}$*

be a normed linear space over $\mathbb{K}$, μ be a σ -finite $\mathcal{Y}$ -valued measure on $(P(\Omega), \mathcal{B}_B(P(\Omega)))$ such that $\mathcal{P} \circ \mu(r_{x_1, x_2}) < \infty$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. By Proposition 12.39, $P(\Omega) = \prod_{i=1}^m \bar{I}_i$ and Ω is a region.

Define $F : \Omega \rightarrow \mathcal{Y}$ by, if $\Omega \neq \emptyset$ then fix $x_0 \in \Omega$, $F(x) = \vartheta_{\mathcal{Y}}$, $\forall x \in \Omega$ with $\pi_{i_0}(x) = \pi_{i_0}(x_0)$ for some $i_0 \in \bar{J} := \{1, \dots, m\}$; and $\forall x \in \Omega$ with $\pi_i(x) \neq \pi_i(x_0)$, $\forall i \in \bar{J}$, $F(x)$ is such that $\Delta_F(r_{\check{x}, \hat{x}}) = \mu(r_{\check{x}, \hat{x}})$, where $\check{x} := x \wedge x_0$ and $\hat{x} := x \vee x_0$. Then, F is a cumulative distribution function of μ , which is called the cumulative distribution function of μ with origin x_0 . Define $G : \Omega \rightarrow \mathbb{R}$ to be the cumulative distribution function of $\mathcal{P} \circ \mu$ with origin x_0 . Then, the following statements hold.

(i) F and G are of locally bounded variation.

(ii) If, in addition, $\mathcal{Y}$ is a Banach space, then $T_F(r_{x_1, x_2}) = \mathcal{P} \circ \mu(r_{x_1, x_2}) = \Delta_G(r_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, and $T_F = \mathcal{P} \circ \mu(P(\Omega))$.

Proof (i) By the assumption on μ , we have $\mathcal{P} \circ \mu(r_{x_1, x_2}) < \infty$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. Then, $r_{x_1, x_2} \in \text{dom}(\mu)$. This implies that $F(x) \in \mathcal{Y}$ and $G(x) \in \mathbb{R}$, $\forall x \in \Omega$. Hence, F and G are well defined.

If $\Omega = \emptyset$, then the result holds trivially. This case is proved. Consider the case $\Omega \neq \emptyset$.

Claim 12.52.1 F is a cumulative distribution function of μ .

Proof of claim: $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. If $x_1 \not\leq x_2$, then by Definition 12.41, we have $\mu(r_{x_1, x_2}) = \mu(\emptyset) = \vartheta_{\mathcal{Y}} = \Delta_F(r_{x_1, x_2})$. On the other hand, if $x_1 \leq x_2$, then let $n := \text{card}(\{i \in \bar{J} \mid (\pi_i(x_1) - \pi_i(x_0))(\pi_i(x_2) - \pi_i(x_0)) < 0\}) \in \{0, \dots, m\}$. We will show that $\mu(r_{x_1, x_2}) = \Delta_F(r_{x_1, x_2})$ by mathematical induction on n .

1° Consider the case $n = 0$. Define $J := \{i \in \bar{J} \mid \pi_i(x_1) < \pi_i(x_0) \text{ or } \pi_i(x_2) < \pi_i(x_0)\}$ and $\bar{n} := \text{card}(J)$.

$\forall j \in \{0, \dots, m\}$, $\forall x \in \text{VRect}_{j, x_1, x_2}$, if $\pi_i(x) \neq \pi_i(x_0)$, $\forall i \in \bar{J}$, then, $\bar{n} = \text{card}(\{i \in \bar{J} \mid \pi_i(x) < \pi_i(x_0)\})$, by the definition of F and Definition 12.41, we have $\Delta_F(r_{\check{x}, \hat{x}}) = \mu(r_{\check{x}, \hat{x}}) = (-1)^{\bar{n}} F(x)$. On the other hand, if $\pi_{i_0}(x) = \pi_{i_0}(x_0)$ for some $i_0 \in \bar{J}$, then, by the definition of F and Definition 12.41, $\Delta_F(r_{\check{x}, \hat{x}}) = \vartheta_{\mathcal{Y}} = \mu(\emptyset) = \mu(r_{\check{x}, \hat{x}}) = F(x) = (-1)^{\bar{n}} F(x)$. Hence, we have $\mu(r_{\check{x}, \hat{x}}) = (-1)^{\bar{n}} F(x)$, $\forall x \in \text{VRect}_{j, x_1, x_2}$, $\forall j \in \{0, \dots, m\}$.

Then, we have

$$\begin{aligned} \mu(r_{x_1, x_2}) &= \mu\left(\left\{x \in P(\Omega) \mid x_1 \leq x \leq x_2\right\}\right) \\ &= \mu(\{x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_2) \leq \pi_i(x_0), \forall i \in J; \\ &\quad \pi_i(x_0) \leq \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_2), \forall i \in \bar{J} \setminus J\}) \\ &= \mu(\{x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J; \end{aligned}$$

$$\begin{aligned}
& \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_2), \forall i \in \bar{J} \setminus J \} \\
& -\mu \left(\bigcup_{i_0=1}^m \{x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \setminus \{i_0\}; \right. \\
& \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_2), \forall i \in (\bar{J} \setminus J) \setminus \{i_0\}; \\
& \pi_i(x_2) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \cap \{i_0\}; \\
& \left. \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_1), \forall i \in (\bar{J} \setminus J) \cap \{i_0\} \} \right) \\
& = \sum_{j=0}^m (-1)^j \sum_{\hat{J} \subseteq \bar{J}, \text{card}(\hat{J})=j} \mu(\{x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_0), \\
& \forall i \in J \setminus \hat{J}; \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_2), \forall i \in (\bar{J} \setminus J) \setminus \hat{J}; \\
& \pi_i(x_2) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \cap \hat{J}; \\
& \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_1), \forall i \in (\bar{J} \setminus J) \cap \hat{J} \}) \\
& \cdot \pi_i(x_j) \quad \quad \quad = \\
& \begin{cases} \pi_i(x_1) & \forall i \in (J \setminus \hat{J}) \cup ((\bar{J} \setminus J) \cap \hat{J}) = J \triangle \hat{J} \\ \pi_i(x_2) & \forall i \in ((\bar{J} \setminus J) \setminus \hat{J}) \cup (J \cap \hat{J}) = \bar{J} \setminus (J \triangle \hat{J}) \end{cases}, \quad \forall i \in \bar{J}. \quad \text{Then,} \\
& \left\{ x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \setminus \hat{J}; \pi_i(x_0) < \pi_i(x) \leq \right. \\
& \pi_i(x_2), \forall i \in (\bar{J} \setminus J) \setminus \hat{J}; \pi_i(x_2) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \cap \hat{J}; \pi_i(x_0) < \\
& \pi_i(x) \leq \pi_i(x_1), \forall i \in (\bar{J} \setminus J) \cap \hat{J} \} = r_{\hat{x}_j, \hat{x}_j}. \text{ By the definition of } x_j, \text{ we have} \\
& x_j \in \text{VRect}_{\text{card}(J \triangle \hat{J}), x_1, x_2} \text{ and } \text{card}(J \triangle \hat{J}) \equiv \text{card}(J) + \text{card}(\hat{J}) = \bar{n} + j \\
& \pmod{2}, \forall \hat{J} \subseteq \bar{J} \text{ with } \text{card}(\hat{J}) = j \in \{0, \dots, m\}. \text{ This leads to the con-} \\
& \text{clusion that}
\end{aligned}$$

$$\begin{aligned}
\mu(r_{x_1, x_2}) &= \sum_{j=0}^m (-1)^j \sum_{\hat{J} \subseteq \bar{J}, \text{card}(\hat{J})=j} \mu(r_{\hat{x}_j, \hat{x}_j}) \\
&= \sum_{j=0}^m (-1)^j \sum_{\hat{J} \subseteq \bar{J}, \text{card}(\hat{J})=j} (-1)^{\bar{n}} F(x_j) \\
&= \sum_{j=0}^m \sum_{\hat{J} \subseteq \bar{J}, \text{card}(\hat{J})=j} (-1)^{\text{card}(J \triangle \hat{J})} F(x_j) \\
&= \sum_{i=0}^m \sum_{\bar{x} \in \text{VRect}_i, x_1, x_2} (-1)^i F(\bar{x}) = \Delta_F(r_{x_1, x_2})
\end{aligned}$$

where the second equality follows from the previous paragraph. This case is proved.

2° Assume that $\mu(r_{x_1, x_2}) = \Delta_F(r_{x_1, x_2})$, $\forall n \leq k-1 \in \{0, \dots, m-1\}$.

3° Consider the case when $n = k \in \bar{J}$. Since $n > 0$, then $\exists i_0 \in \bar{J}$ such that $\pi_{i_0}(x_1) < \pi_{i_0}(x_0) < \pi_{i_0}(x_2)$. Without loss of generality, assume that

$i_0 = 1$. Then, we have $\mu(r_{x_1, x_2}) = \mu(\{x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_2), \forall i \in \bar{J}\}) = \mu(\{x \in P(\Omega) \mid \pi_1(x_1) < \pi_1(x) \leq \pi_1(x_0); \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_2), \forall i \in \bar{J} \setminus \{1\}\}) + \mu(\{x \in P(\Omega) \mid \pi_1(x_0) < \pi_1(x) \leq \pi_1(x_2); \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_2), \forall i \in \bar{J} \setminus \{1\}\}) = \mu(r_{x_1, x_3}) + \mu(r_{x_4, x_2})$, where $x_3, x_4 \in \Omega$ are defined by $\pi_i(x_3) = \begin{cases} \pi_i(x_0) & \text{if } i = 1 \\ \pi_i(x_2) & \forall i \in \bar{J} \setminus \{1\} \end{cases}$ and $\pi_i(x_4) = \begin{cases} \pi_i(x_0) & \text{if } i = 1 \\ \pi_i(x_1) & \forall i \in \bar{J} \setminus \{1\} \end{cases}$; and all equalities follows from the assumption on μ . By the inductive assumption, we have $\mu(r_{x_1, x_2}) = \mu(r_{x_1, x_3}) + \mu(r_{x_4, x_2}) = \Delta_F(r_{x_1, x_3}) + \Delta_F(r_{x_4, x_2}) = \Delta_F(r_{x_1, x_2})$, where the last equality follows from Proposition 12.46. This completes the induction process.

Therefore, F is a cumulative distribution function of μ . $\square$

By Claim 12.52.1, F and G are cumulative distribution functions of μ and $\mathcal{P} \circ \mu$, respectively. We need the following intermediate result.

Claim 12.52.2 $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$, if x_2 is an accumulation point of $\Omega \cap (x_2 + P)$, then $\lim_{\substack{h \rightarrow 0_m \\ h \geq 0_m}} \mathcal{P} \circ \mu(r_{x_1, x_2+h} \setminus r_{x_1, x_2}) = 0$. $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $x_1 \neq x_2$, then $\lim_{\substack{h \rightarrow 0_m, h \geq 0_m \\ x_1+h \leq x_2}} \mathcal{P} \circ \mu(r_{x_1, x_2} \setminus r_{x_1+h, x_2}) = 0$.

Proof of claim: $\forall (h_n)_{n=1}^\infty \subseteq P \cap (\Omega - x_2) \setminus \{0_m\}$ with $\lim_{n \in \mathbb{N}} h_n = 0_m$, we have $\lim_{n \in \mathbb{N}} \bar{h}_n := \lim_{n \in \mathbb{N}} \bigvee_{i=n}^\infty h_i = 0_m$, $0_m \leq \bar{h}_{n+1} \leq \bar{h}_n \neq 0_m$, and $\bar{h}_n \in P \cap (\Omega - x_2) \setminus \{0_m\}$, $\forall n \in \mathbb{N}$. Then, $0 \leq \lim_{n \in \mathbb{N}} \mathcal{P} \circ \mu(r_{x_1, x_2+h_n} \setminus r_{x_1, x_2}) \leq \lim_{n \in \mathbb{N}} \mathcal{P} \circ \mu(r_{x_1, x_2+\bar{h}_n} \setminus r_{x_1, x_2}) = \mathcal{P} \circ \mu(\bigcap_{n=1}^\infty (r_{x_1, x_2+\bar{h}_n} \setminus r_{x_1, x_2})) = \mathcal{P} \circ \mu(\emptyset) = 0$, where the first inequality follows from the fact that $\mathcal{P} \circ \mu$ is a measure; the second inequality follows from Proposition 11.4; and the first equality follows from Proposition 11.5. Then, $\lim_{n \in \mathbb{N}} \mathcal{P} \circ \mu(r_{x_1, x_2+h_n} \setminus r_{x_1, x_2}) = 0$. By Proposition 4.16, we have $\lim_{\substack{h \rightarrow 0_m \\ h \geq 0_m}} \mathcal{P} \circ \mu(r_{x_1, x_2+h} \setminus r_{x_1, x_2}) = 0$.

0. This proves the first statement.

$\forall (h_n)_{n=1}^\infty \subseteq \{h \in \mathbb{R}^m \mid x_1 \leq x_1 + h \leq x_2, h \neq 0_m\}$ with $\lim_{n \in \mathbb{N}} h_n = 0_m$, we have $\lim_{n \in \mathbb{N}} \bar{h}_n := \lim_{n \in \mathbb{N}} \bigvee_{i=n}^\infty h_i = 0_m$ and $0_m \leq \bar{h}_{n+1} \leq \bar{h}_n \neq 0_m$, $\forall n \in \mathbb{N}$. Then, $0 \leq \lim_{n \in \mathbb{N}} \mathcal{P} \circ \mu(r_{x_1, x_2} \setminus r_{x_1+h_n, x_2}) \leq \lim_{n \in \mathbb{N}} \mathcal{P} \circ \mu(r_{x_1, x_2} \setminus r_{x_1+\bar{h}_n, x_2}) = \mathcal{P} \circ \mu(\bigcap_{n=1}^\infty (r_{x_1, x_2} \setminus r_{x_1+\bar{h}_n, x_2})) = \mathcal{P} \circ \mu(\emptyset) = 0$, where the first inequality follows from the fact that $\mathcal{P} \circ \mu$ is a measure; the second inequality follows from Proposition 11.4; and the first equality follows from Proposition 11.5. Then, $\lim_{n \in \mathbb{N}} \mathcal{P} \circ \mu(r_{x_1, x_2} \setminus r_{x_1+h_n, x_2}) = 0$. By Proposition 4.16, we have $\lim_{\substack{h \rightarrow 0_m, h \geq 0_m \\ x_1+h \leq x_2}} \mathcal{P} \circ \mu(r_{x_1, x_2} \setminus r_{x_1+h, x_2}) = 0$.

This completes the proof of the claim. $\square$

$\forall x \in \Omega$, let $J := \{i \in \bar{J} \mid \pi_i(x) < \pi_i(x_0)\}$ and $n := \text{card}(J)$. Let x be an accumulation point of $\Omega \cap (x + P)$. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that, $\forall h \in P \cap (\Omega - x)$ with $h \neq 0_m$ and $h \leq \delta \mathbf{1}_m$, we have $x + h \in \Omega$

and $\{i \in \bar{J} \mid \pi_i(x+h) < \pi_i(x_0)\} = J$. Then, $0 \leq |G(x+h) - G(x)| = |(-1)^n(\mathcal{P} \circ \mu(r_{\tilde{x}_h, \hat{x}_h}) - \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}}))| \leq |\mathcal{P} \circ \mu(r_{\tilde{x}_h, \hat{x}_h}) - \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}})| + |\mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}}) - \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}})| = \mathcal{P} \circ \mu(r_{\tilde{x}_h, \hat{x}_h} \setminus r_{\tilde{x}, \hat{x}}) + \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}} \setminus r_{\tilde{x}_h, \hat{x}}) \leq \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}_h} \setminus r_{\tilde{x}, \hat{x}}) + \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}} \setminus r_{\tilde{x}_h, \hat{x}})$, where $x_h := x+h$; the first equality follows from the definition of G ; the second inequality follows from the fact that $x \leq x_h$ and $\tilde{x} \leq \tilde{x}_h \leq \hat{x} \leq \hat{x}_h$; the second equality follows from Fact 11.3; and the last inequality follows from Proposition 11.4. By Claim 12.52.2, we have $0 \leq \lim_{h \rightarrow 0_m} |G(x+h) - G(x)| \leq \lim_{h \rightarrow 0_m} \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}_h} \setminus r_{\tilde{x}, \hat{x}}) + \lim_{h \rightarrow 0_m} \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}} \setminus r_{\tilde{x}_h, \hat{x}}) = 0$. Hence, by Proposition 12.37, G is continuous on the right. By Proposition 12.49, G is of locally bounded variation.

$\forall x \in \Omega$, let x be an accumulation point of $\Omega \cap (x+P)$, and let J , n , δ , and x_h be as defined in the previous paragraph. $\forall h \in P \cap (\Omega - x)$ with $h \neq 0_m$ and $h \leq \delta 1_m$, we have $0 \leq \|F(x+h) - F(x)\| = \|(-1)^n(\mu(r_{\tilde{x}_h, \hat{x}_h}) - \mu(r_{\tilde{x}, \hat{x}}))\| \leq \|\mu(r_{\tilde{x}_h, \hat{x}_h}) - \mu(r_{\tilde{x}_h, \hat{x}})\| + \|\mu(r_{\tilde{x}_h, \hat{x}}) - \mu(r_{\tilde{x}, \hat{x}})\| = \|\mu(r_{\tilde{x}_h, \hat{x}_h} \setminus r_{\tilde{x}_h, \hat{x}})\| + \|\mu(r_{\tilde{x}, \hat{x}} \setminus r_{\tilde{x}_h, \hat{x}})\| \leq \mathcal{P} \circ \mu(r_{\tilde{x}_h, \hat{x}_h} \setminus r_{\tilde{x}_h, \hat{x}}) + \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}} \setminus r_{\tilde{x}_h, \hat{x}}) \leq \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}_h} \setminus r_{\tilde{x}, \hat{x}}) + \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}} \setminus r_{\tilde{x}_h, \hat{x}})$, where the first equality follows from the definition of F ; the second inequality follows from the fact that $x \leq x_h$ and $\tilde{x} \leq \tilde{x}_h \leq \hat{x} \leq \hat{x}_h$; the second equality and the third inequality follow from the fact that μ is a locally finite $\mathcal{Y}$ -valued measure; and the last inequality follows from Proposition 11.4. By Claim 12.52.2, we have $0 \leq \lim_{h \rightarrow 0_m} \|F(x+h) - F(x)\| \leq \lim_{h \rightarrow 0_m} \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}_h} \setminus r_{\tilde{x}, \hat{x}}) + \lim_{h \rightarrow 0_m} \mathcal{P} \circ \mu(r_{\tilde{x}, \hat{x}} \setminus r_{\tilde{x}_h, \hat{x}}) = 0$. Hence, by Proposition 12.37, F is continuous on the right and F satisfies (i) of Definition 12.41.

$\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$, we have

$$\begin{aligned}
T_F(r_{x_1, x_2}) &= \sup_{\substack{n \in \mathbb{Z}_+, \bar{x}_1, \bar{x}_1, \dots, \bar{x}_n \in \mathbb{F}_{x_1, x_2}, \bar{x}_i \leq \bar{x}_i, \forall i \in \{1, \dots, n\} \\ \bigcup_{i=1}^n r_{\bar{x}_i, \bar{x}_i} = r_{x_1, x_2}, r_{\bar{x}_i, \bar{x}_i} \cap r_{\bar{x}_j, \bar{x}_j} = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n \|\Delta_F(r_{\bar{x}_i, \bar{x}_i})\| \\
&= \sup_{\substack{n \in \mathbb{Z}_+, \bar{x}_1, \bar{x}_1, \dots, \bar{x}_n \in \mathbb{F}_{x_1, x_2}, \bar{x}_i \leq \bar{x}_i, \forall i \in \{1, \dots, n\} \\ \bigcup_{i=1}^n r_{\bar{x}_i, \bar{x}_i} = r_{x_1, x_2}, r_{\bar{x}_i, \bar{x}_i} \cap r_{\bar{x}_j, \bar{x}_j} = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n \|\mu(r_{\bar{x}_i, \bar{x}_i})\| \\
&\leq \sup_{\substack{n \in \mathbb{Z}_+, \bar{x}_1, \bar{x}_1, \dots, \bar{x}_n \in \mathbb{F}_{x_1, x_2}, \bar{x}_i \leq \bar{x}_i, \forall i \in \{1, \dots, n\} \\ \bigcup_{i=1}^n r_{\bar{x}_i, \bar{x}_i} = r_{x_1, x_2}, r_{\bar{x}_i, \bar{x}_i} \cap r_{\bar{x}_j, \bar{x}_j} = \emptyset, \forall 1 \leq i < j \leq n}} \sum_{i=1}^n \mathcal{P} \circ \mu(r_{\bar{x}_i, \bar{x}_i}) \\
&= \mathcal{P} \circ \mu(r_{x_1, x_2}) < \infty
\end{aligned}$$

where the second equality follows from the fact that F is a cumulative distribution function of μ ; the first inequality follows from Definition 11.108; the third equality follows from the fact that $\mathcal{P} \circ \mu$ is a measure; and the last inequality follows from the assumption on μ . Then, F satisfies (ii) of Definition 12.41.

$\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$, define $\bar{G} : \bar{r}_{x_1, x_2} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\bar{G}(x) = T_F(r_{x_1, x})$, $\forall x \in \bar{r}_{x_1, x_2}$. $\forall x \in \bar{r}_{x_1, x_2}$ with $x \neq x_2$, we have

$$\begin{aligned} 0 &\leq \lim_{\substack{h \rightarrow 0_m \\ h \geq 0_m}} (\bar{G}(x+h) - \bar{G}(x)) = \lim_{\substack{h \rightarrow 0_m \\ h \geq 0_m}} (T_F(r_{x_1, x+h}) - T_F(r_{x_1, x})) \\ &= \lim_{\substack{h \rightarrow 0_m \\ h \geq 0_m}} \sum_{(\bar{x}, \tilde{x}) \in M} T_F(r_{\bar{x}, \tilde{x}}) \leq \lim_{\substack{h \rightarrow 0_m \\ h \geq 0_m}} \sum_{(\bar{x}, \tilde{x}) \in M} \mathcal{P} \circ \mu(r_{\bar{x}, \tilde{x}}) \\ &= \lim_{\substack{h \rightarrow 0_m \\ h \geq 0_m}} \mathcal{P} \circ \mu(r_{x_1, x+h} \setminus r_{x_1, x}) = 0 \end{aligned}$$

where the first inequality follows from the monotonicity of $\bar{G}$; the first equality follows from the definition of $\bar{G}$; $M := \{(\bar{x}, \tilde{x}) \in \bar{r}_{x_1, x} \times \bar{r}_{x, x+h} \mid i \in \{0, \dots, m-1\}, J \subseteq \bar{J}, \text{card}(J) = i, \pi_k(\bar{x}) = \pi_k(x_1), \forall k \in J, \pi_k(\tilde{x}) = \pi_k(x), \forall k \in \bar{J} \setminus J, \pi_k(\tilde{x}) = \pi_k(x), \forall k \in J, \pi_k(\tilde{x}) = \pi_k(x+h), \forall k \in \bar{J} \setminus J\}$; the second equality follows from Proposition 12.47; the second inequality follows from the last paragraph; the third equality follows from the fact that $\mathcal{P} \circ \mu$ is a measure; and the last equality follows from Claim 12.52.2. Hence, by Proposition 12.37, $\bar{G}$ is continuous on the right. Thus, F satisfies (iii) of Definition 12.41.

Therefore, F is of locally bounded variation. This proves (i).

(ii) Let $\mathcal{Y}$ be a Banach space. By Theorem 12.50 and (i), μ is the unique σ -finite $\mathcal{Y}$ -valued measure on the measurable space $(P(\Omega), \mathcal{B}_B(P(\Omega)))$ such that F is a cumulative distribution function of μ . Theorem 12.50, coupled with (i), further concludes that $T_F(r_{x_1, x_2}) = \mathcal{P} \circ \mu(r_{x_1, x_2}) = \Delta_G(r_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, and $T_F = \mathcal{P} \circ \mu(P(\Omega))$.

This completes the proof of the proposition. $\square$

Lemma 12.53 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone P , $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a region with the subset topology $\mathcal{O}$, and $F : \Omega \rightarrow \mathbb{R}$. Assume that the following holds.*

(i) *F is continuous on the right.*

(ii) *F is monotonically nondecreasing, i. e., $F(x_1) \leq F(x_2)$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$.*

Then, F is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable.

Proof Consider the special case when Ω is a rectangle. $\forall \alpha \in \mathbb{R}$, let $B := \{\bar{x} \in \Omega \mid F(\bar{x}) < \alpha\}$. We will show that $B \in \mathcal{O} \subseteq \mathcal{B}_B(\Omega) \subseteq \mathcal{B}_B(\mathbb{R}^m)$. Then, by Proposition 11.35, F is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable.

$\forall x \in B$, $F(x) < \alpha$. Since F is continuous on the right, then $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $F(\bar{x}) < \alpha$, $\forall \bar{x} \in \Omega \cap (x + P)$ with $\bar{x} \leq x + \delta \mathbf{1}_m$. $\forall \bar{x} \in \{\tilde{x} \in \Omega \mid \tilde{x} \leq x + \delta \mathbf{1}_m\} =: O \in \mathcal{O}$, since $\bar{x}, x \in \Omega$ and Ω is a rectangle, we have $x \vee \bar{x} \in \Omega \cap (x + P)$ with $x \vee \bar{x} \leq x + \delta \mathbf{1}_m$. Then, $F(\bar{x}) \leq F(x \vee \bar{x}) < \alpha$.

This implies that $\bar{x} \in B$. By the arbitrariness of $\bar{x}$, we have $O \subseteq B$. By the arbitrariness of x , we have $B \in \mathcal{O} \subseteq \mathcal{B}_B(\Omega) \subseteq \mathcal{B}_B(\mathbb{R}^m)$.

Now, consider the general case when Ω is a region. Then, $\exists N \subseteq \mathbb{N}$, $\exists (\bar{x}_i)_{i \in N}, (\tilde{x}_i)_{i \in N} \subseteq \Omega$ with $\bar{x}_i \leq \tilde{x}_i$ and $\bar{\Gamma}_{\bar{x}_i, \tilde{x}_i} \subseteq \Omega$, $\forall i \in N$, such that $\Omega = \bigcup_{i \in N} \bar{\Gamma}_{\bar{x}_i, \tilde{x}_i}$. $\forall \alpha \in \mathbb{R}$, let $B := \{\hat{x} \in \Omega \mid F(\hat{x}) < \alpha\} = \bigcup_{i \in N} (B \cap \bar{\Gamma}_{\bar{x}_i, \tilde{x}_i}) =: \bigcup_{i \in N} B_i$. By the special case, $B_i \in \mathcal{B}_B(\mathbb{R}^m)$, $\forall i \in N$. Then, $B \in \mathcal{B}_B(\mathbb{R}^m)$. By Proposition 11.35, F is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable.

This completes the proof of the lemma. $\square$

Proposition 12.54 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a region with the subset topology $\mathcal{O}$, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable, $\mu \in \mathcal{M}_f(\mathcal{P}(\Omega), \mathcal{B}_B(\mathcal{P}(\Omega)), \mathcal{Y})$, and $F : \Omega \rightarrow \mathcal{Y}$ be the cumulative distribution function of μ as defined by $F(x) = \mu(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\})$, $\forall x \in \Omega$, as shown in Proposition 12.51. Then, F is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable.*

Proof We will prove the result in three steps. In the first step, we consider the special case $\mathcal{Y} = \mathbb{R}$. By Jordan Decomposition Theorem 11.162, there exists a unique pair of mutually singular finite measures μ_+ and μ_- on the measurable space $(\mathcal{P}(\Omega), \mathcal{B}_B(\mathcal{P}(\Omega)))$ such that $\mu = \mu_+ - \mu_-$. Furthermore, $\mathcal{P} \circ \mu = \mu_+ + \mu_-$. Let $F_+ : \Omega \rightarrow [0, \infty) \subset \mathbb{R}$ and $F_- : \Omega \rightarrow [0, \infty) \subset \mathbb{R}$ be the cumulative distribution functions of μ_+ and μ_- , respectively, as defined by $F_+(x) = \mu_+(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\})$ and $F_-(x) = \mu_-(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\})$, $\forall x \in \Omega$. By Proposition 12.51, F_+ and F_- are of locally bounded variation, and hence, are continuous on the right. By their definitions, $F_+(x_1) \leq F_+(x_2)$ and $F_-(x_1) \leq F_-(x_2)$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. Hence, F_+ and F_- are monotonically nondecreasing. By Lemma 12.53, F_+ and F_- are $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. Clearly, $F(x) = \mu(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\}) = \mu_+(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\}) - \mu_-(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\}) = F_+(x) - F_-(x)$, $\forall x \in \Omega$. By Propositions 11.38, 11.39, and 7.23, F is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. This completes the first step.

In the second step, we consider the special case $\mathcal{Y} = \mathbb{C}$. By Proposition 11.164, $\mu = \mu_r + i\mu_i$, where μ_r and μ_i are finite $\mathbb{R}$ -valued measures on $(\mathcal{P}(\Omega), \mathcal{B}_B(\mathcal{P}(\Omega)))$. Let $F_r : \Omega \rightarrow \mathbb{R}$ and $F_i : \Omega \rightarrow \mathbb{R}$ be defined by $F_r(x) = \mu_r(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\})$ and $F_i(x) = \mu_i(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\})$, $\forall x \in \Omega$. By Proposition 12.51, F_r and F_i are cumulative distribution functions of μ_r and μ_i , respectively. By the first step, F_r and F_i are $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. Clearly, $F(x) = \mu(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\}) = \mu_r(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\}) + i\mu_i(\{\bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x\}) = F_r(x) + iF_i(x)$, $\forall x \in \Omega$. By Propositions 11.38, 11.39, and 7.23, F is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. This completes the second step.

In the last step, we consider the general case where $\mathcal{Y}$ is a separable normed linear space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable. Fix any $y_* \in \mathcal{Y}^*$. Define $\mu_{y_*} : \mathcal{B}_B(P(\Omega)) \rightarrow \mathbb{K}$ by $\mu_{y_*}(E) = \langle\langle y_*, \mu(E) \rangle\rangle$, $\forall E \in \mathcal{B}_B(P(\Omega))$. By Proposition 11.136, μ_{y_*} is a finite $\mathbb{K}$ -valued measure on $(P(\Omega), \mathcal{B}_B(P(\Omega)))$ and $\mathcal{P} \circ \mu_{y_*} \leq \|y_*\| \mathcal{P} \circ \mu$. Define $F_{y_*} : \Omega \rightarrow \mathbb{K}$ by $F_{y_*}(x) = \langle\langle y_*, F(x) \rangle\rangle$, $\forall x \in \Omega$. Clearly, $F_{y_*}(x) = \langle\langle y_*, F(x) \rangle\rangle = \langle\langle y_*, \mu(\{\bar{x} \in P(\Omega) \mid \bar{x} \leq x\}) \rangle\rangle = \mu_{y_*}(\{\bar{x} \in P(\Omega) \mid \bar{x} \leq x\})$, $\forall x \in \Omega$. Then, F_{y_*} is a cumulative distribution function of μ_{y_*} . By the first and the second steps, we have F_{y_*} is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. By Proposition 11.170, F is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. This completes the last step, and hence, the proof of the proposition. $\square$

Proposition 12.55 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle with the subset topology $\mathcal{O}$, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable, μ be a σ -finite $\mathcal{Y}$ -valued measure on $(P(\Omega), \mathcal{B}_B(P(\Omega)))$ such that $\mathcal{P} \circ \mu(r_{x_1, x_2}) < \infty$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$, and $F : \Omega \rightarrow \mathcal{Y}$ be the cumulative distribution function of μ with origin $x_0 \in \Omega$ when $\Omega \neq \emptyset$. Then, F is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable.*

Proof The result is trivial if $\Omega = \emptyset$. Consider the case $\Omega \neq \emptyset$. Since Ω is a rectangle, then, by Proposition 12.52, $P(\Omega)$ is also a rectangle and Ω is a region. Then, $\exists (\bar{x}_i)_{i=1}^\infty, (\tilde{x}_i)_{i=1}^\infty \subseteq \Omega$ such that $\bar{x}_i \leq x_0 \leq \tilde{x}_i$, $\forall i \in \mathbb{N}$, $\bigcup_{i=1}^\infty \bar{r}_{\bar{x}_i, \tilde{x}_i} = \Omega$, and $P(\Omega) = \bigcup_{i=1}^\infty r_{\bar{x}_i, \tilde{x}_i}$.

Fix any $i \in \mathbb{N}$, let $\bar{E}_i := \bar{r}_{\bar{x}_i, \tilde{x}_i} \in \mathcal{B}_B(\mathbb{R}^m)$ and $E_i := r_{\bar{x}_i, \tilde{x}_i} \in \mathcal{B}_B(\mathbb{R}^m)$. Note that $\mathcal{P} \circ \mu(E_i) < \infty$, by the assumption on μ . Let $(E_i, \mathcal{B}_{E_i}, \mu_{E_i})$ be the finite $\mathcal{Y}$ -valued measure subspace of $(P(\Omega), \mathcal{B}_B(P(\Omega)), \mu)$, $\mathcal{O}_{\bar{E}_i}$ be the subset topology on $\bar{E}_i \subseteq \Omega$, and $\bar{\mathcal{E}}_i := (\bar{E}_i, \mathcal{O}_{\bar{E}_i})$. Define $\bar{F} : \bar{E}_i \rightarrow \mathcal{Y}$ by $\bar{F}(x) = \mu_{E_i}(\{\bar{x} \in E_i \mid \bar{x} \leq x\})$, $\forall x \in \bar{E}_i$. By Proposition 12.51, $\bar{F}$ is a cumulative distribution function of μ_{E_i} and is of locally bounded variation. By Proposition 12.54, $\bar{F}$ is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. $\forall x \in \bar{E}_i$, let $\tilde{x} := x \wedge x_0 \in \bar{E}_i$ and $\hat{x} := x \vee x_0 \in \bar{E}_i$. Then, $\Delta_F(r_{\tilde{x}, \hat{x}}) = \mu(r_{\tilde{x}, \hat{x}}) = \mu_{E_i}(r_{\tilde{x}, \hat{x}}) = \Delta_{\bar{F}}(r_{\tilde{x}, \hat{x}})$, where the first equality follows from the fact that F is a cumulative distribution function of μ ; the second equality follows from Proposition 11.115; and the third equality follows from the fact that $\bar{F}$ is a cumulative distribution function of μ_{E_i} . If $\exists k_0 \in \bar{J} := \{1, \dots, m\}$ such that $\pi_{k_0}(x) = \pi_{k_0}(x_0)$, then $F(x) = \vartheta_{\mathcal{Y}} = \Delta_F(r_{\tilde{x}, \hat{x}})$. Otherwise, we have $(-1)^{n(x)} F(x) = \Delta_F(r_{\tilde{x}, \hat{x}})$, where $n(x) = \text{card}(\{k \in \bar{J} \mid \pi_k(x) < \pi_k(x_0)\})$. Hence, we have $F(x) = (-1)^{n(x)} \Delta_F(r_{\tilde{x}, \hat{x}}) = (-1)^{n(x)} \Delta_{\bar{F}}(r_{\tilde{x}, \hat{x}}) = (-1)^{n(x)} \sum_{\bar{j} \subseteq \bar{J}} (-1)^{\text{card}(\bar{J})} \bar{F}(\bar{x}_{\bar{j}})$, where $\bar{x}_{\bar{j}} \in \bar{E}_i$ is defined by $\pi_k(\bar{x}_{\bar{j}}) = \begin{cases} \pi_k(\tilde{x}) & \forall k \in \bar{J} \\ \pi_k(\hat{x}) & \forall k \in \bar{J} \setminus \bar{J} \end{cases}$. $\forall j \in \{0, \dots, 2^m - 1\}$, let $j_1 \dots j_m$ be the representation of j in base 2, $J_j := \{k \in \bar{J} \mid j_k = 0\}$, $n_j := \text{card}(J_j)$. Define $O_j := \{x \in$

$\bar{E}_i \mid \pi_k(x) > \pi_k(x_0), \forall k \in \bar{J} \setminus J_j, \pi_k(x) < \pi_k(x_0), \forall k \in J_j \} \in \mathcal{O}_{\bar{E}_i}$. Let $D := \{x \in \bar{E}_i \mid \pi_{i_0}(x) = \pi_{i_0}(x_0) \text{ for some } i_0 \in \bar{J}\} \in \mathcal{B}_B(\mathbb{R}^m)$. Clearly, $\bar{E}_i = (\bigcup_{j=0}^{2^m-1} O_j) \cup D$. $F|_D(x) = \vartheta_{\mathcal{Y}}, \forall x \in D$. Then, $F|_D$ is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. Fix any $j \in \{0, \dots, 2^m - 1\}$. Denote $J_j =: J$ and $n_j =: n$. Then, $F|_{O_j}(x) = (-1)^n \sum_{\hat{J} \subseteq J} (-1)^{\text{card}(\hat{J})} \bar{F}(\bar{x}_{\hat{J}}), \forall x \in O_j$. Note that $\pi_k(\bar{x}_{\hat{J}}) = \begin{cases} \pi_k(\tilde{x}) & \forall k \in \hat{J} \\ \pi_k(\hat{x}) & \forall k \in \bar{J} \setminus \hat{J} \end{cases} = \begin{cases} \pi_k(x) & \forall k \in \bar{J} \setminus (J \triangle \hat{J}) \\ \pi_k(x_0) & \forall k \in J \triangle \hat{J} \end{cases}$, $\forall k \in \bar{J}, \forall \hat{J} \subseteq \bar{J}$. We may define function $\tilde{F}_j : O_j \rightarrow \mathcal{Y}$ by $\tilde{F}_j(x) = \bar{F}(\bar{x}_{\hat{J}}), \forall x \in O_j$. Since $\bar{E}_i$ is a closed rectangle, it can be viewed as product of a closed set $E_{i,\delta}$ in the $J \triangle \hat{J}$ collection of coordinates and another closed set $E_{i,\sigma}$ in the $\bar{J} \setminus (J \triangle \hat{J})$ collection of coordinates. $\forall x \in \bar{E}_i, \pi_\delta(x) \in E_{i,\delta}$ gives the one coordinate and $\pi_\sigma(x) \in E_{i,\sigma}$ gives the other coordinate. Then, $x = (\pi_\delta(x), \pi_\sigma(x))$. For any open set $O_Y \in \mathcal{O}_{\mathcal{Y}}$, $\bar{F}_{\text{inv}}(O_Y) \in \mathcal{B}_B(\bar{E}_i) \subseteq \mathcal{B}_B(\mathbb{R}^m)$, since $\bar{F}$ is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. Define $A_1 := \bar{F}_{\text{inv}}(O_Y)_{\pi_\delta(x_0)} := \{s \in E_{i,\sigma} \mid (\pi_\delta(x_0), s) \in \bar{F}_{\text{inv}}(O_Y)\}$. By Proposition 12.25, $A_1 \in \mathcal{B}_B(E_{i,\sigma})$. By Proposition 11.24, $E_{i,\delta} \times A_1 \in \mathcal{B}_B(\bar{E}_i)$. Then, $\tilde{F}_{j\text{inv}}(O_Y) = O_j \cap (E_{i,\delta} \times A_1) \in \mathcal{B}_B(\bar{E}_i) \subseteq \mathcal{B}_B(\mathbb{R}^m)$. Hence, $\tilde{F}_j$ is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. Clearly, $F|_{O_j} = (-1)^n \sum_{\hat{J} \subseteq J} (-1)^{\text{card}(\hat{J})} \tilde{F}_j$. By Propositions 7.23, 11.39, and 11.38, $F|_{O_j}$ is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. By the arbitrariness of j and Proposition 11.41, $F|_{\bar{E}_i}$ is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. By the arbitrariness of i and Proposition 11.41, F is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. This completes the proof of the proposition. $\square$

Example 12.56 Let $m \in \mathbb{N}$, $\mathbb{R}$ be the second countable σ -finite real Banach measure space as defined in Example 11.28, $\mathbb{R}^m$ be the second countable σ -finite product topological measure space according to Proposition 12.32. Since $\mathbb{R}^m$ is a real Banach space, then $\mathbb{R}^m =: ((\mathbb{R}^m, \mathbb{R}, |\cdot|), \mathcal{B}_B(\mathbb{R}^m), \mu_{Bm})$ is a second countable σ -finite real Banach measure space. μ_{Bm} is the m -dimensional Borel measure. By Carathéodory Extension Theorem 11.19, μ_{Bm} induces the m -dimensional Lebesgue outer measure $\mu_{Lmo} : \mathbb{R}^m 2 \rightarrow [0, \infty] \subset \mathbb{R}_e$, which is defined by $\mu_{Lmo}(E) = \inf_{(A_i)_{i=1}^\infty \subseteq \mathcal{B}_B(\mathbb{R}^m), E \subseteq \bigcup_{i=1}^\infty A_i} \sum_{i=1}^\infty \mu_{Bm}(A_i), \forall E \subseteq \mathbb{R}^m$. $\mathcal{B}_{Lm} := \{E \subseteq \mathbb{R}^m \mid E \text{ is measurable with respect to } \mu_{Lmo}\}$ is the collection of m -dimensional Lebesgue measurable sets, and $\mu_{Lm} := \mu_{Lmo}|_{\mathcal{B}_{Lm}}$ is the m -dimensional Lebesgue measure. By Carathéodory Extension Theorem 11.19, $(\mathbb{R}^m, \mathcal{B}_{Lm}, \mu_{Lm})$ is the completion of the m -dimensional Borel measure space $(\mathbb{R}^m, \mathcal{B}_B(\mathbb{R}^m), \mu_{Bm})$, and is said to be the m -dimensional Lebesgue measure space. When $m = 1$, $\mu_{Bm}, \mu_{Lmo}, \mu_{Lm}$, and $\mathcal{B}_{Lm}$ are also denoted by μ_B, μ_{Lo}, μ_L , and $\mathcal{B}_L$, respectively.

$\forall E \in \mathbb{R}^m 2$, we will show that

$$\mu_{Lmo}(E) = \inf_{(O_i)_{i=1}^\infty \subseteq \mathcal{O}_{\mathbb{R}^m}, E \subseteq \bigcup_{i=1}^\infty O_i} \sum_{i=1}^\infty \mu_{Bm}(O_i) =: \bar{\mu}_o(E)$$

Since $\mathcal{O}_{\mathbb{R}^m} \subseteq \mathcal{B}_B(\mathbb{R}^m)$, then $\mu_{Lmo}(E) \leq \bar{\mu}_o(E)$. On the other hand, if $\mu_{Lmo}(E) = \infty$, then $\infty = \mu_{Lmo}(E) \leq \bar{\mu}_o(E) \leq \infty$ and $\mu_{Lmo}(E) = \bar{\mu}_o(E)$. If $\mu_{Lmo}(E) < \infty$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists (A_i)_{i=1}^\infty \subseteq \mathcal{B}_B(\mathbb{R}^m)$ such that $E \subseteq \bigcup_{i=1}^\infty A_i$ and $\mu_{Lmo}(E) \leq \sum_{i=1}^\infty \mu_{Bm}(A_i) < \mu_{Lmo}(E) + \epsilon/2$. $\forall i \in \mathbb{N}$, since $\mathbb{R}^m$ is a topological measure space, then $\exists O_i \in \mathcal{O}_{\mathbb{R}^m}$ with $A_i \subseteq O_i$ such that $\mu_{Bm}(O_i \setminus A_i) < 2^{-i-1}\epsilon$. Then, $E \subseteq \bigcup_{i=1}^\infty O_i$ and $\bar{\mu}_o(E) \leq \sum_{i=1}^\infty \mu_{Bm}(O_i) = \sum_{i=1}^\infty (\mu_{Bm}(A_i) + \mu_{Bm}(O_i \setminus A_i)) < \mu_{Lmo}(E) + \epsilon/2 + \epsilon/2 = \mu_{Lmo}(E) + \epsilon$. By the arbitrariness of ϵ , we have $\bar{\mu}_o(E) \leq \mu_{Lmo}(E)$. Hence, $\mu_{Lmo}(E) = \bar{\mu}_o(E)$.

By the translational invariance of $\mathbb{R}$, we have $\mathbb{R}^m$ is translational invariant. Then, by the above, μ_{Lmo} and μ_{Lm} are translational invariant. $\diamond$

Definition 12.57 (Lebesgue-Stieltjes Integral) Let $m \in \mathbb{N}$, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a region with the subset topology, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $F : \Omega \rightarrow \mathcal{Y}$ be of locally bounded variation, $(P(\Omega), \mathcal{B}_B(P(\Omega)), \mu)$ be the unique σ -finite $\mathcal{Y}$ -valued measure space that admits F as a cumulative distribution function as prescribed in Theorem 12.50, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $U \in \mathcal{B}_B(P(\Omega))$, and $g : U \rightarrow B(\mathcal{Y}, \mathcal{Z})$ be $\mathcal{B}_B(P(\Omega))$ -measurable (or $\mathcal{B}_B(\mathbb{R}^m)$ -measurable). We will denote the integral $\int_U g(x) d\mu(x)$ by $\int_U g dF$ or $\int_U g(x) dF(x)$, whenever the first integral makes sense.

12.5 Absolute and Lipschitz Continuity

Definition 12.58 Let $I \subseteq \mathbb{R}$ be an interval and $\mathcal{Y}$ be a normed linear space. A function $f : I \rightarrow \mathcal{Y}$ is said to be absolutely continuous at $x_0 \in I$ if $\exists a \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta(\epsilon) \in (0, a] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall x_0 - a \leq x_1 \leq \bar{x}_1 \leq x_2 \leq \bar{x}_2 \leq \dots \leq x_n \leq \bar{x}_n \leq x_0 + a$ with $x_i, \bar{x}_i \in I$, $i = 1, \dots, n$, and $\sum_{i=1}^n \mu_B(r_{x_i, \bar{x}_i}) = \sum_{i=1}^n (\bar{x}_i - x_i) < \delta(\epsilon)$, we have $\sum_{i=1}^n \|\Delta_F(r_{x_i, \bar{x}_i})\| = \sum_{i=1}^n \|f(\bar{x}_i) - f(x_i)\| < \epsilon$. We will say that f is absolutely continuous if it is absolutely continuous at any $x_0 \in I$. We will say f is absolutely continuous on $E \subseteq I$ if it is absolutely continuous at any $x_0 \in E$.

Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, and $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$. $F : \Omega \rightarrow \mathcal{Y}$ is said to be absolutely continuous at $x_0 \in \Omega$ if the following conditions holds.

- (i) When $m > 1$, $\forall i_0 \in \bar{J} := \{1, \dots, m\}$, let $D_{i_0} := \{x \in \mathbb{R}^{m-1} \mid \exists \bar{x} \in \Omega \ni \cdot \pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0) \text{ and } \pi_{\hat{i}_0}(\bar{x}) = x\}$, where $\pi_{\hat{i}_0} : \mathbb{R}^m \rightarrow \mathbb{R}^{m-1}$ is defined by $\pi_{\hat{i}_0}(x) = (\pi_1(x), \dots, \pi_{i_0-1}(x), \pi_{i_0+1}(x), \dots, \pi_m(x)) \in \mathbb{R}^{m-1}$, $\forall x \in \mathbb{R}^m$, and $F_{i_0} : D_{i_0} \rightarrow \mathcal{Y}$ be defined by $F_{i_0}(x) = F(\bar{x})$, $\forall x \in D_{i_0}$, $\forall \bar{x} \in \Omega$ with $\pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0)$ and $\pi_{\hat{i}_0}(\bar{x}) = x$. The function F_{i_0} is absolutely continuous at $\pi_{\hat{i}_0}(x_0)$.

(ii) $\exists a \in (0, \infty) \subset \mathbb{R}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R},$
 $\forall n \in \mathbb{N}, \forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{r}_{x_0 - a\mathbf{1}_m, x_0 + a\mathbf{1}_m} \cap \Omega$ with $\bar{x}_i \leq \tilde{x}_i, \forall i \in$
 $\{1, \dots, n\}, (r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) <$
 $\delta(\epsilon),$ we have $\sum_{i=1}^n \|\Delta_F(r_{\bar{x}_i, \tilde{x}_i})\| < \epsilon.$

F is said to be absolutely continuous if it is absolutely continuous at any $x_0 \in \Omega$. F is said to be absolutely continuous on $E \subseteq \Omega$ if it is absolutely continuous at any $x_0 \in E$.

In the above definition, the definition for general $\mathbb{R}^m$ case subsumes the definition for $\mathbb{R}$ case. We explicitly specify the $\mathbb{R}$ case for its direct implications.

Proposition 12.59 *Let $m \in \mathbb{N}, \mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $F : \Omega \rightarrow \mathcal{Y}$ be absolutely continuous at $x_0 \in \Omega$. Then, F is continuous at x_0 .*

Proof We will prove the result using mathematical induction on m .

1° $m = 1$. By Definition 12.58, F is continuous at x_0 . This case is proved.

2° Assume that the result holds for $m \leq k - 1 \in \mathbb{N}$.

3° Consider the case when $m = k \in \{2, 3, \dots\}$. By (i) of Definition 12.58, $\forall i_0 \in \bar{J} := \{1, \dots, m\}, F_{i_0} : D_{i_0} \rightarrow \mathcal{Y}$ is absolutely continuous at $\pi_{i_0}^\wedge(x_0)$, where F_{i_0} and D_{i_0} are as defined in Definition 12.58. Clearly, $D_{i_0} \in \mathcal{B}_B(\mathbb{R}^{m-1})$ is a rectangle since Ω is a rectangle. By the inductive assumption, F_{i_0} is continuous at $\pi_{i_0}^\wedge(x_0)$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta_{i_0}(\epsilon) \in (0, \infty) \subset \mathbb{R}$ such that $\|F_{i_0}(x) - F_{i_0}(\pi_{i_0}^\wedge(x_0))\| < \epsilon, \forall x \in \mathcal{B}_{\mathbb{R}^{m-1}}(\pi_{i_0}^\wedge(x_0), \delta_{i_0}(\epsilon)) \cap D_{i_0}.$

By (ii) of Definition 12.58, $\exists a \in (0, \infty) \subset \mathbb{R}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \bar{\delta}(\epsilon) \in (0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N}, \forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{r}_{x_0 - a\mathbf{1}_m, x_0 + a\mathbf{1}_m} \cap \Omega$ with $\bar{x}_i \leq \tilde{x}_i, \forall i \in \{1, \dots, n\}, (r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) < \bar{\delta}(\epsilon),$ we have $\sum_{i=1}^n \|\Delta_F(r_{\bar{x}_i, \tilde{x}_i})\| < \epsilon.$ Take $\delta(\epsilon) := \min\{a, \sqrt[m]{\bar{\delta}(2^{-m}\epsilon)}, \delta_1(2^{-m}\epsilon), \dots, \delta_m(2^{-m}\epsilon)\} \in (0, \infty) \subset \mathbb{R}.$ $\forall x \in \mathcal{B}_{\mathbb{R}^m}(x_0, \delta(\epsilon)) \cap \Omega,$ we have $\|F(x) - F(x_0)\| \leq \|\Delta_G(r_{\tilde{x}, \hat{x}})\| + \sum_{J \subset \bar{J}} \|F(x_J) - F(x_0)\|,$ where $\tilde{x} := x \wedge x_0$ and $\hat{x} := x \vee x_0, G : \Omega \rightarrow \mathcal{Y}$ is defined by $G(x) = F(x) - F(x_0), \forall x \in \Omega,$ and $x_J \in \bar{r}_{\tilde{x}, \hat{x}}$ is defined by $\pi_i(x_J) = \begin{cases} \pi_i(x) & \forall i \in J \\ \pi_i(x_0) & \forall i \in \bar{J} \setminus J \end{cases}, \forall J \subset \bar{J}.$

Clearly, $\tilde{x}, \hat{x} \in \bar{r}_{x_0 - a\mathbf{1}_m, x_0 + a\mathbf{1}_m} \cap \Omega$ since $a > |x - x_0| \geq \max\{|\tilde{x} - x_0|, |\hat{x} - x_0|\} \geq 0.$ Since $\mu_{Bm}(r_{\tilde{x}, \hat{x}}) \leq |x - x_0|^m < \bar{\delta}(2^{-m}\epsilon),$ then $\|\Delta_F(r_{\tilde{x}, \hat{x}})\| = \|\Delta_G(r_{\tilde{x}, \hat{x}})\| < 2^{-m}\epsilon. \forall J \subset \bar{J},$ fix any $i_0 \in \bar{J} \setminus J,$ we have $\|F(x_J) - F(x_0)\| = \|F_{i_0}(\pi_{i_0}^\wedge(x_J)) - F_{i_0}(\pi_{i_0}^\wedge(x_0))\|.$ Clearly, $\pi_{i_0}^\wedge(x_J) \in D_{i_0}$ and $|\pi_{i_0}^\wedge(x_J) - \pi_{i_0}^\wedge(x_0)| = |x_J - x_0| \leq |x - x_0| < \delta_{i_0}(2^{-m}\epsilon).$ Then, $\|F(x_J) - F(x_0)\| = \|F_{i_0}(\pi_{i_0}^\wedge(x_J)) - F_{i_0}(\pi_{i_0}^\wedge(x_0))\| < 2^{-m}\epsilon.$ Hence, $\|F(x) - F(x_0)\| < \epsilon.$ This shows that F is continuous at $x_0.$

This completes the induction process. This completes the proof of the proposition. $\square$

Proposition 12.60 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $F : \Omega \rightarrow \mathcal{Y}$, and $x_0 \in \Omega$. Then, F is absolutely continuous at x_0 if, and only if,*

- (a) $\forall i \in \{0, \dots, m-1\}$, $\forall J \subset \bar{J} := \{1, \dots, m\}$ with $\text{card}(J) = i$, let $\pi_J : \mathbb{R}^m \rightarrow \mathbb{R}^i$ be such that $\pi_J(\bar{x}) = (\pi_l(\bar{x}))_{l \in J}$, $\forall \bar{x} \in \mathbb{R}^m$, $\pi_{\bar{J} \setminus J} : \mathbb{R}^m \rightarrow \mathbb{R}^{m-i}$ be such that $\pi_{\bar{J} \setminus J}(\bar{x}) = (\pi_l(\bar{x}))_{l \in \bar{J} \setminus J}$, $\forall \bar{x} \in \mathbb{R}^m$, $M : \mathbb{R}^i \times \mathbb{R}^{m-i} \rightarrow \mathbb{R}^m$ be such that $M(x_a, x_b) = \bar{x}$ with $x_a = \pi_J(\bar{x})$ and $x_b = \pi_{\bar{J} \setminus J}(\bar{x})$, $\forall x_a \in \mathbb{R}^i$ and $\forall x_b \in \mathbb{R}^{m-i}$, $x_f := \pi_J(x_0)$, $\tilde{x}_f := \pi_{\bar{J} \setminus J}(x_0)$, $\tilde{\Omega}_J := \pi_{\bar{J} \setminus J}(\Omega)$, and $F_J : \tilde{\Omega}_J \rightarrow \mathcal{Y}$ be defined by $F_J(\bar{x}) = F(M(x_f, \bar{x}))$, $\forall \bar{x} \in \tilde{\Omega}_J$, $\exists a_J \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_J(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_j)_{j=1}^n, (\tilde{x}_j)_{j=1}^n \subseteq \bar{\mathbf{r}}_{\tilde{x}_f - a_J \mathbf{1}_{m-i}, \tilde{x}_f + a_J \mathbf{1}_{m-i}} \cap \tilde{\Omega}_J$ with $\bar{x}_j \leq \tilde{x}_j$, $\forall j \in \{1, \dots, n\}$, $(\mathbf{r}_{\bar{x}_j, \tilde{x}_j})_{j=1}^n$ being pairwise disjoint, and $\sum_{j=1}^n \mu_{B(m-i)}(\mathbf{r}_{\bar{x}_j, \tilde{x}_j}) < \delta_J(\epsilon)$, we have $\sum_{j=1}^n \|\Delta_{F_J}(\mathbf{r}_{\bar{x}_j, \tilde{x}_j})\| < \epsilon$.

Proof We will prove the result using mathematical induction on m .
 1° The result is clearly a restatement of Definition 12.58. Hence, the case is proved.

2° Assume that the result holds for $m \leq k-1 \in \mathbb{N}$.

3° Consider the case when $m = k \in \{2, 3, \dots\}$.

“Sufficiency” We need to show that F is absolutely continuous at x_0 assuming (a). Let $i = 0$, (a) implies (ii) of Definition 12.58. $\forall i_0 \in \bar{J}$, let π_{i_0} , $\pi_{\hat{i}_0}$, D_{i_0} , and F_{i_0} be as defined in (i) of Definition 12.58. By (a) holds for any $i \in \{1, \dots, m-1\}$ and any $J \subset \bar{J}$ with $\text{card}(J) = i$ such that $i_0 \in J$, we have (a) holds for F_{i_0} at $\pi_{\hat{i}_0}(x_0)$. By the inductive assumption, F_{i_0} is absolutely continuous at $\pi_{\hat{i}_0}(x_0)$. Hence, (i) of Definition 12.58 holds for F at x_0 . Then, F is absolutely continuous at x_0 .

“Necessity” We need to show that (a) holds assuming that F is absolutely continuous at x_0 . By (ii) of Definition 12.58, (a) holds when $i = 0$. By (i) of Definition 12.58, $\forall i_0 \in \bar{J}$, let π_{i_0} , $\pi_{\hat{i}_0}$, D_{i_0} , and F_{i_0} be as defined in (i) of Definition 12.58, we have F_{i_0} is absolutely continuous at $\pi_{\hat{i}_0}(x_0)$. By the inductive assumption, (a) holds for F_{i_0} . Then, (a) holds for F , $\forall i \in \{1, \dots, m-1\}$, $\forall J \subset \bar{J}$ with $\text{card}(J) = i$ and $i_0 \in J$. Hence, (a) holds for F . This completes the proof for the necessity.

This completes the induction process.

This completes the proof of the proposition. $\square$

Proposition 12.61 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $\mathcal{Y}$ and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $f : \Omega \rightarrow \mathcal{Y}$, $g : \Omega \rightarrow \mathcal{Z}$, $x_0 \in \Omega$, and $h : \Omega \rightarrow \mathcal{Y} \times \mathcal{Z}$ be defined by $h(x) = (f(x), g(x))$, $\forall x \in \Omega$. Then, h is*

absolutely continuous at x_0 if, and only if, f and g are absolutely continuous at x_0 .

Proof We will prove the result using mathematical induction on m .
 1° $m = 1$. “Sufficiency” By f being absolutely continuous at x_0 , $\exists a_f \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_f(\epsilon) \in (0, a_f] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall x_0 - a_f \leq x_1 \leq \bar{x}_1 \leq x_2 \leq \bar{x}_2 \leq \dots \leq x_n \leq \bar{x}_n \leq x_0 + a_f$ with $x_i, \bar{x}_i \in I$, $i = 1, \dots, n$, and $\sum_{i=1}^n (\bar{x}_i - x_i) < \delta_f(\epsilon)$, we have $\sum_{i=1}^n \|f(\bar{x}_i) - f(x_i)\| < \epsilon$. By g being absolutely continuous at x_0 , $\exists a_g \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_g(\epsilon) \in (0, a_g] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall x_0 - a_g \leq x_1 \leq \bar{x}_1 \leq x_2 \leq \bar{x}_2 \leq \dots \leq x_n \leq \bar{x}_n \leq x_0 + a_g$ with $x_i, \bar{x}_i \in I$, $i = 1, \dots, n$, and $\sum_{i=1}^n (\bar{x}_i - x_i) < \delta_g(\epsilon)$, we have $\sum_{i=1}^n \|g(\bar{x}_i) - g(x_i)\| < \epsilon$. Let $a := \min\{a_f, a_g\} \in (0, +\infty) \subset \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta(\epsilon) := \min\{\delta_f(\epsilon/2), \delta_g(\epsilon/2)\} \in (0, a] \subset \mathbb{R}$. $\forall n \in \mathbb{N}$, $\forall x_0 - a \leq x_1 \leq \bar{x}_1 \leq x_2 \leq \bar{x}_2 \leq \dots \leq x_n \leq \bar{x}_n \leq x_0 + a$ with $x_i, \bar{x}_i \in I$, $i = 1, \dots, n$, and $\sum_{i=1}^n (\bar{x}_i - x_i) < \delta(\epsilon)$, we have $\sum_{i=1}^n \|h(\bar{x}_i) - h(x_i)\| = \sum_{i=1}^n \|(f(\bar{x}_i) - f(x_i), g(\bar{x}_i) - g(x_i))\| = \sum_{i=1}^n (\|f(\bar{x}_i) - f(x_i)\|^2 + \|g(\bar{x}_i) - g(x_i)\|^2)^{1/2} \leq \sum_{i=1}^n (\|f(\bar{x}_i) - f(x_i)\| + \|g(\bar{x}_i) - g(x_i)\|) < \epsilon$. Hence, h is absolutely continuous at x_0 .

“Necessity” By h being absolutely continuous at x_0 , $\exists a \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta(\epsilon) \in (0, a] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall x_0 - a \leq x_1 \leq \bar{x}_1 \leq x_2 \leq \bar{x}_2 \leq \dots \leq x_n \leq \bar{x}_n \leq x_0 + a$ with $x_i, \bar{x}_i \in I$, $i = 1, \dots, n$, and $\sum_{i=1}^n (\bar{x}_i - x_i) < \delta(\epsilon)$, we have $\epsilon > \sum_{i=1}^n \|h(\bar{x}_i) - h(x_i)\| = \sum_{i=1}^n (\|f(\bar{x}_i) - f(x_i)\|^2 + \|g(\bar{x}_i) - g(x_i)\|^2)^{1/2} \geq \sum_{i=1}^n \|f(\bar{x}_i) - f(x_i)\|$. Hence, f is absolutely continuous at x_0 . By symmetry, g is absolutely continuous at x_0 . This case is proved.

2° Assume that the result holds for $m \leq k - 1 \in \mathbb{N}$.

3° Consider the case $m = k \in \{2, 3, \dots\}$.

“Sufficiency” Assume that f and g are absolutely continuous at x_0 . $\forall i_0 \in \bar{J} := \{1, \dots, m\}$, let $D_{i_0} \subseteq \mathbb{R}^{m-1}$, $\pi_{i_0}^\wedge : \mathbb{R}^m \rightarrow \mathbb{R}^{m-1}$, $f_{i_0} : D_{i_0} \rightarrow \mathcal{Y}$, $g_{i_0} : D_{i_0} \rightarrow \mathcal{Z}$, and $h_{i_0} : D_{i_0} \rightarrow \mathcal{Y} \times \mathcal{Z}$ be defined as in Definition 12.58. By (i) of Definition 12.58, f_{i_0} and g_{i_0} are absolutely continuous at $\pi_{i_0}^\wedge(x_0)$. By the inductive assumption, h_{i_0} is absolutely continuous at $\pi_{i_0}^\wedge(x_0)$. Then, h satisfies (i) of Definition 12.58.

By (ii) of Definition 12.58, $\exists a_f \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_f(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{x}_0 - a_f \mathbf{1}_m, x_0 + a_f \mathbf{1}_m \cap \Omega$ with $\bar{x}_i \leq \tilde{x}_i$, $\forall i \in \{1, \dots, n\}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) < \delta_f(\epsilon)$, we have $\sum_{i=1}^n \|\Delta_f(r_{\bar{x}_i, \tilde{x}_i})\| < \epsilon$; $\exists a_g \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_g(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{x}_0 - a_g \mathbf{1}_m, x_0 + a_g \mathbf{1}_m \cap \Omega$ with $\bar{x}_i \leq \tilde{x}_i$, $\forall i \in \{1, \dots, n\}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) < \delta_g(\epsilon)$, we have $\sum_{i=1}^n \|\Delta_g(r_{\bar{x}_i, \tilde{x}_i})\| < \epsilon$. Let $a := a_f \wedge a_g \in (0, \infty) \subset \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta(\epsilon) := \min\{\delta_f(\epsilon/2), \delta_g(\epsilon/2)\} \in (0, \infty) \subset \mathbb{R}$. $\forall n \in \mathbb{N}$, $\forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{x}_0 - a \mathbf{1}_m, x_0 + a \mathbf{1}_m \cap \Omega$ with $\bar{x}_i \leq \tilde{x}_i$, $\forall i \in \{1, \dots, n\}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) < \delta(\epsilon)$, we have $\sum_{i=1}^n \|\Delta_h(r_{\bar{x}_i, \tilde{x}_i})\| =$

$\sum_{i=1}^n \|(\Delta_f(r_{\bar{x}_i, \tilde{x}_i}), \Delta_g(r_{\bar{x}_i, \tilde{x}_i}))\| = \sum_{i=1}^n (\|\Delta_f(r_{\bar{x}_i, \tilde{x}_i})\|^2 + \|\Delta_g(r_{\bar{x}_i, \tilde{x}_i})\|^2)^{1/2} \leq \sum_{i=1}^n \|\Delta_f(r_{\bar{x}_i, \tilde{x}_i})\| + \sum_{i=1}^n \|\Delta_g(r_{\bar{x}_i, \tilde{x}_i})\| < \epsilon$. Hence, h satisfies (ii) of Definition 12.58. Then, h is absolutely continuous at x_0 .

“Necessity” Let h be absolutely continuous at x_0 . $\forall i_0 \in \bar{J}$, let D_{i_0} , $\pi_{\hat{i}_0}$, f_{i_0} , g_{i_0} , and h_{i_0} be defined as in Definition 12.58. By (i) of Definition 12.58, h_{i_0} is absolutely continuous at $\pi_{\hat{i}_0}(x_0)$. By the inductive assumption, f_{i_0} and g_{i_0} are absolutely continuous at x_0 . Hence, f and g satisfy (i) of Definition 12.58.

By (ii) of Definition 12.58, $\exists a \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{r}_{x_0-a\mathbf{1}_m, x_0+a\mathbf{1}_m} \cap \Omega$ with $\bar{x}_i \leq \tilde{x}_i$, $\forall i \in \{1, \dots, n\}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) < \delta(\epsilon)$, we have $\epsilon > \sum_{i=1}^n \|\Delta_h(r_{\bar{x}_i, \tilde{x}_i})\| \geq \sum_{i=1}^n \|\Delta_f(r_{\bar{x}_i, \tilde{x}_i})\|$. Hence, f satisfies (ii) of Definition 12.58. Then, f is absolutely continuous at x_0 . By symmetry, g is absolutely continuous at x_0 .

This completes the induction process. This completes the proof of the proposition. $\square$

Proposition 12.62 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $F : \Omega \rightarrow \mathcal{Y}$, and $\Omega_1, \Omega_2 \subseteq \Omega$ be rectangles in $\mathbb{R}^m$ such that $\Omega_1 \cup \Omega_2 = \Omega$. Then, the following statements hold.*

(i) *If F is absolutely continuous on $E \subseteq \Omega$, then $F|_{\Omega_1}$ is absolutely continuous on $E \cap \Omega_1$.*

(ii) *If $F|_{\Omega_1}$ and $F|_{\Omega_2}$ are absolutely continuous and Ω_1 and Ω_2 are both relatively open or both relatively closed (with respect to the subset topology $\mathcal{O}$), then F is absolutely continuous.*

Proof (i) $\forall x_0 \in E \cap \Omega_1$, since F is absolutely continuous at x_0 , then, $F|_{\Omega_1}$ is absolutely continuous at x_0 . By the arbitrariness of x_0 , $F|_{\Omega_1}$ is absolutely continuous on $E \cap \Omega_1$.

(ii) We will distinguish two exhaustive cases: Case 1: Ω_1 and Ω_2 are relatively open in Ω ; Case 2: Ω_1 and Ω_2 are relatively closed in Ω .

Case 1: Ω_1 and Ω_2 are relatively open in Ω . Fix any $x_0 \in \Omega$. We will prove the statement that “if $F|_{\Omega_i}$ is absolutely continuous at x_0 when $x_0 \in \Omega_i$, and Ω_1 and Ω_2 are relatively open in Ω , then F is absolutely continuous at x_0 ” by mathematical induction on m .

1° $m = 1$. Without loss of generality, assume $x_0 \in \Omega_1$. Since Ω_1 is open relative to Ω , then $\exists a \in (0, \infty) \subset \mathbb{R}$ such that $(x_0 - a, x_0 + a) \cap \Omega = (x_0 - a, x_0 + a) \cap \Omega$. By $F|_{\Omega_1}$ being absolutely continuous at x_0 , we can easily show that that F is absolutely continuous at x_0 . This completes the first step in the induction process.

2° Assume that the result holds for $m \leq k - 1 \in \mathbb{N}$.

3° Consider the case when $m = k \in \{2, 3, \dots\}$. Without loss of generality, assume $x_0 \in \Omega_1$. $\forall i_0 \in \bar{J} := \{1, \dots, m\}$, let D_{i_0} and $F_{i_0} : D_{i_0} \rightarrow \mathcal{Y}$ be as

defined in Definition 12.58. Let $D_{i_0,j} := \{x \in \mathbb{R}^{m-1} \mid \exists \bar{x} \in \Omega_j \ni \pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0) \text{ and } \pi_{\hat{i}_0}(\bar{x}) = x\}$ and $F_{i_0,j} : D_{i_0,j} \rightarrow \mathcal{Y}$ be defined by $F_{i_0,j}(x) = F|_{\Omega_j}(\bar{x})$, $\forall x \in D_{i_0,j}$, where $\bar{x} \in \Omega_j$ is such that $\pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0)$ and $\pi_{\hat{i}_0}(\bar{x}) = x$, $j = 1, 2$. Fix any $j \in \{1, 2\}$. Since $\Omega_j \subseteq \Omega$ is a relatively open rectangle, then $D_{i_0,j} \subseteq D_{i_0}$ is a relatively open rectangle. By $F|_{\Omega_j}$ being absolutely continuous at x_0 when $x_0 \in \Omega_j$, we have $F_{i_0,j}$ is absolutely continuous at $\pi_{\hat{i}_0}(x_0)$ when $\pi_{\hat{i}_0}(x_0) \in D_{i_0,j}$. It is easy to see that $F_{i_0,j} = F_{i_0}|_{D_{i_0,j}}$ and $D_{i_0,1} \cup D_{i_0,2} = D_{i_0}$. By the inductive assumption, F_{i_0} is absolutely continuous at $\pi_{\hat{i}_0}(x_0) \in D_{i_0}$. Hence, F satisfies (i) of Definition 12.58.

Since Ω_1 is relative open in Ω , then $\exists a \in (0, \infty) \subset \mathbb{R}$ such that $r_{x_0-a\mathbf{1}_m, x_0+a\mathbf{1}_m}^\circ \cap \Omega_1 = r_{x_0-a\mathbf{1}_m, x_0+a\mathbf{1}_m}^\circ \cap \Omega$. By $F|_{\Omega_1}$ being absolutely continuous at x_0 and (ii) of Definition 12.58, $\exists a_F \in (0, a) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{r}_{x_0-a_F\mathbf{1}_m, x_0+a_F\mathbf{1}_m} \cap \Omega_1 = \bar{r}_{x_0-a_F\mathbf{1}_m, x_0+a_F\mathbf{1}_m} \cap \Omega$ with $\bar{x}_i \leq \tilde{x}_i$, $\forall i \in \{1, \dots, n\}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) < \delta(\epsilon)$, we have $\sum_{i=1}^n \|\Delta_F(r_{\bar{x}_i, \tilde{x}_i})\| < \epsilon$. Hence, F satisfies (ii) of Definition 12.58.

Hence, F is absolutely continuous at x_0 . This completes the induction process.

By the arbitrariness of x_0 , F is absolutely continuous. This completes the proof for Case 1.

Case 2: Ω_1 and Ω_2 are relatively closed in Ω . Fix any $x_0 \in \Omega$. We will prove the statement that “if $F|_{\Omega_i}$ is absolutely continuous at x_0 when $x_0 \in \Omega_i$, and Ω_1 and Ω_2 are relatively closed in Ω , then F is absolutely continuous at x_0 ” by mathematical induction on m .

1° $m = 1$. Without loss of generality, assume $x_0 \in \Omega_1$. We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: $x_0 \notin \Omega_2$; Case 2b: $x_0 \in \Omega_2$. Case 2a: $x_0 \notin \Omega_2$. Since Ω_2 is relatively closed in Ω and $x_0 \in \Omega \setminus \Omega_2$, then, by Proposition 4.10, $\exists a \in (0, \infty) \subset \mathbb{R}$ such that $(x_0 - a, x_0 + a) \cap \Omega \subseteq \Omega \setminus \Omega_2 \subseteq \Omega_1$. $F|_{\Omega_1}$ is absolutely continuous at x_0 implies that F is absolutely continuous at x_0 . This subcase is proved. Case 2b: $x_0 \in \Omega_2$. Then, $x_0 \in \Omega_1 \cap \Omega_2$. By $F|_{\Omega_i}$ being absolutely continuous at x_0 , $\exists a_i \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_i(\epsilon) \in (0, a_i] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall x_0 - a_i \leq x_1 \leq \bar{x}_1 \leq \dots \leq x_n \leq \bar{x}_n \leq x_0 + a_i$ with $x_j, \bar{x}_j \in \Omega_i$, $j = 1, \dots, n$, and $\sum_{j=1}^n (\bar{x}_j - x_j) < \delta_i(\epsilon)$, we have $\sum_{j=1}^n \|F(\bar{x}_j) - F(x_j)\| < \epsilon$, $i = 1, 2$. Take $a := \min\{a_1, a_2\} \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta(\epsilon) := \min\{\delta_1(\epsilon/2), \delta_2(\epsilon/2)\} \in (0, a] \subset \mathbb{R}$. $\forall n \in \mathbb{N}$, $\forall x_0 - a \leq x_1 \leq \bar{x}_1 \leq \dots \leq x_n \leq \bar{x}_n \leq x_0 + a$ with $x_j, \bar{x}_j \in \Omega$, $j = 1, \dots, n$, and $\sum_{j=1}^n (\bar{x}_j - x_j) < \delta(\epsilon)$. Without loss of generality, assume $\exists j_0 \in \{1, \dots, n-1\}$ such that $\bar{x}_{j_0} \leq x_0 \leq x_{j_0+1}$. (In case j_0 does not exist, add a pair of x_0 's into the collection of x_j 's and $\bar{x}_j$'s.) Then, $x_1 \in \Omega = \Omega_1 \cup \Omega_2$. Without loss of generality, assume $x_1 \in \Omega_1$. Then $x_1, \bar{x}_1, \dots, x_{j_0}, \bar{x}_{j_0} \in \Omega_1$, since Ω_1 is an rectangle (interval, since $m = 1$), and $\sum_{j=1}^{j_0} (\bar{x}_j - x_j) < \delta(\epsilon) \leq \delta_1(\epsilon/2)$. Then, $\sum_{j=1}^{j_0} \|F(\bar{x}_j) - F(x_j)\| < \epsilon/2$. By a similar argument, we have

$\sum_{j=j_0+1}^n \|F(\bar{x}_j) - F(x_j)\| < \epsilon/2$. Then, $\sum_{j=1}^n \|F(\bar{x}_j) - F(x_j)\| < \epsilon$. Hence, F is absolutely continuous at x_0 . This subcase is proved.

Hence, in both subcases, F is absolutely continuous at x_0 . This completes the proof for the first step of the induction process.

2° Assume that the result holds for $m \leq k-1 \in \mathbb{N}$.

3° Consider the case when $m = k \in \{2, 3, \dots\}$. Without loss of generality, assume $x_0 \in \Omega_1$. $\forall i_0 \in \bar{J}$, let D_{i_0} and $F_{i_0} : D_{i_0} \rightarrow \mathcal{Y}$ be as defined in Definition 12.58. Let $D_{i_0,j} := \{x \in \mathbb{R}^{m-1} \mid \exists \bar{x} \in \Omega_j \ni \cdot \pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0) \text{ and } \pi_{\hat{i}_0}(\bar{x}) = x\}$ and $F_{i_0,j} : D_{i_0,j} \rightarrow \mathcal{Y}$ be defined by $F_{i_0,j}(x) = F|_{\Omega_j}(\bar{x})$, $\forall x \in D_{i_0,j}$, where $\bar{x} \in \Omega_j$ is such that $\pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0)$ and $\pi_{\hat{i}_0}(\bar{x}) = x$, $j = 1, 2$. Fix any $j \in \{1, 2\}$. Since $\Omega_j \subseteq \Omega$ is a relatively closed rectangle, then $D_{i_0,j} \subseteq D_{i_0}$ is a relatively closed rectangle. By $F|_{\Omega_j}$ being absolutely continuous at x_0 when $x_0 \in \Omega_j$, we have $F_{i_0,j}$ is absolutely continuous at $\pi_{\hat{i}_0}(x_0)$ when $\pi_{\hat{i}_0}(x_0) \in D_{i_0,j}$. It is easy to see that $F_{i_0,j} = F_{i_0}|_{D_{i_0,j}}$ and $D_{i_0,1} \cup D_{i_0,2} = D_{i_0}$. By the inductive assumption, F_{i_0} is absolutely continuous at $\pi_{\hat{i}_0}(x_0) \in D_{i_0}$. Hence, F satisfies (i) of Definition 12.58.

We will further distinguish two exhaustive and mutually exclusives subcases: Case 2 α : $x_0 \notin \Omega_2$; Case 2 β : $x_0 \in \Omega_2$.

Case 2 α : $x_0 \notin \Omega_2$. Then, $x_0 \in \Omega \setminus \Omega_2$. Since Ω_2 is relatively closed in Ω , by Proposition 4.10, $\exists a \in (0, \infty) \subset \mathbb{R}$ such that $r_{x_0-a\mathbf{1}_m, x_0+a\mathbf{1}_m}^o \cap \Omega \subseteq \Omega \setminus \Omega_2 \subseteq \Omega_1$. By $F|_{\Omega_1}$ being absolutely continuous at x_0 , $\exists \bar{a} \in (0, a) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{r}_{x_0-\bar{a}\mathbf{1}_m, x_0+\bar{a}\mathbf{1}_m} \cap \Omega_1 = \bar{r}_{x_0-\bar{a}\mathbf{1}_m, x_0+\bar{a}\mathbf{1}_m} \cap \Omega$ with $\bar{x}_i \leq \tilde{x}_i$, $\forall i \in \{1, \dots, n\}$, and $(r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) < \delta(\epsilon)$, we have $\sum_{i=1}^n \|\Delta_F(r_{\bar{x}_i, \tilde{x}_i})\| < \epsilon$. Hence, F satisfies (ii) of Definition 12.58. Then, F is absolutely continuous at x_0 . This subcase is proved.

Case 2 β : $x_0 \in \Omega_2$. Then, $x_0 \in \Omega_1 \cap \Omega_2$. Fix any $i \in \{1, 2\}$. By $F|_{\Omega_i}$ being absolutely continuous at x_0 and (ii) of Definition 12.58, $\exists a_i \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_i(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_j)_{j=1}^n, (\tilde{x}_j)_{j=1}^n \subseteq \bar{r}_{x_0-a_i\mathbf{1}_m, x_0+a_i\mathbf{1}_m} \cap \Omega_i$ with $\bar{x}_j \leq \tilde{x}_j$, $\forall j \in \{1, \dots, n\}$, $(r_{\bar{x}_j, \tilde{x}_j})_{j=1}^n$ being pairwise disjoint, and $\sum_{j=1}^n \mu_{Bm}(r_{\bar{x}_j, \tilde{x}_j}) < \delta_i(\epsilon)$, we have $\sum_{j=1}^n \|\Delta_F(r_{\bar{x}_j, \tilde{x}_j})\| < \epsilon$.

Take $a := \min\{a_1, a_2\} \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta(\epsilon) := \min\{\delta_1(\epsilon/2), \delta_2(\epsilon/2)\} \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_j)_{j=1}^n, (\tilde{x}_j)_{j=1}^n \subseteq \bar{r}_{x_0-a\mathbf{1}_m, x_0+a\mathbf{1}_m} \cap \Omega$ with $\bar{x}_j \leq \tilde{x}_j$, $\forall j \in \{1, \dots, n\}$, $(r_{\bar{x}_j, \tilde{x}_j})_{j=1}^n$ being pairwise disjoint, and $\sum_{j=1}^n \mu_{Bm}(r_{\bar{x}_j, \tilde{x}_j}) < \delta(\epsilon)$. $\forall j \in \{1, \dots, n\}$, since $\Omega = \Omega_1 \cup \Omega_2$ and Ω_1 and Ω_2 are relatively closed rectangles in the rectangle Ω , then $r_{\bar{x}_j, \tilde{x}_j} = r_{\bar{x}_j, 1, \tilde{x}_j, 1} \cup r_{\bar{x}_j, 2, \tilde{x}_j, 2}$, where $\bar{x}_{j,k}, \tilde{x}_{j,k} \in \bar{r}_{x_0-a\mathbf{1}_m, x_0+a\mathbf{1}_m} \cap \Omega_k$, $\bar{x}_{j,k} \leq \tilde{x}_{j,k}$, $k = 1, 2$, and the two sets in the union are disjoint. To see the above, note that $r_{\bar{x}_j, \tilde{x}_j} \cap \Omega_k$ is a rectangle, $k = 1, 2$. If $\bar{r}_{\bar{x}_j, \tilde{x}_j} \subseteq \Omega_l$, for some $l \in \{1, 2\}$, then let $\bar{x}_{j,l} = \bar{x}_j$, $\tilde{x}_{j,l} = \tilde{x}_j$, $\bar{x}_{j,3-l} = x_0$ and $\tilde{x}_{j,3-l} = x_0$. The desired result holds. Consider the remaining pos-

sibility, we must have $\bar{x}_j \in \Omega_l \setminus \Omega_{3-l}$ and $\tilde{x}_j \in \Omega_{3-l} \setminus \Omega_l$, for some $l \in \{1, 2\}$. Without loss of generality, assume $l = 1$. Ω_1 and Ω_2 are m -dimensional rectangles and none contains the other, their union is a m -dimensional rectangle Ω , in which both Ω_1 and Ω_2 are relatively closed. The only possibility for this to happen is that $\pi_i(\Omega_1) = \pi_i(\Omega_2) = \pi_i(\Omega)$, $\forall i \in \bar{J} \setminus \{i_0\}$, for some $i_0 \in \bar{J}$, and $\pi_{i_0}(\Omega_1)$ and $\pi_{i_0}(\Omega_2)$ are overlapping relatively closed intervals in $\pi_{i_0}(\Omega)$. Let $\bar{x}_{j,1} = \bar{x}_j \in \Omega_1 \setminus \Omega_2$. Since $\bar{r}_{\bar{x}_j, \tilde{x}_j} \subseteq \Omega$, then $\bar{r}_{\bar{x}_j, \tilde{x}_j} \cap \Omega_1 \subseteq \Omega$ is a relatively closed rectangle with respect to $\bar{r}_{\bar{x}_j, \tilde{x}_j}$. Then, $\emptyset \neq \bar{r}_{\bar{x}_j, \tilde{x}_j} \cap \Omega_1 =: \bar{r}_{\bar{x}_{j,1}, \tilde{x}_{j,1}} \subseteq \bar{r}_{\bar{x}_j, \tilde{x}_j} \subseteq \Omega$. This implies that $\bar{x}_{j,1}, \tilde{x}_{j,1} \in \bar{r}_{x_0-a\mathbf{1}_m, x_0+a\mathbf{1}_m} \cap \Omega_1$ and $\bar{x}_{j,1} \leq \tilde{x}_{j,1}$. Since $\tilde{x}_j \in \Omega_2 \setminus \Omega_1$, then $\emptyset \neq \bar{r}_{\bar{x}_j, \tilde{x}_j} \cap \Omega_2 =: \bar{r}_{\hat{x}_j, \tilde{x}_j} \subseteq \bar{r}_{\bar{x}_j, \tilde{x}_j} \subseteq \Omega$, where $\hat{x}_j \in \bar{r}_{x_0-a\mathbf{1}_m, x_0+a\mathbf{1}_m} \cap \Omega_2$ and $\hat{x}_j \leq \tilde{x}_j$. Clearly, $\bar{x}_{j,1} \in \Omega_1 \cap \Omega_2$ since Ω_2 is relatively closed in Ω and $\hat{x}_j \in \Omega_1 \cap \Omega_2$ since Ω_1 is relatively closed in Ω . Then, $\hat{x}_j \in \bar{r}_{\bar{x}_j, \tilde{x}_j} \cap \Omega_1 = \bar{r}_{\bar{x}_{j,1}, \tilde{x}_{j,1}}$ and $\hat{x}_j \leq \tilde{x}_{j,1}$. This implies that $\bar{r}_{\bar{x}_{j,1}, \tilde{x}_{j,1}} \subseteq \Omega_1 \cap \Omega_2$. Let $\tilde{x}_{j,2} = \tilde{x}_j$ and $\bar{x}_{j,2} \in \bar{r}_{\hat{x}_j, \tilde{x}_{j,1}}$ be such that $\pi_l(\bar{x}_{j,2}) = \begin{cases} \pi_l(\hat{x}_j) & \text{if } \pi_l(\tilde{x}_{j,1}) = \pi_l(\tilde{x}_j) \\ \pi_l(\tilde{x}_{j,1}) & \text{if } \pi_l(\tilde{x}_{j,1}) < \pi_l(\tilde{x}_j) \end{cases}, \forall l \in \bar{J}$. Then, $\bar{x}_{j,2} \leq \tilde{x}_{j,2}$ and $\bar{x}_{j,2}, \tilde{x}_{j,2} \in \Omega_2 \cap \bar{r}_{x_0-a\mathbf{1}_m, x_0+a\mathbf{1}_m}$. Clearly, $r_{\bar{x}_{j,1}, \tilde{x}_{j,1}} \cap r_{\bar{x}_{j,2}, \tilde{x}_{j,2}} = \emptyset$ and $r_{\bar{x}_{j,1}, \tilde{x}_{j,1}} \cup r_{\bar{x}_{j,2}, \tilde{x}_{j,2}} = r_{\bar{x}_j, \tilde{x}_j}$ since $\bar{r}_{\bar{x}_{j,1}, \tilde{x}_{j,1}} \cup \bar{r}_{\hat{x}_j, \tilde{x}_j} = \bar{r}_{\bar{x}_j, \tilde{x}_j}$. These $\bar{x}_{j,1}, \bar{x}_{j,2}, \tilde{x}_{j,1}$, and $\tilde{x}_{j,2}$ satisfy the conditions specified above.

Then, $\sum_{j=1}^n \mu_{Bm}(r_{\bar{x}_j, \tilde{x}_j}) = \sum_{k=1}^2 \sum_{j=1}^m \mu_{Bm}(r_{\bar{x}_{j,k}, \tilde{x}_{j,k}}) < \delta(\epsilon)$ and $\sum_{j=1}^m \mu_{Bm}(r_{\bar{x}_{j,k}, \tilde{x}_{j,k}}) < \delta_k(\epsilon/2)$, $k = 1, 2$. This implies that $\sum_{j=1}^n \|\Delta_F(r_{\bar{x}_j, \tilde{x}_j})\| = \sum_{j=1}^n \|\Delta_F(r_{\bar{x}_{j,1}, \tilde{x}_{j,1}}) + \Delta_F(r_{\bar{x}_{j,2}, \tilde{x}_{j,2}})\| \leq \sum_{k=1}^2 \sum_{j=1}^n \|\Delta_F(r_{\bar{x}_{j,k}, \tilde{x}_{j,k}})\| < \epsilon$, where the first equality follows from Proposition 12.46. Hence, F satisfies (ii) of Definition 12.58. Therefore, F is absolutely continuous at x_0 . This subcase is proved.

In both subcases, F is absolutely continuous at x_0 . This completes the induction process.

By the arbitrariness of x_0 , F is absolutely continuous. This completes the proof for Case 2.

In both cases, F is absolutely continuous. This completes the proof of the proposition. $\square$

Definition 12.63 Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}_1$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}_2$, $U \subseteq \mathcal{X}$, $f : U \rightarrow \mathcal{Y}$, and $x_0 \in U$. f is said to be locally Lipschitz at x_0 (with Lipschitz constant $L_0 \in [0, \infty) \subset \mathbb{R}$) if $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x_1, x_2 \in U \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, we have $\|f(x_1) - f(x_2)\| \leq L_0 \|x_1 - x_2\|$. f is said to be locally Lipschitz on U if it is locally Lipschitz at x , $\forall x \in U$. f is said to be Lipschitz on U (with Lipschitz constant $L \in [0, \infty) \subset \mathbb{R}$) if $\forall x_1, x_2 \in U$, we have $\|f(x_1) - f(x_2)\| \leq L \|x_1 - x_2\|$.

Clearly, if f is locally Lipschitz at x_0 , then f is continuous at x_0 ; if f is Lipschitz on U , then it is uniformly continuous; if f is Lipschitz on U , then it is locally Lipschitz on U .

Definition 12.64 Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}_1$, Y be a set, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}_2$, $U \subseteq \mathcal{X}$, $f : U \times Y \rightarrow \mathcal{Z}$, and $x_0 \in U$. f is said to be locally Lipschitz at x_0 (with Lipschitz constant $L_0 \in [0, \infty) \subset \mathbb{R}$) uniformly over Y if $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x_1, x_2 \in U \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta)$, $\forall y \in Y$, we have $\|f(x_1, y) - f(x_2, y)\| \leq L_0 \|x_1 - x_2\|$. f is said to be locally Lipschitz on U uniformly over Y if it is locally Lipschitz at x uniformly over Y , $\forall x \in U$. f is said to be Lipschitz on U (with Lipschitz constant $L \in [0, \infty) \subset \mathbb{R}$) uniformly over Y if $\forall x_1, x_2 \in U$, $\forall y \in Y$, we have $\|f(x_1, y) - f(x_2, y)\| \leq L \|x_1 - x_2\|$.

Clearly, if f is Lipschitz on U uniformly over Y , then it is locally Lipschitz on U uniformly over Y .

Proposition 12.65 Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}_1$, Y be a set, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}_2$, $U \subseteq \mathcal{X}$ be compact, $f : U \times Y \rightarrow \mathcal{Z}$ be locally Lipschitz on U uniformly over Y . Assume that $\forall x \in U$, $\sup_{y \in Y} \|f(x, y)\| < +\infty$. Then, f is Lipschitz on U uniformly over Y .

Proof $\forall x \in U$, by the local Lipschitz continuity of f , $\exists L_x \in [0, \infty) \subset \mathbb{R}$, $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$, $\forall x_1, x_2 \in O_x := U \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x)$, $\forall y \in Y$, we have $\|f(x_1, y) - f(x_2, y)\| \leq L_x \|x_1 - x_2\|$. Let $\hat{O}_x := \mathcal{B}_{\mathcal{X}}(x, \delta_x/2)$. Then, $U \subseteq \bigcup_{x \in U} \hat{O}_x$, which is an open covering of U . By the compactness of U , there exists a finite set $X_N \subseteq U$ such that $U \subseteq \bigcup_{x \in X_N} \hat{O}_x$. Let $L_M := \max\{\max_{x \in X_N} L_x, \frac{2}{\min_{x \in X_N} \delta_x} \max_{\hat{x}_1, \hat{x}_2 \in X_N} (L_{\hat{x}_1} \delta_{\hat{x}_1}/2 + L_{\hat{x}_2} \delta_{\hat{x}_2}/2 + \sup_{y \in Y} \|f(\hat{x}_1, y)\| + \sup_{y \in Y} \|f(\hat{x}_2, y)\|), 0\} \in [0, \infty) \subset \mathbb{R}$. $\forall x_1, x_2 \in U$, $\forall y \in Y$, $\exists \hat{x}_1 \in X_N$ such that $x_1 \in \hat{O}_{\hat{x}_1}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x_2 \in \hat{O}_{\hat{x}_1}$; Case 2: $x_2 \notin \hat{O}_{\hat{x}_1}$. Case 1: $x_2 \in \hat{O}_{\hat{x}_1}$. Then, $x_1, x_2 \in \hat{O}_{\hat{x}_1}$ and $\|f(x_1, y) - f(x_2, y)\| \leq L_{\hat{x}_1} \|x_1 - x_2\| \leq L_M \|x_1 - x_2\|$. Case 2: $x_2 \notin \hat{O}_{\hat{x}_1}$. Then, $\|x_2 - \hat{x}_1\| \geq \delta_{\hat{x}_1}$ and $\|x_2 - x_1\| \geq \|x_2 - \hat{x}_1\| - \|x_1 - \hat{x}_1\| > \delta_{\hat{x}_1}/2 \geq \frac{\min_{x \in X_N} \delta_x}{2} > 0$. $\exists \hat{x}_2 \in X_N$ such that $x_2 \in \hat{O}_{\hat{x}_2}$. Then, $\|f(x_1, y) - f(x_2, y)\| \leq \|f(x_1, y) - f(\hat{x}_1, y)\| + \|f(\hat{x}_1, y) - f(\hat{x}_2, y)\| + \|f(\hat{x}_2, y) - f(x_2, y)\| \leq L_{\hat{x}_1} \|x_1 - \hat{x}_1\| + \|f(\hat{x}_1, y)\| + \|f(\hat{x}_2, y)\| + L_{\hat{x}_2} \|x_2 - \hat{x}_2\| \leq \sup_{y \in Y} \|f(\hat{x}_1, y)\| + \sup_{y \in Y} \|f(\hat{x}_2, y)\| + L_{\hat{x}_1} \delta_{\hat{x}_1}/2 + L_{\hat{x}_2} \delta_{\hat{x}_2}/2 \leq L_M \frac{\min_{x \in X_N} \delta_x}{2} \leq L_M \|x_1 - x_2\|$. Hence, in both cases, we have $\|f(x_1, y) - f(x_2, y)\| \leq L_M \|x_1 - x_2\|$. Then, f is Lipschitz on U uniformly over Y . This completes the proof of the proposition. $\square$

Proposition 12.66 Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, Y be a set, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $U \subseteq \mathcal{X}$ be convex, $f : U \times Y \rightarrow \mathcal{Z}$ be partial differentiable with respect to x . Assume that $\exists L \in [0, \infty) \subset \mathbb{R}$ such that $\left\| \frac{\partial f}{\partial x}(x, y) \right\| \leq L$, $\forall (x, y) \in U \times Y$. Then, f is Lipschitz on U with Lipschitz constant L uniformly over Y .

Proof $\forall x_1, x_2 \in U, \forall y \in Y$, by Mean Value Theorem 9.23, $\exists t \in (0, 1) \subset \mathbb{R}$ such that $\|f(x_1, y) - f(x_2, y)\| \leq \left\| \frac{\partial f}{\partial x}(tx_1 + (1-t)x_2, y) \right\| \|x_1 - x_2\| \leq L \|x_1 - x_2\|$. Hence, f is Lipschitz on U with Lipschitz constant L uniformly over Y . This completes the proof of the proposition. $\square$

Proposition 12.67 *Let $I \subseteq \mathbb{R}$ be an interval, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}_1$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}_2$, $U \subseteq \mathcal{Y}$, $f : I \rightarrow U$ be absolutely continuous at $x_0 \in I$, $g : U \rightarrow \mathcal{Z}$ be locally Lipschitz at $y_0 := f(x_0) \in U$. Then, $g \circ f$ is absolutely continuous at x_0 .*

Proof By g being locally Lipschitz at y_0 , then $\exists L_0 \in [0, \infty) \subset \mathbb{R}$, $\exists \delta_g \in (0, \infty) \subset \mathbb{R}$, $\forall y_1, y_2 \in U \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta_g)$, we have $\|g(y_1) - g(y_2)\| \leq L_0 \|y_1 - y_2\|$. Since f is absolutely continuous at x_0 , then it is continuous at x_0 by Proposition 12.59. Then, $\exists a_0 \in (0, \infty) \subset \mathbb{R}$ such that $\forall x \in \mathcal{B}_{\mathbb{R}}(x_0, a_0) \cap I$, we have $f(x) \in U \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta_g)$. By the absolute continuity of f at x_0 , $\exists a_f \in (0, a_0) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_f(\epsilon) \in (0, a_f] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall x_0 - a_f \leq x_1 \leq \bar{x}_1 \leq x_2 \leq \bar{x}_2 \leq \dots \leq x_n \leq \bar{x}_n \leq x_0 + a_f$ with $x_i, \bar{x}_i \in I$, $i = 1, \dots, n$, and $\sum_{i=1}^n (\bar{x}_i - x_i) < \delta_f(\epsilon)$, we have $\sum_{i=1}^n \|f(\bar{x}_i) - f(x_i)\| < \epsilon$. For the function $g \circ f$, take $a = a_f \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta(\epsilon) = \delta_f(\epsilon/(1 + L_0)) \in (0, a] \subset \mathbb{R}$. $\forall n \in \mathbb{N}$, $\forall x_0 - a \leq x_1 \leq \bar{x}_1 \leq x_2 \leq \bar{x}_2 \leq \dots \leq x_n \leq \bar{x}_n \leq x_0 + a$ with $x_i, \bar{x}_i \in I$, $i = 1, \dots, n$, and $\sum_{i=1}^n (\bar{x}_i - x_i) < \delta(\epsilon)$, we have $f(\bar{x}_i), f(x_i) \in U \cap \mathcal{B}_{\mathcal{Y}}(y_0, \delta_g)$ and $\sum_{i=1}^n \|g \circ f(\bar{x}_i) - g \circ f(x_i)\| \leq \sum_{i=1}^n L_0 \|f(\bar{x}_i) - f(x_i)\| \leq L_0 \frac{\epsilon}{1 + L_0} < \epsilon$. Hence, $g \circ f$ is absolutely continuous at x_0 . This completes the proof of the proposition. $\square$

Proposition 12.68 *Let $I \subseteq \mathbb{R}$ be an interval, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $f : I \rightarrow \mathcal{Y}$ and $g : I \rightarrow \mathcal{Y}$ be absolutely continuous at $x_0 \in I$, $A : I \rightarrow \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be absolutely continuous at x_0 , and $h : I \rightarrow \mathcal{Z}$ be defined by $h(x) = A(x)f(x)$, $\forall x \in I$. Then, $f + g$ and h are absolutely continuous at x_0 .*

Proof By Propositions 12.61, 9.14, 12.67, and 12.66, we have that $f + g$ is absolutely continuous at x_0 .

By Propositions 12.61, 9.17, 12.67, and 12.66, we have that h is absolutely continuous at x_0 .

This completes the proof of the proposition. $\square$

Proposition 12.69 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_{\mathbb{B}}(\mathbb{R}^m)$ be a rectangle, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $F : \Omega \rightarrow \mathcal{Y}$ be absolutely continuous at $x_0 \in \Omega$, and $A \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$. Then, $h : \Omega \rightarrow \mathcal{Z}$, defined by $h(x) = Af(x)$, $\forall x \in \Omega$, is absolutely continuous at x_0 .*

Proof We will prove this result using mathematical induction on m .
1° $m = 1$. Clearly, the result follows from Proposition 12.68.

2° Assume that the result holds for $m \leq k \in \mathbb{N}$.

3° Consider the case when $m = k + 1$. We will prove that h is absolutely continuous at x_0 by Definition 12.58. Fix any $i_0 \in \bar{J} := \{1, \dots, m\}$, let D_{i_0} , $\pi_{\hat{i}_0}$, and F_{i_0} be defined as in (i) of Definition 12.58. Let $h_{i_0} : D_{i_0} \rightarrow \mathbb{Z}$ be defined by $h_{i_0}(x) = h(\bar{x}) = AF(\bar{x}) = AF_{i_0}(x)$, $\forall x \in D_{i_0} \subseteq \mathbb{R}^k$, $\forall \bar{x} \in \Omega$ with $\pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0)$ and $\pi_{\hat{i}_0}(\bar{x}) = x$. By the assumption, F is absolutely continuous at x_0 , then F_{i_0} is absolutely continuous at $\pi_{\hat{i}_0}(x_0)$. By inductive assumption, h_{i_0} is absolutely continuous at $\pi_{\hat{i}_0}(x_0)$. Thus, (i) of Definition 12.58 holds.

Since F is absolutely continuous at x_0 , by (ii) of Definition 12.58, $\exists a \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{\mathbb{F}}_{x_0 - a\mathbf{1}_m, x_0 + a\mathbf{1}_m} \cap \Omega$ with $\bar{x}_i \leq \tilde{x}_i$, $\forall i \in \{1, \dots, n\}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) < \delta(\epsilon)$, we have $\sum_{i=1}^n \|\Delta_F(r_{\bar{x}_i, \tilde{x}_i})\| < \frac{\epsilon}{1 + \|A\|}$. It is easy to see that $\Delta_h(r_{\bar{x}_i, \tilde{x}_i}) = A\Delta_F(r_{\bar{x}_i, \tilde{x}_i})$, $\forall i \in \{1, \dots, n\}$. Then, $\sum_{i=1}^n \|\Delta_h(r_{\bar{x}_i, \tilde{x}_i})\| \leq \|A\| \sum_{i=1}^n \|\Delta_F(r_{\bar{x}_i, \tilde{x}_i})\| < \epsilon$. Hence, (ii) of Definition 12.58 holds.

Thus, h is absolutely continuous at x_0 .

This completes the induction process and therefore the proof of the proposition. $\square$

Proposition 12.70 *Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}_1$, Y be a set, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}_2$, $U \subseteq \mathcal{X}$ be endowed with subset topology $\mathcal{O}$, $f : U \times Y \rightarrow \mathcal{Z}$, $g : U \rightarrow \mathbb{R}$ be continuous, $(c_i)_{i \in \mathbb{Z}} \subset \mathbb{R}$ be a sequence of strictly increasing real numbers with $\lim_{i \rightarrow -\infty} c_i = -\infty$ and $\lim_{i \rightarrow \infty} c_i = \infty$, $U_i := \{x \in U \mid c_i \leq g(x) \leq c_{i+1}\}$, $\forall i \in \mathbb{Z}$. Assume that*

(i) U is locally convex according to Definition 7.124;

(ii) $f|_{U_i \times Y}$ is locally Lipschitz on U_i uniformly over Y , $\forall i \in \mathbb{Z}$.

Then, f is locally Lipschitz on U uniformly over Y .

Proof Fix any $x_0 \in U = \bigcup_{i \in \mathbb{Z}} U_i$. Then, $g(x_0) \in \mathbb{R}$. We will distinguish two exhaustive and mutually exclusives cases: Case 1: $\exists i \in \mathbb{Z}$ such that $c_i < g(x_0) < c_{i+1}$; Case 2: $\exists i \in \mathbb{Z}$ such that $g(x_0) = c_i$. Case 1: $\exists i \in \mathbb{Z}$ such that $c_i < g(x_0) < c_{i+1}$. Then, $x_0 \in U_i$. Since g is continuous, then $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathcal{X}}(x_0, \delta_1) \cap U \subseteq U_i$. By $f|_{U_i \times Y}$ being locally Lipschitz on U_i uniformly over Y , we have $f|_{U_i \times Y}$ is locally Lipschitz at x_0 uniformly over Y . By Definition 12.64, $\exists L_1 \in [0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \delta_1) \subset \mathbb{R}$, $\forall x_1, x_2 \in \mathcal{B}_{\mathcal{X}}(x_0, \delta) \cap U_i = \mathcal{B}_{\mathcal{X}}(x_0, \delta) \cap U$, $\forall y \in Y$, we have $\|f(x_1, y) - f(x_2, y)\| \leq L_1 \|x_1 - x_2\|$. Hence, f is locally Lipschitz at x_0 uniformly over Y .

Case 2: $\exists i \in \mathbb{Z}$ such that $g(x_0) = c_i$. Then, $x_0 \in U_{i-1} \cap U_i$. By (i) and the continuity of g , $\exists \delta_{x_0} \in (0, \infty) \subset \mathbb{R}$ such that $U \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta_{x_0}) \subseteq U_{i-1} \cup U_i$ is convex. By $f|_{U_j \times Y}$ being locally Lipschitz on U_j uniformly over Y , we have $f|_{U_j \times Y}$ is locally Lipschitz at x_0 uniformly over Y , $j = i - 1, i$.

By Definition 12.64, $\exists L_j \in [0, \infty) \subset \mathbb{R}$, $\exists \delta_j \in (0, \delta_{x_0}] \subset \mathbb{R}$, $\forall x_1, x_2 \in \mathcal{B}_X(x_0, \delta_j) \cap U_j$, $\forall y \in Y$, we have $\|f(x_1, y) - f(x_2, y)\| \leq L_j \|x_1 - x_2\|$, $j = i-1, i$. Take $L := L_{i-1} \vee L_i \in [0, \infty) \subset \mathbb{R}$ and $\delta := \delta_1 \wedge \delta_2 \in (0, \delta_{x_0}] \subset \mathbb{R}$. $\forall x_1, x_2 \in \mathcal{B}_X(x_0, \delta) \cap U =: D$, $\forall y \in Y$, by Proposition 6.40, D is convex. If $\exists j \in \{i-1, i\}$ such that $x_1, x_2 \in U_j$, then $\|f(x_1, y) - f(x_2, y)\| \leq L_j \|x_1 - x_2\| \leq L \|x_1 - x_2\|$. On the other hand, without loss of generality, assume that $x_1 \in U_{i-1}$ and $x_2 \in U_i$. Consider the line segment connecting x_1 and x_2 .

Claim 12.70.1 $\exists t_0 \in [0, 1] \subset \mathbb{R}$ such that $x_{t_0} := t_0 x_1 + (1 - t_0)x_2 \in U_{i-1} \cap U_i \cap D$.

Proof of claim: Since D is convex, then $x_t := tx_1 + (1 - t)x_2 \in D \subseteq U_{i-1} \cup U_i$, $\forall t \in [0, 1] \subset \mathbb{R}$. Define

$$t_0 := \sup \{t \in [0, 1] \subset \mathbb{R} \mid x_t \in U_i\} =: \sup A$$

Clearly, $0 \in A$ and $0 \leq t_0 \leq 1$. Suppose $t_0 \notin A$. This leads to $t_0 \in (0, 1] \subset \mathbb{R}$. Then, $\exists (t_n)_{n=1}^\infty \subseteq [0, t_0) \cap A$ such that $\lim_{n \in \mathbb{N}} t_n = t_0$. $\forall n \in \mathbb{N}$, by $t_n \in A$, we have $g(x_{t_n}) \geq c_i$. Then, $g(x_{t_0}) = \lim_{n \in \mathbb{N}} g(x_{t_n}) \geq c_i$, $x_{t_0} \in U_i$, and $t_0 \in A$. This is a contradiction. Hence, $t_0 \in A$ and $x_{t_0} \in U_2$. If $t_0 = 1$, then $x_{t_0} = x_1 \in U_{i-1}$ and $x_{t_0} \in U_{i-1} \cap U_i$. Hence, the claim holds. On the other hand, if $t_0 < 1$, then $\forall t \in (t_0, 1] \subset \mathbb{R}$, $t \notin A$ and $x_t \in U_{i-1}$. This implies that $g(x_t) \leq c_i$, $\forall t \in (t_0, 1] \subset \mathbb{R}$. Then, $g(x_{t_0}) = \lim_{t \rightarrow t_0^+} g(x_t) \leq c_i$ and $x_{t_0} \in U_{i-1}$. Hence, $x_{t_0} \in U_{i-1} \cap U_i$. The claim holds as well. This completes the proof of the claim. $\square$

By Claim 12.70.1, $\|f(x_1, y) - f(x_2, y)\| \leq \|f(x_1, y) - f(x_{t_0}, y)\| + \|f(x_{t_0}, y) - f(x_2, y)\| \leq L_{i-1} \|x_1 - x_{t_0}\| + L_i \|x_2 - x_{t_0}\| = L_{i-1} (1 - t_0) \|x_1 - x_2\| + L_i t_0 \|x_1 - x_2\| \leq L \|x_1 - x_2\|$.

In both cases, we have that $\|f(x_1, y) - f(x_2, y)\| \leq L \|x_1 - x_2\|$. Then, f is locally Lipschitz at x_0 uniformly over Y .

In both cases, we have that f is locally Lipschitz at x_0 uniformly over Y . By the arbitrariness of x_0 , f is locally Lipschitz on U uniformly over Y . This completes the proof of the proposition. $\square$

Definition 12.71 Let $a, b \in \mathbb{R}$, $I := r_{a \wedge b, a \vee b} \subset \mathbb{R}$ be the semi-open interval with a and b as end points, $\mathbb{I} := ((I, |\cdot|), \mathcal{B}, \mu)$ be the finite metric measure subspace of $\mathbb{R}$, $\mathcal{Y}$ be a normed linear space, and $f : I \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. We will write

$$\int_a^b f(x) dx := \int_a^b f d\mu_B := \begin{cases} \int_I f d\mu_B & \text{if } b \geq a \\ -\int_I f d\mu_B & \text{if } b < a \end{cases}$$

whenever the right-hand-side makes sense.

By Lemma 11.73, the interval I may be $\bar{r}_{a \wedge b, a \vee b}$, or $r_{a \wedge b, a \vee b}$, or $r_{a \wedge b, a \vee b}^\circ$, which will not change the value of the integral.

Fact 12.72 Let $a, b, c \in \mathbb{R}$, $I := r_{a \wedge b \wedge c, a \vee b \vee c} \subset \mathbb{R}$, $\mathbb{I} := ((I, |\cdot|), \mathcal{B}, \mu)$ be the finite metric measure subspace of $\mathbb{R}$, $\mathcal{Y}$ be a separable Banach space, and $f : I \rightarrow \mathcal{Y}$ be absolutely integrable over $\mathbb{I}$. Then, $\int_a^c f(x) dx + \int_c^b f(x) dx = \int_a^b f(x) dx \in \mathcal{Y}$.

Proof We will distinguish six exhaustive and mutually exclusives cases: Case 1: $a \leq c \leq b$; Case 2: $a \leq b \leq c$; Case 3: $b \leq a \leq c$; Case 4: $b \leq c \leq a$; Case 5: $c \leq a \leq b$; Case 6: $c \leq b \leq a$. Case 1: $a \leq c \leq b$. Let $J_1 := r_{a,c} \subset \mathbb{R}$, $J_2 := r_{c,b} \subset \mathbb{R}$. By Proposition 11.92, $\int_a^b f(x) dx = \int_I f d\mu_B = \int_{J_1} f d\mu_B + \int_{J_2} f d\mu_B = \int_a^c f(x) dx + \int_c^b f(x) dx \in \mathcal{Y}$. Then, $\int_a^c f(x) dx + \int_c^b f(x) dx = \int_a^b f(x) dx \in \mathcal{Y}$.

Case 2: $a \leq b \leq c$. By Case 1, $\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx$. Then, $\int_a^b f(x) dx = \int_a^c f(x) dx - \int_b^c f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$.

Case 3: $b \leq a \leq c$. By Case 1, $\int_b^c f(x) dx = \int_b^a f(x) dx + \int_a^c f(x) dx$. Then, $\int_a^b f(x) dx = -\int_b^a f(x) dx = \int_a^c f(x) dx - \int_b^c f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$.

Case 4: $b \leq c \leq a$. By Case 1, $\int_b^a f(x) dx = \int_b^c f(x) dx + \int_c^a f(x) dx$. Then, $\int_a^b f(x) dx = -\int_b^a f(x) dx = -\int_b^c f(x) dx - \int_c^a f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$.

Case 5: $c \leq a \leq b$. By Case 1, $\int_c^b f(x) dx = \int_c^a f(x) dx + \int_a^b f(x) dx$. Then, $\int_a^b f(x) dx = \int_c^b f(x) dx - \int_c^a f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$.

Case 6: $c \leq b \leq a$. By Case 1, $\int_c^a f(x) dx = \int_c^b f(x) dx + \int_b^a f(x) dx$. Then, $\int_a^b f(x) dx = -\int_b^a f(x) dx = \int_c^b f(x) dx - \int_c^a f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$.

This completes the proof of the fact. $\square$

Proposition 12.73 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle with the subset topology $\mathcal{O}$, $((P(\Omega), |\cdot|), \mathcal{B}, \mu)$ be the σ -finite metric measure subspace of $\mathbb{R}^m$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $F : \Omega \rightarrow \mathcal{Y}$ be absolutely continuous. Then, the following statements hold.

(i) F is of locally bounded variation.

(ii) If, in addition, $\mathcal{Y}$ is a Banach space over $\mathbb{K}$, then there exists a unique σ -finite $\mathcal{Y}$ -valued measure space $(P(\Omega), \mathcal{B}, \nu)$ such that ν admits F as a cumulative distribution function. Furthermore, $\mathcal{P} \circ \nu(r_{x_1, x_2}) = T_F(r_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$, and $\mathcal{P} \circ \nu(P(\Omega)) = T_F$ and $\mathcal{P} \circ \nu \ll \mu$.

Proof (i) By Proposition 12.52, Ω is a region and $P(\Omega) \in \mathcal{B}_B(\mathbb{R}^m)$ is a rectangle. By Proposition 12.59, F is continuous. Then, (i) of Definition 12.41 holds.

$\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$, consider the closed rectangle $\bar{r}_{x_1, x_2} \subseteq \Omega$. $\forall x_0 \in \bar{r}_{x_1, x_2}$, F is absolutely continuous at x_0 . By (ii) of Definition 12.58, $\exists a_{x_0} \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_{x_0}(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{r}_{x_0 - a_{x_0} \mathbf{1}_m, x_0 + a_{x_0} \mathbf{1}_m} \cap \Omega$ with $\bar{x}_i \leq \tilde{x}_i$, $\forall i \in \{1, \dots, n\}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) < \delta_{x_0}(\epsilon)$, we have $\sum_{i=1}^n \|\Delta_F(r_{\bar{x}_i, \tilde{x}_i})\| < \epsilon$. Then, $\bar{r}_{x_1, x_2} \subseteq \bigcup_{x \in \bar{r}_{x_1, x_2}} r_{x - a_x \mathbf{1}_m, x + a_x \mathbf{1}_m}^\circ$. By the compactness of $\bar{r}_{x_1, x_2}$, $\exists$ a finite set $N \subseteq \bar{r}_{x_1, x_2}$ such that $\bar{r}_{x_1, x_2} \subseteq \bigcup_{x \in N} r_{x - a_x \mathbf{1}_m, x + a_x \mathbf{1}_m}^\circ$. Let $\epsilon_0 := 1 \in \mathbb{R}$, $a := \max_{x \in N} a_x \in (0, \infty) \subset \mathbb{R}$, and $\delta := \min_{x \in N} \delta_x(\epsilon_0) \in (0, \infty) \subset \mathbb{R}$. $\forall n \in \mathbb{Z}_+$, $\forall (\check{x}_i)_{i=1}^n, (\hat{x}_i)_{i=1}^n \subseteq \bar{r}_{x_1, x_2}$ with $\check{x}_i \leq \hat{x}_i$, $\forall i \in \{1, \dots, n\}$, $\bigcup_{i=1}^n r_{\check{x}_i, \hat{x}_i} = r_{x_1, x_2}$, and sets in the union being pairwise disjoint, we have, by Proposition 12.46,

$$\begin{aligned} \sum_{i=1}^n \|\Delta_F(r_{\check{x}_i, \hat{x}_i})\| &= \sum_{i=1}^n \left\| \sum_{x \in N} \sum_{j=1}^{n_x} \Delta_F(r_{\check{x}_i, x, j, \hat{x}_i, x, j}) \right\| \\ &\leq \sum_{x \in N} \sum_{i=1}^n \sum_{j=1}^{n_x} \|\Delta_F(r_{\check{x}_i, x, j, \hat{x}_i, x, j})\| \end{aligned}$$

where $n_x \in \mathbb{Z}_+$, $\forall x \in N$, $(\check{x}_{i, x, j})_{x \in N, j \in \{1, \dots, n_x\}}, (\hat{x}_{i, x, j})_{x \in N, j \in \{1, \dots, n_x\}} \subseteq \bar{r}_{\check{x}_i, \hat{x}_i}$, $\check{x}_{i, x, j} \leq \hat{x}_{i, x, j}$, $\forall x \in N$, $\forall j \in \{1, \dots, n_x\}$, $\bigcup_{x \in N} \bigcup_{j=1}^{n_x} r_{\check{x}_{i, x, j}, \hat{x}_{i, x, j}} = r_{\check{x}_i, \hat{x}_i}$, the sets in the union are pairwise disjoint, and $\bigcup_{j=1}^{n_x} r_{\check{x}_{i, x, j}, \hat{x}_{i, x, j}} \subseteq r_{x - a_x \mathbf{1}_m, x + a_x \mathbf{1}_m}$, $\forall x \in N$, $i = 1, \dots, n$. $\forall x \in N$, by further increasing n_x such that $\sum_{i=1}^n \mu_{Bm}(r_{\check{x}_{i, x, j}, \hat{x}_{i, x, j}}) < \delta_x(\epsilon_0)$, $j = 1, \dots, n_x$. It is easy to see that $n_x \leq 1 + \lfloor 2^m a_x^m / \delta_x(\epsilon_0) \rfloor$. Then, $\sum_{i=1}^n \|\Delta_F(r_{\check{x}_i, \hat{x}_i})\| \leq \sum_{x \in N} \sum_{j=1}^{n_x} \sum_{i=1}^n \|\Delta_F(r_{\check{x}_{i, x, j}, \hat{x}_{i, x, j}})\| < \sum_{x \in N} \sum_{j=1}^{n_x} 1 = \text{card}(N)(1 + \lfloor 2^m a^m / \delta \rfloor)$. Hence, $T_F(r_{x_1, x_2}) \leq \text{card}(N)(1 + \lfloor 2^m a^m / \delta \rfloor) < \infty$. Hence, F satisfies (ii) of Definition 12.41.

$\forall x_1, x_2 \in \Omega$ with $x_1 < x_2$, define $\bar{G} : \bar{r}_{x_1, x_2} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\bar{G}(x) = T_F(r_{x_1, x})$, $\forall x \in \bar{r}_{x_1, x_2}$. We will show that $\bar{G}$ is continuous on the right. Fix any $x_0 \in \bar{r}_{x_1, x_2}$. Consider the sets $S_J := \{x \in \bar{r}_{x_1, x_2} \mid \pi_i(x) = \pi_i(x_0), \forall i \in J\}$, $\forall J \subseteq \bar{J} := \{1, \dots, m\}$ with $J \neq \emptyset$. Fix any $J \subseteq \bar{J}$ with $J \neq \emptyset$. Clearly, S_J is compact. $\forall x \in S_J$, by (ii) of Definition 12.58, $\exists a_x \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_x(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_i)_{i=1}^n, (\tilde{x}_i)_{i=1}^n \subseteq \bar{r}_{x - a_x \mathbf{1}_m, x + a_x \mathbf{1}_m} \cap \bar{r}_{x_1, x_2}$ with $\bar{x}_i \leq \tilde{x}_i$, $\forall i \in \{1, \dots, n\}$, $(r_{\bar{x}_i, \tilde{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu_{Bm}(r_{\bar{x}_i, \tilde{x}_i}) < \delta_x(\epsilon)$, we have $\sum_{i=1}^n \|\Delta_F(r_{\bar{x}_i, \tilde{x}_i})\| < \epsilon$. Then, $S_J \subseteq \bigcup_{x \in S_J} r_{x - a_x \mathbf{1}_m, x + a_x \mathbf{1}_m}^\circ$. By the compactness of S_J , $\exists$ a finite set $N_J \subseteq S_J$ such that $S_J \subseteq \bigcup_{x \in N_J} r_{x - a_x \mathbf{1}_m, x + a_x \mathbf{1}_m}^\circ$. Let $N := \bigcup_{J \subseteq \bar{J}, J \neq \emptyset} N_J$. Let $\bar{a} := \max_{x \in N} a_x \in (0, \infty) \subset \mathbb{R}$ and $\tilde{a} := \min_{x \in N} a_x \in (0, \infty) \subset \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta := \min_{x \in N} \delta_x(\epsilon / (1 + \text{card}(N))) \in (0, \infty) \subset \mathbb{R}$. Let $\tilde{\delta}$ be the minimum positive real root of $\prod_{i=1}^m (\pi_i(x_0) - \pi_i(x_1) + \delta) - \prod_{i=1}^m (\pi_i(x_0) - \pi_i(x_1)) = \delta/2$ and $\bar{\delta} := \tilde{\delta} \wedge \tilde{a} \in (0, \infty) \subset \mathbb{R}$. $\forall \bar{x} \in \mathcal{B}_{\mathbb{R}^m}(x_0, \bar{\delta}) \cap \bar{r}_{x_1, x_2}$ with $x_0 \leq \bar{x}$, we have $|\bar{G}(\bar{x}) - \bar{G}(x_0)| = |T_F(r_{x_1, \bar{x}}) - T_F(r_{x_1, x_0})| = \sum_{J \subseteq \bar{J}, J \neq \emptyset} T_F(r_{\check{x}_J, \hat{x}_J})$, where $\check{x}_J \in \bar{r}_{x_1, x_0}$ is such that $\pi_i(\check{x}_J) = \pi_i(x_0)$, $\forall i \in \bar{J}$, $\pi_i(\hat{x}_J) = \pi_i(x_1)$,

$\forall i \in \bar{J} \setminus J$; and $\hat{x}_J \in \bar{r}_{x_0, \bar{x}}$ is such that $\pi_i(\hat{x}_J) = \pi_i(\bar{x})$, $\forall i \in J$, $\pi_i(\hat{x}_J) = \pi_i(x_0)$, $\forall i \in \bar{J} \setminus J$; and the second equality follows from Proposition 12.47. $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$, $r_{\hat{x}_J, \hat{x}_J} \subseteq \bigcup_{x \in N_J} r_{x-a_x \mathbf{1}_m, x+a_x \mathbf{1}_m}^\circ$ implies that $r_{\hat{x}_J, \hat{x}_J} = \bigcup_{x \in N_J} (r_{x-a_x \mathbf{1}_m, x+a_x \mathbf{1}_m}^\circ \cap r_{\hat{x}_J, \hat{x}_J})$. Then, we have $|\bar{G}(\bar{x}) - \bar{G}(x_0)| \leq \sum_{x \in N} \sum_{J \subseteq \bar{J}, J \neq \emptyset} T_F(r_{\hat{x}_J, \hat{x}_J} \cap r_{x-a_x \mathbf{1}_m, x+a_x \mathbf{1}_m})$. Note that, $\forall x \in N$, $\sum_{J \subseteq \bar{J}, J \neq \emptyset} \mu_{Bm}(r_{\hat{x}_J, \hat{x}_J} \cap r_{x-a_x \mathbf{1}_m, x+a_x \mathbf{1}_m}) \leq \sum_{J \subseteq \bar{J}, J \neq \emptyset} \mu_{Bm}(r_{\hat{x}_J, \hat{x}_J}) \leq \prod_{i=1}^m (\pi_i(x_0) - \pi_i(x_1) + \delta) - \prod_{i=1}^m (\pi_i(x_0) - \pi_i(x_1)) \leq \delta/2 < \delta$. Then, $\sum_{J \subseteq \bar{J}, J \neq \emptyset} T_F(r_{\hat{x}_J, \hat{x}_J} \cap r_{x-a_x \mathbf{1}_m, x+a_x \mathbf{1}_m}) < \epsilon/(1 + \text{card}(N))$. Hence, $|\bar{G}(\bar{x}) - \bar{G}(x_0)| < \sum_{x \in N} \epsilon/(1 + \text{card}(N)) < \epsilon$. This shows that $\bar{G}$ is continuous on the right at x_0 . By the arbitrariness of x_0 , we have $\bar{G}$ is continuous on the right. Hence, (iii) of Definition 12.41 holds. Therefore, F is of locally bounded variation. This completes the proof of (i).

(ii) By Theorem 12.50, there exists a unique σ -finite $\mathcal{Y}$ -valued measure space $(P(\Omega), \mathcal{B}, \nu)$ such that F is a cumulative distribution function of ν , $\mathcal{P} \circ \nu(r_{x_1, x_2}) = T_F(r_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$, and $\mathcal{P} \circ \nu(P(\Omega)) = T_F$.

All we need to show is $\mathcal{P} \circ \nu \ll \mu$. First, we consider the special case $\Omega = \bar{r}_{x_1, x_2} \subseteq \mathbb{R}^m$, where $x_1, x_2 \in \mathbb{R}^m$ with $x_1 \leq x_2$. $\forall x \in \bar{r}_{x_1, x_2}$, F being absolutely continuous at x implies that $\exists \lambda_x \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_x(\epsilon) \in (0, \lambda_x] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\hat{x}_i)_{i=1}^n, (\bar{x}_i)_{i=1}^n \subseteq \Omega \cap \bar{r}_{x-\lambda_x \mathbf{1}_m, x+\lambda_x \mathbf{1}_m}$ with $\hat{x}_i \leq \bar{x}_i$, $\forall i = 1, \dots, n$, $(r_{\hat{x}_i, \bar{x}_i})_{i=1}^n$ being pairwise disjoint, and $\sum_{i=1}^n \mu(r_{\hat{x}_i, \bar{x}_i}) < \delta_x(\epsilon)$, we have $\sum_{i=1}^n \|\Delta_F(r_{\hat{x}_i, \bar{x}_i})\| < \epsilon$. Then, $\bar{r}_{x_1, x_2} \subseteq \bigcup_{x \in \bar{r}_{x_1, x_2}} r_{x-\lambda_x \mathbf{1}_m, x+\lambda_x \mathbf{1}_m}^\circ \subset \mathbb{R}^m$. By the compactness of $\bar{r}_{x_1, x_2}$, there exists a finite set $X_N := \{\tilde{x}_1, \dots, \tilde{x}_p\} \subseteq \bar{r}_{x_1, x_2}$, where $p \in \mathbb{N}$, such that $\bar{r}_{x_1, x_2} \subseteq \bigcup_{i=1}^p r_{\tilde{x}_i - \lambda_{\tilde{x}_i} \mathbf{1}_m, \tilde{x}_i + \lambda_{\tilde{x}_i} \mathbf{1}_m}^\circ \subset \mathbb{R}^m$.

$\forall E \in \mathcal{B}$ with $\mu(E) = 0$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta(\epsilon) := \min_{i \in \{1, \dots, p\}} \delta_{\tilde{x}_i}(\frac{\epsilon}{2p}) \in (0, \infty) \subset \mathbb{R}$. By Proposition 11.23, $\exists O \in \mathcal{O}_{\mathbb{R}^m}$ with $E \subseteq O$ such that $\mu_{Bm}(O \setminus E) < \delta(\epsilon)$. Then, $\mu_{Bm}(O) = \mu_{Bm}(O \setminus E) + \mu_{Bm}(E) = \mu_{Bm}(O \setminus E) < \delta(\epsilon)$. By Proposition 12.40, O is a region with $P(O) = O$. By Definition 12.38, there exists $N \subseteq \mathbb{N}$ and $\exists (\hat{x}_{i,1})_{i \in N}, (\hat{x}_{i,2})_{i \in N} \subseteq O$ such that $\hat{x}_{i,1} \leq \hat{x}_{i,2}$, $\forall i \in \mathbb{N}$, and $O = \bigcup_{i \in N} r_{\hat{x}_{i,1}, \hat{x}_{i,2}}$ and the sets in the union are pairwise disjoint. Then, $\mu_{Bm}(O) = \sum_{i \in N} \mu_{Bm}(r_{\hat{x}_{i,1}, \hat{x}_{i,2}}) < \delta(\epsilon)$ and $\mathcal{P} \circ \nu(E) \leq \mathcal{P} \circ \nu(O \cap r_{x_1, x_2}) = \sum_{i \in N} \mathcal{P} \circ \nu(r_{\hat{x}_{i,1}, \hat{x}_{i,2}} \cap r_{x_1, x_2})$. By Definition 12.41, for any fixed $i \in N$, define $r_{\tilde{x}_{i,1}, \tilde{x}_{i,2}} := r_{\hat{x}_{i,1}, \hat{x}_{i,2}} \cap r_{x_1, x_2}$ with $\tilde{x}_{i,j} \in \bar{r}_{x_1, x_2}$, $j = 1, 2$, and $\tilde{x}_{i,1} \leq \tilde{x}_{i,2}$, we have $O \cap r_{x_1, x_2} = \bigcup_{i \in N} r_{\tilde{x}_{i,1}, \tilde{x}_{i,2}}$. By possibly splitting rectangles into smaller disjoint pieces in the collection, without loss of generality, assume that $\forall i \in N$, $\exists j \in \{1, \dots, p\}$ such that $r_{\tilde{x}_{i,1}, \tilde{x}_{i,2}} \subseteq r_{\tilde{x}_j - \lambda_{\tilde{x}_j} \mathbf{1}_m, \tilde{x}_j + \lambda_{\tilde{x}_j} \mathbf{1}_m}$. Then, $\forall i \in N$, $\mathcal{P} \circ \nu(r_{\tilde{x}_{i,1}, \tilde{x}_{i,2}}) = T_F(r_{\tilde{x}_{i,1}, \tilde{x}_{i,2}}) = \sup_{n \in \mathbb{Z}_+, \tilde{x}_{1,1}, \tilde{x}_{1,2}, \dots, \tilde{x}_{n,1}, \tilde{x}_{n,2} \in \bar{r}_{\tilde{x}_{i,1}, \tilde{x}_{i,2}}, \tilde{x}_{j,1} \leq \tilde{x}_{j,2}, \forall j \in \{1, \dots, n\}} \sum_{j=1}^n \|\Delta_F(r_{\tilde{x}_{j,1}, \tilde{x}_{j,2}})\|$. $\bigcup_{j=1}^n r_{\tilde{x}_{j,1}, \tilde{x}_{j,2}} = r_{\tilde{x}_{i,1}, \tilde{x}_{i,2}}, r_{\tilde{x}_{k,1}, \tilde{x}_{k,2}} \cap r_{\tilde{x}_{j,1}, \tilde{x}_{j,2}} = \emptyset, \forall 1 \leq k < j \leq n$. Then, there exists $n_i \in \mathbb{N}$, $(\tilde{x}_{i,j,1})_{j=1}^{n_i}, (\tilde{x}_{i,j,2})_{j=1}^{n_i} \subseteq \bar{r}_{\tilde{x}_{i,1}, \tilde{x}_{i,2}}$ such that $\tilde{x}_{i,j,1} \leq \tilde{x}_{i,j,2}$, $\forall j \in \{1, \dots, n_i\}$, $\bigcup_{j=1}^{n_i} r_{\tilde{x}_{i,j,1}, \tilde{x}_{i,j,2}} = r_{\tilde{x}_{i,1}, \tilde{x}_{i,2}}$ and the sets in the union are pairwise disjoint, and $\mathcal{P} \circ \nu(r_{\tilde{x}_{i,1}, \tilde{x}_{i,2}}) < \sum_{j=1}^{n_i} \|\Delta_F(r_{\tilde{x}_{i,j,1}, \tilde{x}_{i,j,2}})\| +$

$2^{-i-1}\epsilon$. Fix any finite subset $N_f \subseteq N$, we have $\sum_{i \in N_f} \mathcal{P} \circ \nu(r_{\hat{x}_{i,1}, \hat{x}_{i,2}} \cap r_{x_1, x_2}) = \sum_{i \in N_f} \mathcal{P} \circ \nu(r_{\tilde{x}_{i,1}, \tilde{x}_{i,2}}) < \sum_{i \in N_f} \sum_{j=1}^{n_i} \|\Delta_F(r_{\tilde{x}_{i,j,1}, \tilde{x}_{i,j,2}})\| + \epsilon/2 =: \sum_{l=1}^{\tilde{n}} \|\Delta_F(r_{x_{j,1}, x_{j,2}})\| + \epsilon/2$, where $\tilde{n} \in \mathbb{Z}_+$ and $(x_{l,1})_{l=1}^{\tilde{n}}, (x_{l,2})_{l=1}^{\tilde{n}} \subseteq \bar{r}_{x_1, x_2}$ with $\sum_{i=1}^{\tilde{n}} \mu(r_{x_{i,1}, x_{i,2}}) = \sum_{i=1}^{\tilde{n}} \mu_{Bm}(r_{x_{i,1}, x_{i,2}}) \leq \mu_{Bm}(O \cap r_{x_1, x_2}) < \delta(\epsilon)$. We may rearrange and partition the collection into p segments: $r_{x_{1,1}, x_{1,2}}, \dots, r_{x_{i_1,1}, x_{i_1,2}} \subseteq r_{\tilde{x}_1 - \lambda_{\tilde{x}_1} \mathbf{1}_m, \tilde{x}_1 + \lambda_{\tilde{x}_1} \mathbf{1}_m}; r_{x_{i_1+1,1}, x_{i_1+1,2}}, \dots, r_{x_{i_2,1}, x_{i_2,2}} \subseteq r_{\tilde{x}_2 - \lambda_{\tilde{x}_2} \mathbf{1}_m, \tilde{x}_2 + \lambda_{\tilde{x}_2} \mathbf{1}_m}; \dots; r_{x_{i_{p-1}+1,1}, x_{i_{p-1}+1,2}}, \dots, r_{x_{i_p,1}, x_{i_p,2}} \subseteq r_{\tilde{x}_p - \lambda_{\tilde{x}_p} \mathbf{1}_m, \tilde{x}_p + \lambda_{\tilde{x}_p} \mathbf{1}_m}$. Note that $\forall j \in \{1, \dots, p\}$, $\sum_{l=i_{j-1}+1}^{i_j} \mu_{Bm}(r_{x_{l,1}, x_{l,2}}) < \delta(\epsilon) \leq \delta_{\tilde{x}_j}(\frac{\epsilon}{2p})$, which implies that $\sum_{l=i_{j-1}+1}^{i_j} \|\Delta_F(r_{x_{j,1}, x_{j,2}})\| < \frac{\epsilon}{2p}$. Then, we have $\sum_{i \in N_f} \mathcal{P} \circ \nu(r_{\hat{x}_{i,1}, \hat{x}_{i,2}} \cap r_{x_1, x_2}) < \epsilon$. By the arbitrariness of N_f , we have $\mathcal{P} \circ \nu(E) \leq \mathcal{P} \circ \nu(O \cap r_{x_1, x_2}) = \sum_{i \in N} \mathcal{P} \circ \nu(r_{\hat{x}_{i,1}, \hat{x}_{i,2}} \cap r_{x_1, x_2}) = \sum_{i \in N} \mathcal{P} \circ \nu(r_{\tilde{x}_{i,1}, \tilde{x}_{i,2}}) \leq \epsilon$. By the arbitrariness of ϵ , we have $\mathcal{P} \circ \nu(E) = 0$. By the arbitrariness of E , we have $\mathcal{P} \circ \nu \ll \mu$. This proves that $\mathcal{P} \circ \nu \ll \mu$ in the special case.

Consider the general case when Ω is a rectangle. By Proposition 12.39, Ω is a region and $P(\Omega)$ is a rectangle. By Definition 12.38, $\exists N \subseteq \mathbb{N}$ and $(\bar{x}_i)_{i \in N}, (\tilde{x}_i)_{i \in N} \subseteq \Omega$ with $\bar{x}_i \leq \tilde{x}_i, \forall i \in N$, such that $\Omega = \bigcup_{i \in N} \bar{r}_{\bar{x}_i, \tilde{x}_i}$, $P(\Omega) = \bigcup_{i \in N} \bar{r}_{\bar{x}_i, \tilde{x}_i}$, and the sets in the second union are pairwise disjoint. $\forall E \in \mathcal{B}$ with $\mu(E) = 0$. Then, $\mu(E \cap r_{\bar{x}_i, \tilde{x}_i}) = 0, \forall i \in N$. By the special case, we have $\mathcal{P} \circ \nu(E \cap r_{\bar{x}_i, \tilde{x}_i}) = 0, \forall i \in N$. Then, we have $\mathcal{P} \circ \nu(E) = \sum_{i \in N} \mathcal{P} \circ \nu(E \cap r_{\bar{x}_i, \tilde{x}_i}) = 0$. By the arbitrariness of E , we have $\mathcal{P} \circ \nu \ll \mu$. This completes the proof of the proposition. $\square$

Lemma 12.74 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space and ν be a measure on $(X, \mathcal{B})$. The following statements hold.

- (i) $\nu \ll \mu$ and ν is finite implies that $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}, \forall E \in \mathcal{B}$ with $\mu(E) < \delta$, we have $\nu(E) < \epsilon$.
- (ii) $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}, \forall E \in \mathcal{B}$ with $\mu(E) < \delta$, we have $\nu(E) < \epsilon$. Then, $\nu \ll \mu$.

Proof (i) We will prove the result by an argument of contradiction. Suppose the result does not hold. Then, $\exists \epsilon_0 \in (0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N}, \exists E_n \in \mathcal{B}$ with $\mu(E_n) < 2^{-n-1}$ such that $\nu(E_n) > \epsilon_0$. Let $\bar{E}_n := \bigcup_{j=n}^{\infty} E_j, \forall n \in \mathbb{N}$. Then, $\bar{E}_n \supseteq \bar{E}_{n+1}$ and $\mu(\bar{E}_n) \leq \sum_{j=n}^{\infty} \mu(E_j) < 2^{-n}$. By Proposition 11.5, we have $0 = \lim_{n \in \mathbb{N}} \mu(\bar{E}_n) = \mu(\bigcap_{n=1}^{\infty} \bar{E}_n)$. Then, by $\nu \ll \mu$ and Proposition 11.5, we have $0 = \nu(\bigcap_{n=1}^{\infty} \bar{E}_n) = \lim_{n \in \mathbb{N}} \nu(\bar{E}_n)$. This contradicts with the fact that $\nu(\bar{E}_n) \geq \nu(E_n) > \epsilon_0, \forall n \in \mathbb{N}$. Hence, the result holds.

(ii) $\forall E \in \mathcal{B}$ with $\mu(E) = 0$. By the assumption, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\forall \bar{E} \in \mathcal{B}$ with $\mu(\bar{E}) < \delta$, we have $\nu(\bar{E}) < \epsilon$. In particular, $\mu(E) = 0 < \delta$, and this implies that $\nu(E) < \epsilon$. By the arbitrariness of ϵ , we have $\nu(E) = 0$. By the arbitrariness of E , we have $\nu \ll \mu$.

This completes the proof of the lemma. $\square$

Proposition 12.75 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, ν be a σ -finite $\mathcal{Y}$ -valued measure on the measurable space $(P(\Omega), \mathcal{B}_B(P(\Omega)))$, and $\mathbb{X} := (P(\Omega), \mathcal{B}_B(P(\Omega)), \mu)$ be the metric measure subspace of $\mathbb{R}^m$. Assume that $\mathcal{P} \circ \nu \ll \mu$ and $\mathcal{P} \circ \nu(r_{x_1, x_2}) < \infty$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. By Proposition 12.52, we may define $F : \Omega \rightarrow \mathcal{Y}$ to be the cumulative distribution function of ν with origin x_s , where x_s is an arbitrary fixed vector in Ω if $\Omega \neq \emptyset$. Then, F is of locally bounded variation and is absolutely continuous.*

In particular, if $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $M \subseteq \mathcal{Y}$ be a separable subspace, and $f : P(\Omega) \rightarrow M$ be $\mathcal{B}$ -measurable. We may define ν to be the $\mathcal{Y}$ -valued measure with kernel f over $\mathbb{X}$ as outlined in Proposition 11.116. Then, f is the Radon-Nikodym derivative of ν with respect to μ . Assume that, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$, f is absolutely integrable on r_{x_1, x_2} . Then, ν satisfies the assumption of the previous paragraph. Let $F : \Omega \rightarrow \mathcal{Y}$ be the cumulative distribution function of ν with origin at x_s as defined in Proposition 12.52. Then, F is absolutely continuous.

Proof We will prove that F is absolutely continuous at x_0 , $\forall x_0 \in \Omega$, by Proposition 12.60. Fix any $x_0 \in \Omega$, any $i \in \{0, \dots, m-1\}$, and any $J \subset \bar{J} := \{1, \dots, m\}$ with $\text{card}(J) = i$. Let $\pi_J : \mathbb{R}^m \rightarrow \mathbb{R}^i$, $\pi_{\bar{J} \setminus J} : \mathbb{R}^m \rightarrow \mathbb{R}^{m-i}$, $M : \mathbb{R}^i \times \mathbb{R}^{m-i} \rightarrow \mathbb{R}^m$, x_f , $\tilde{x}_f$, $\tilde{\Omega}_J$, and $F_J : \tilde{\Omega}_J \rightarrow \mathcal{Y}$ be as defined in (a) Proposition 12.60. Define $\tilde{x} := M(\pi_J(x_s), \pi_{\bar{J} \setminus J}(x_0)) \in \Omega$, $E := r_{x_f \wedge \pi_J(x_s), x_f \vee \pi_J(x_s)}$, and $E_n := r_{\tilde{x}_f - \mathbf{1}_{m-i}/n, \tilde{x}_f + \mathbf{1}_{m-i}/n} \cap \tilde{\Omega}_J$, $\forall n \in \mathbb{N}$. Clearly, $V_n := M(E \times E_n) \subseteq \Omega$ and $\mu_{Bm}(P(V_n)) = \mu_{Bm}(V_n) = \mu_{Bi}(E)\mu_{B(m-i)}(E_n) \leq (2/n)^{m-i}\mu_{Bi}(E) < \infty$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} \mu_{Bm}(P(V_n)) = \mu_{Bm}(\bigcap_{n=1}^{\infty} P(V_n)) = 0$, by Proposition 11.5. Then, by $\mathcal{P} \circ \nu \ll \mu$, we have $\mathcal{P} \circ \nu(\bigcap_{n=1}^{\infty} P(V_n)) = 0$. Since Ω is a rectangle, then $\exists n_0 \in \mathbb{N}$ such that $P(V_{n_0}) = r_{x_1, x_2}$ for some $x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. Then, $\mathcal{P} \circ \nu(P(V_{n_0})) = \mathcal{P} \circ \nu(r_{x_1, x_2}) < \infty$. Let $a_J := 1/(2n_0) \in (0, \infty) \subset \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by Lemma 12.74, $\exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall E \in \mathcal{B}_B(P(V_{n_0}))$ with $\mu(E) = \mu_{Bm}(E) < \delta(\epsilon)$, we have $\mathcal{P} \circ \nu(E) < \epsilon$. Take $\bar{\delta}_J(\epsilon) := \delta(\epsilon)/(1 + \mu_{Bi}(E)) \in (0, \infty) \subset \mathbb{R}$. $\forall n \in \mathbb{N}$, $\forall (\bar{x}_j)_{j=1}^n, (\tilde{x}_j)_{j=1}^n \subseteq \bar{r}_{\tilde{x}_f - a_J \mathbf{1}_{m-i}, \tilde{x}_f + a_J \mathbf{1}_{m-i}} \cap \tilde{\Omega}_J$ with $\bar{x}_j \leq \tilde{x}_j$, $\forall j \in \{1, \dots, n\}$, $(r_{\bar{x}_j, \tilde{x}_j})_{j=1}^n$ being pairwise disjoint, and $\sum_{j=1}^n \mu_{B(m-i)}(r_{\bar{x}_j, \tilde{x}_j}) < \bar{\delta}_J(\epsilon)$, we have $\sum_{j=1}^n \|\Delta_{F_J}(r_{\bar{x}_j, \tilde{x}_j})\| = \sum_{j=1}^n \|\Delta_F(M(E \times r_{\bar{x}_j, \tilde{x}_j}))\| = \sum_{j=1}^n \|\nu(M(E \times r_{\bar{x}_j, \tilde{x}_j}))\| \leq \sum_{j=1}^n \mathcal{P} \circ \nu(M(E \times r_{\bar{x}_j, \tilde{x}_j})) = \mathcal{P} \circ \nu(M(E \times (\bigcup_{j=1}^n r_{\bar{x}_j, \tilde{x}_j})))$. Since $\bigcup_{j=1}^n r_{\bar{x}_j, \tilde{x}_j} \subseteq E_{n_0}$ with $\mu_{B(m-i)}(\bigcup_{j=1}^n r_{\bar{x}_j, \tilde{x}_j}) < \bar{\delta}_J(\epsilon)$, then $\mu_{Bm}(M(E \times (\bigcup_{j=1}^n r_{\bar{x}_j, \tilde{x}_j}))) = \mu_{Bi}(E)\mu_{Bm-i}(\bigcup_{j=1}^n r_{\bar{x}_j, \tilde{x}_j}) < \delta(\epsilon)$ and $M(E \times (\bigcup_{j=1}^n r_{\bar{x}_j, \tilde{x}_j})) = \bigcup_{j=1}^n M(E \times r_{\bar{x}_j, \tilde{x}_j}) = \bigcup_{j=1}^n P(M(E \times r_{\bar{x}_j, \tilde{x}_j})) \subseteq P(\bigcup_{j=1}^n M(E \times r_{\bar{x}_j, \tilde{x}_j})) = P(M(E \times (\bigcup_{j=1}^n r_{\bar{x}_j, \tilde{x}_j}))) \subseteq P(M(E \times E_{n_0})) =$

$P(V_{n_0})$, where the second equality, the first set containment, the third equality, and the second set containment follow from Proposition 12.39. Hence, $\mathcal{P} \circ \nu(M(E \times (\bigcup_{j=1}^n r_{\bar{x}_j, \bar{x}_j}))) < \epsilon$. Therefore, $\sum_{j=1}^n \|\Delta_{F_j}(r_{\bar{x}_j, \bar{x}_j})\| < \epsilon$.

By Proposition 12.60, F is absolutely continuous at x_0 . By the arbitrariness of x_0 , F is absolutely continuous. By Proposition 12.73, F is of locally bounded variation. This completes the proof of the first paragraph of the proposition.

By Proposition 11.116, $(P(\Omega), \mathcal{B}_B(P(\Omega)), \nu)$ is a σ -finite $\mathcal{Y}$ -valued measure space and $\mathcal{P} \circ \nu = \bar{\nu}$, where $\bar{\nu}$ is a σ -finite measure on $(P(\Omega), \mathcal{B}_B(P(\Omega)))$ defined by $\bar{\nu}(E) = \int_E \mathcal{P} \circ f d\mu \in [0, \infty] \subset \mathbb{R}_e$, $\forall E \in \mathcal{B}_B(P(\Omega))$. $\forall E \in \mathcal{B}_B(P(\Omega))$ with $\mu(E) = 0$, we have $\mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ f d\mu_{Bm} = 0$, where the last equality follows from Definition 11.79 and Proposition 11.83. Hence, $\mathcal{P} \circ \nu \ll \mu$. By Definition 11.166, f is the Radon-Nikodym derivative of ν with respect to μ . $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$, we have $r_{x_1, x_2} \subseteq P(\Omega)$. By the assumption, we have $\mathcal{P} \circ \nu(r_{x_1, x_2}) < \infty$. Thus, ν satisfies all the assumption of the first paragraph of the proposition. This implies that F is absolutely continuous. This completes the proof of the proposition. $\square$

The preceding results states that if a function of a rectangle in $\mathbb{R}^m$ into a normed linear space, then, the function being absolutely continuous implies it is of locally bounded variation. If the normed linear space is a Banach space, then there exists a unique Banach space valued measure whose cumulative distribution function is the given function, and this measure is absolutely continuous with respect to the restriction of μ_{Bm} to the principal. This will further lead to the existence of a unique Radon-Nikodym derivative of the Banach space valued measure with respect to the restriction of μ_{Bm} , if the Banach space is separable and reflexive whose dual is also separable. When, the range space is only a Banach space, then the Banach space valued measure may or may not admit a Radon-Nikodym derivative with respect to the restriction of μ_{Bm} to the principal.

Starting from a measurable function of the principal of a rectangle in $\mathbb{R}^m$ to a separable subspace of a Banach space, we may define a σ -finite Banach space valued measure on the principal. The cumulative distribution function of the Banach space valued measure with origin at some x_0 inside the rectangle exists and is absolutely continuous under the mild condition that the measurable function we started with is absolutely integrable on any semi-open rectangles in $\mathbb{R}^m$ with corners inside the rectangle. In this case, we may further conclude that f is the Radon-Nikodym derivative of the Banach space valued measure with respect to the restriction of μ_{Bm} to the principal.

12.6 Fundamental Theorem of Calculus

Definition 12.76 Let $\mathcal{X} := (X, \rho)$ be a metric space and $S \subseteq X$. The diameter of S is defined to be $\text{dia}(S) = \max\{\sup_{x_1, x_2 \in S} \rho(x_1, x_2), 0\} \in$

$[0, \infty] \subset \mathbb{R}_e$.

Definition 12.77 Let $m \in \mathbb{N}$, $E \subseteq \mathbb{R}^m$ and $\mathcal{I} \subseteq \mathcal{B}_B(\mathbb{R}^m)$ be a collection of nondegenerate rectangles in $\mathbb{R}^m$ (that is, $\exists$ nondegenerate intervals $I_1, \dots, I_m$ such that $U = \prod_{i=1}^m I_i$, $\forall U \in \mathcal{I}$), i. e., $\mu_{Bm}(U) > 0$, $\forall U \in \mathcal{I}$. We will say that $\mathcal{I}$ covers E in the sense of Vitali with index $c \in (0, \frac{1}{\sqrt{m}}] \subset \mathbb{R}$ if $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x \in E$, $\exists U \in \mathcal{I}$ such that $x \in U$, $\text{dia}(U) < \epsilon$, and the shortest side of the rectangle U has length at least $c \text{dia}(U)$.

Lemma 12.78 (Vitali) Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $E \subseteq \mathbb{R}^m$ with $\mu_{Lmo}(E) < +\infty$ and $\mathcal{I} \subseteq \mathcal{B}_B(\mathbb{R}^m)$ be a collection of nondegenerate rectangles in $\mathbb{R}^m$ that covers E in the sense of Vitali with index $c \in (0, \frac{1}{\sqrt{m}}] \subset \mathbb{R}$. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists a finite pairwise disjoint subcollection $\{U_1, \dots, U_n\} \subseteq \mathcal{I}$, where $n \in \mathbb{Z}_+$, such that $\mu_{Lmo}(E \setminus (\bigcup_{i=1}^n U_i)) < \epsilon$.

Proof Let $\hat{\mathcal{I}} := \{\bar{U} \subseteq \mathbb{R}^m \mid U \in \mathcal{I}\}$. Then, $\hat{\mathcal{I}}$ is a collection of nondegenerate closed rectangles in $\mathbb{R}^m$ that covers E in the sense of Vitali with index c . We will first show that the result holds for $\hat{\mathcal{I}}$. Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. By Example 12.56, $\mu_{Lmo}(E) = \inf_{(O_i)_{i=1}^\infty \in \mathcal{O}_{\mathbb{R}^m}, E \subseteq \bigcup_{i=1}^\infty O_i} \sum_{i=1}^\infty \mu_{Bm}(O_i) < +\infty$. Then, $\exists O \in \mathcal{O}_{\mathbb{R}^m}$ such that $E \subseteq O$ and $\mu_{Bm}(O) < +\infty$. By neglecting rectangles in $\hat{\mathcal{I}}$, we may without loss of generality, assume that $U \subseteq O$ and the shortest side of U has length at least $c \text{dia}(U)$, $\forall U \in \hat{\mathcal{I}}$. $\forall k \in \mathbb{Z}_+$, assume that pairwise disjoint $U_1, \dots, U_k \in \hat{\mathcal{I}}$ has already be chosen. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $E \subseteq \bigcup_{i=1}^k U_i$; Case 2: $E \setminus (\bigcup_{i=1}^k U_i) \neq \emptyset$. Case 1: $E \subseteq \bigcup_{i=1}^k U_i$. Then, $\mu_{Lmo}(E \setminus (\bigcup_{i=1}^k U_i)) = 0$. The result holds for $\hat{\mathcal{I}}$. Case 2: $E \setminus (\bigcup_{i=1}^k U_i) \neq \emptyset$. Note that, for any closed rectangle $U \in \hat{\mathcal{I}}$ with $\text{dia}(U) =: \bar{p} > 0$, let $x_0 \in U$ be the center of U , such that $\bar{x}_{x_0 - \frac{1}{2}c\bar{p}\mathbf{1}_m, x_0 + \frac{1}{2}c\bar{p}\mathbf{1}_m} \subseteq U \subseteq \bar{x}_{x_0 - \frac{1}{2}\hat{p}\mathbf{1}_m, x_0 + \frac{1}{2}\hat{p}\mathbf{1}_m}$, where $\hat{p} := \sqrt{1 - (m-1)c^2\bar{p}}$ is the upper bound for the longest side of the rectangle U . This implies that $c^m\bar{p}^m \leq \mu_{Bm}(U) \leq \hat{p}^m = (1 - (m-1)c^2)^{m/2}\bar{p}^m$. Let l_k be the supremum of the diameters of rectangles in $\hat{\mathcal{I}}$ that do not intersect $U_1, \dots, U_k$. Clearly, $l_k \leq \frac{\sqrt{\mu_{Bm}(O)}}{c} < +\infty$. Then, $\exists x_0 \in E \setminus (\bigcup_{i=1}^k U_i)$. Since $\bigcup_{i=1}^k U_i$ is closed, then, by Proposition 4.10, $d := \inf_{x \in \bigcup_{i=1}^k U_i} |x - x_0| > 0$. Then, by the assumption, $\exists \hat{U} \in \hat{\mathcal{I}}$ such that $x_0 \in \hat{U}$ and $\text{dia}(\hat{U}) < d$. Then, $\hat{U} \cap (\bigcup_{i=1}^k U_i) = \emptyset$. This shows that $l_k \geq \text{dia}(\hat{U}) > 0$. Hence, $\exists U_{k+1} \in \hat{\mathcal{I}}$ such that $U_1, \dots, U_{k+1} \in \hat{\mathcal{I}}$ are pairwise disjoint and $\text{dia}(U_{k+1}) \geq l_k/2$. Inductively, we either have the result holds for $\hat{\mathcal{I}}$ or $\exists (U_k)_{k=1}^\infty \subseteq \hat{\mathcal{I}}$ such that $\text{dia}(U_{k+1}) \geq l_k/2$, $\forall k \in \mathbb{Z}_+$ and the sequence is pairwise disjoint.

In the latter case, we have $O \supseteq \bigcup_{k=1}^\infty U_k$ and $+\infty > \mu_{Bm}(O) \geq \mu_{Bm}(\bigcup_{k=1}^\infty U_k) = \sum_{k=1}^\infty \mu_{Bm}(U_k) \geq \sum_{k=0}^\infty 2^{-m}c^m l_k^m$. Note that $(l_k)_{k=0}^\infty$ is nonincreasing. Hence, $\lim_{k \in \mathbb{N}} l_k = 0$. Then, $\exists n \in \mathbb{N}$ such that $\sum_{i=n+1}^\infty \mu_{Bm}(U_i) < \pi^{-m/2}2^m 5^{-m}c^m \Gamma(\frac{m}{2} + 1)\epsilon$, where we have referred

to the special Gamma function. Let $R := E \setminus (\bigcup_{k=1}^n U_k)$. Without loss of generality, assume $\text{dia}(U_k) =: p_k \in (0, \infty) \subset \mathbb{R}$ and the center of U_k is x_k , $\forall k \in \mathbb{N}$. $\forall \bar{x} \in R$, since $\bigcup_{k=1}^n U_k$ is closed, then, by Proposition 4.10, $d := \inf_{x \in \bigcup_{k=1}^n U_k} |x - \bar{x}| > 0$. By the assumption, $\exists U \in \hat{\mathcal{I}}$ such that $\bar{x} \in U$ and $0 < \text{dia}(U) < d$. Then, $U \cap (\bigcup_{k=1}^n U_k) = \emptyset$. By the fact that $\lim_{k \in \mathbb{N}} l_k = 0$, we have $\exists i_0 \in \{n+1, n+2, \dots\}$ such that $U \cap U_{i_0} \neq \emptyset$ and $U \cap U_k = \emptyset$, $\forall k \in \{1, \dots, i_0 - 1\}$. Then, $\text{dia}(U) \leq l_{i_0-1} \leq 2\text{dia}(U_{i_0}) = 2p_{i_0}$. Let $\hat{x} \in U \cap U_{i_0}$. Then, $|\bar{x} - x_{i_0}| \leq |\bar{x} - \hat{x}| + |\hat{x} - x_{i_0}| \leq \text{dia}(U) + \frac{1}{2}p_{i_0} \leq \frac{5}{2}p_{i_0}$. Then, $\bar{x} \in V_{i_0} := \bar{B}_{\mathbb{R}^m}(x_{i_0}, \frac{5}{2}p_{i_0}) \subset \mathbb{R}^m$. Then, $\bar{x} \in \bigcup_{k=n+1}^{\infty} V_k := \bigcup_{k=n+1}^{\infty} \bar{B}_{\mathbb{R}^m}(x_k, \frac{5}{2}p_k) \subset \mathbb{R}^m$. Hence, $R \subseteq \bigcup_{k=n+1}^{\infty} V_k$. Then, $\mu_{Lmo}(R) \leq \mu_{Lmo}(\bigcup_{k=n+1}^{\infty} V_k) = \sum_{k=n+1}^{\infty} \mu_{Bm}(V_k) = \sum_{k=n+1}^{\infty} \frac{\pi^{m/2} 2^{-m} 5^m p_k^m}{\Gamma(\frac{m}{2}+1)} \leq \frac{\pi^{m/2} 2^{-m} 5^m c^{-m}}{\Gamma(\frac{m}{2}+1)} \sum_{k=n+1}^{\infty} c^m p_k^m \leq \frac{\pi^{m/2} 2^{-m} 5^m c^{-m}}{\Gamma(\frac{m}{2}+1)} \cdot \sum_{k=n+1}^{\infty} \mu_{Bm}(U_k) < \epsilon$, where the second equality follows from (Mathematics Handbook Editors Group, 1979, pp. 320). Hence, the result holds for $\hat{\mathcal{I}}$.

Then, $\exists n \in \mathbb{Z}_+$, $\exists \{U_1, \dots, U_n\} \subseteq \mathcal{I}$ such that $\mu_{Lmo}(E \setminus (\bigcup_{i=1}^n \bar{U}_i)) < \epsilon$. Note that $\mu_{Lmo}(J) := \mu_{Lmo}((\bigcup_{i=1}^n \bar{U}_i) \setminus (\bigcup_{i=1}^n U_i)) = 0$. Then, $J \in \mathcal{B}_L$ and $\mu_{Lmo}(E \setminus (\bigcup_{i=1}^n U_i)) = \mu_{Lmo}((E \setminus (\bigcup_{i=1}^n U_i)) \cap J) + \mu_{Lmo}((E \setminus (\bigcup_{i=1}^n U_i)) \setminus J) \leq \mu_{Lmo}(E \setminus (\bigcup_{i=1}^n \bar{U}_i)) < \epsilon$, where the equality follows from Definition 11.15; and the first inequality follows from fact that $\mu_{Lmo}((E \setminus (\bigcup_{i=1}^n U_i)) \cap J) \leq \mu_{Lmo}(J) = 0$ and the fact that $(E \setminus (\bigcup_{i=1}^n U_i)) \setminus J = E \setminus (\bigcup_{i=1}^n \bar{U}_i)$.

This completes the proof of the lemma. $\square$

Definition 12.79 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $E \in \mathcal{O}_{\mathbb{R}^m}$ with $\mu_{Bm}(E) > 0$, $\mathbb{E} := ((E, |\cdot|), \mathcal{B}, \mu)$ be the σ -finite metric measure subspace of $\mathbb{R}^m$ as defined in Proposition 11.29, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, $c \in (0, \frac{1}{\sqrt{m}}] \subset \mathbb{R}$, and $f \in \bar{L}_1(\mathbb{E}, \mathcal{Y})$. A point $x \in E$ is said to be a rectangular Lebesgue point with regularity c of f if $\limsup_{r \rightarrow 0+} \sup_{\substack{r_{x_1, x_2} \subset \mathbb{R}^m \text{ with } \frac{1}{2}(x_1+x_2)=x, \\ |x_2-x_1|=r, \min(x_2-x_1) \geq cr}} \frac{1}{\mu(r_{x_1, x_2} \cap E)} \int_{r_{x_1, x_2} \cap E} \|f(y) - f(x)\| d\mu(y) = 0$.

Proposition 12.80 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $E \in \mathcal{O}_{\mathbb{R}^m}$ with $\mu_{Bm}(E) > 0$, $\mathbb{E} := ((E, |\cdot|), \mathcal{B}, \mu)$ be the σ -finite metric measure subspace of $\mathbb{R}^m$ as defined in Proposition 11.29, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, $c \in (0, \frac{1}{\sqrt{m}}] \subset \mathbb{R}$, $f \in \bar{L}_1(\mathbb{E}, \mathcal{Y})$, and $A := \{x \in E \mid x \text{ is a rectangular Lebesgue point of } f \text{ with regularity } c\}$. Then, $A \in \mathcal{B}_{Lm}$ and $\mu_{Lm}(E \setminus A) = 0$.

Proof Since $\mathbb{E}$ is σ -finite, then $\exists (E_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $\bigcup_{n=1}^{\infty} E_n = E$ and $\mu(E_n) < +\infty$, $\forall n \in \mathbb{N}$. $\forall r \in (0, \infty) \subset \mathbb{R}$, $\forall x \in E$, define $T_{r,c}f : E \rightarrow [0, \infty) \subset \mathbb{R}$ and $M_{r,c}f : E \rightarrow [0, \infty) \subset \mathbb{R}$ by, $\forall x \in E$,

$$(T_{r,c}f)(x) = \sup_{\substack{r_{x_1, x_2} \subset \mathbb{R}^m \text{ with } \frac{1}{2}(x_1+x_2)=x, \\ |x_2-x_1|=r, \min(x_2-x_1) \geq cr}} \frac{1}{\mu(r_{x_1, x_2} \cap E)}$$

$$(M_{r,c}f)(x) = \sup_{\substack{r_{x_1, x_2} \subset \mathbb{R}^m \text{ with } \frac{1}{2}(x_1 + x_2) = x, \\ |x_2 - x_1| = r, \min(x_2 - x_1) \geq cr}} \frac{1}{\mu(r_{x_1, x_2} \cap E)} \int_{r_{x_1, x_2} \cap E} \mathcal{P} \circ f \, d\mu$$

$\forall x \in E$, since $E \in \mathcal{O}_{\mathbb{R}^m}$ then $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $r_{x-\delta \mathbf{1}_m, x+\delta \mathbf{1}_m}^\circ \subseteq E$. $\forall r_{x_1, x_2} \subset \mathbb{R}^m$ with $x = \frac{1}{2}(x_1 + x_2)$, $|x_2 - x_1| = r > 0$ and $\min(x_2 - x_1) \geq cr$, we have $\mu(r_{x_1, x_2} \cap E) = \mu_{Bm}(r_{x_1, x_2} \cap E) \geq \mu_{Bm}(r_{x_1, x_2} \cap r_{x-\delta \mathbf{1}_m, x+\delta \mathbf{1}_m}^\circ) \geq ((2\delta) \wedge (cr))^m > 0$. Then, $(M_{r,c}f)(x) \leq ((2\delta) \wedge (cr))^{-m} \|f\|_1 < \infty$ and $(T_{r,c}f)(x) \leq (M_{r,c}f)(x) + \|f(x)\| < \infty$. Hence, $T_{r,c}f$ and $M_{r,c}f$ are well-defined. Define $T_cf : E \rightarrow [0, \infty] \subset \mathbb{R}_e$ and $M_cf : E \rightarrow [0, \infty] \subset \mathbb{R}_e$ by

$$(T_cf)(x) = \limsup_{r \rightarrow 0^+} (T_{r,c}f)(x) \quad (M_cf)(x) = \limsup_{r \rightarrow 0^+} (M_{r,c}f)(x); \quad \forall x \in E$$

Clearly, T_cf and M_cf are well-defined. Then, $T_{r,c}$ and $M_{r,c}$ are functions of $\bar{L}_1(\mathbb{E}, \mathcal{Y})$ to nonnegative real-valued functions of E , and T_c and M_c are functions of $\bar{L}_1(\mathbb{E}, \mathcal{Y})$ to nonnegative extended real-valued functions of E . Then, $A = \{x \in E \mid (T_cf)(x) = 0\}$ and $E \setminus A = \{x \in E \mid (T_cf)(x) > 0\} = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} A_{n,k} := \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} \{x \in E_n \mid (T_cf)(x) > 1/k\}$.

We will show that $\mu_{Lmo}(A_{n,k}) = 0, \forall n \in \mathbb{N}, \forall k \in \mathbb{N}$, which implies that $A_{n,k} \in \mathcal{B}_{Lm}$ and $\mu_{Lm}(A_{n,k}) = 0, \forall n \in \mathbb{N}, \forall k \in \mathbb{N}$. This further implies that $E \setminus A \in \mathcal{B}_{Lm}$, $\mu_{Lm}(E \setminus A) = 0$, and $A \in \mathcal{B}_{Lm}$.

Fix any $n \in \mathbb{N}$ and any $k \in \mathbb{N}$. Suppose $\mu_{Lmo}(A_{n,k}) > 0$. By monotonicity of outer measures, we have $\mu_{Lmo}(A_{n,k}) \leq \mu_{Lmo}(E_n) = \mu_{Lm}(E_n) = \mu(E_n) < +\infty$. Let $\epsilon_0 = \frac{\mu_{Lmo}(A_{n,k})}{5k} \in (0, +\infty) \subset \mathbb{R}$. By Propositions 11.182 and 4.11, there exists a continuous function $g : E \rightarrow \mathcal{Y}$ such that $g \in \bar{L}_1(\mathbb{E}, \mathcal{Y})$ and $\|g - f\|_1 < \epsilon_0$. Let $h := f - g$. Then, $\|h\|_1 < \epsilon_0$. Note that $(T_{r,c}f)(x) \leq (T_{r,c}g)(x) + (T_{r,c}h)(x), \forall r \in (0, \infty) \subset \mathbb{R}, \forall x \in E$. Then, by Proposition 3.85, $(T_cf)(x) \leq (T_cg)(x) + (T_ch)(x), \forall x \in E$. Since g is continuous, we have $(T_cg)(x) = 0, \forall x \in E$. Then, $(T_cf)(x) \leq (T_ch)(x), \forall x \in E$. Note that

$$(T_{r,c}h)(x) \leq (M_{r,c}h)(x) + \|h(x)\|; \quad \forall r \in (0, \infty) \subset \mathbb{R}, \forall x \in E$$

Then, by Proposition 3.85, we have $(T_ch)(x) \leq (M_ch)(x) + \|h(x)\|, \forall x \in E$. Hence, $(T_cf)(x) \leq (M_ch)(x) + \|h(x)\|, \forall x \in E$. Then, $A_{n,k} \subseteq \{x \in E_n \mid (M_ch)(x) > \frac{1}{2k}\} \cup \{x \in E_n \mid \|h(x)\| > \frac{1}{2k}\} =: \tilde{A}_{n,k} \cup \tilde{A}_{n,k}$. By monotonicity of outer measures, we have $\mu_{Lmo}(\tilde{A}_{n,k}) \leq \mu_{Lmo}(\{x \in E \mid \|h(x)\| > \frac{1}{2k}\}) =: \mu_{Lmo}(\hat{A}_k) = \mu_{Bm}(\hat{A}_k) = \int_E \chi_{\hat{A}_k, E} \, d\mu \leq \int_E (2k\mathcal{P} \circ h) \, d\mu = 2k\|h\|_1 < 2k\epsilon_0$, where the second equality follows from the fact that $\hat{A}_k \in \mathcal{B}$ since h is $\mathcal{B}$ -measurable. By countable subadditivity of outer measures, we have $\mu_{Lmo}(A_{n,k}) \leq \mu_{Lmo}(\tilde{A}_{n,k}) + \mu_{Lmo}(\tilde{A}_{n,k}) < \mu_{Lmo}(\tilde{A}_{n,k}) + 2k\epsilon_0$. Hence, $\mu_{Lmo}(\tilde{A}_{n,k}) > 3k\epsilon_0$. Clearly, $\mu_{Lmo}(\tilde{A}_{n,k}) \leq \mu_{Lmo}(E_n) < +\infty$.

Define $\mathcal{I} := \{r_{x_1, x_2} \subseteq \mathbb{R}^m \mid x = \frac{1}{2}(x_1 + x_2) \in E, |x_2 - x_1| =: r > 0, \min(x_2 - x_1) \geq cr, \int_{r_{x_1, x_2} \cap E} \mathcal{P} \circ h \, d\mu > \frac{1}{2k} \mu(r_{x_1, x_2} \cap E)\}$. Then, $\mathcal{I}$ covers $\tilde{A}_{n,k}$ in the sense of Vitali with index c . By Vitali's Lemma 12.78, there exists pairwise disjoint rectangles $(r_{x_{i,1}, x_{i,2}})_{i=1}^{\bar{m}} \subseteq \mathcal{I}$ with $\bar{m} \in \mathbb{Z}_+$ such that $\mu_{Lmo}(\tilde{A}_{n,k} \setminus (\bigcup_{i=1}^{\bar{m}} r_{x_{i,1}, x_{i,2}})) < \epsilon_0$. Let $B = (\bigcup_{i=1}^{\bar{m}} r_{x_{i,1}, x_{i,2}}) \cap E \in \mathcal{B}$. Then, $\mu(B) = \mu_{Lmo}(B) \geq \mu_{Lmo}(B \cap \tilde{A}_{n,k}) = \mu_{Lmo}(\tilde{A}_{n,k}) - \mu_{Lmo}(\tilde{A}_{n,k} \setminus B) > (3k-1)\epsilon_0$, where the first inequality follows from the monotonicity of outer measures; and the second equality follows from the measurability of B . Then, $\|h\|_1 = \int_E \mathcal{P} \circ h \, d\mu \geq \int_B \mathcal{P} \circ h \, d\mu = \sum_{i=1}^{\bar{m}} \int_{r_{x_{i,1}, x_{i,2}} \cap E} \mathcal{P} \circ h \, d\mu > \sum_{i=1}^{\bar{m}} \frac{1}{2k} \mu(r_{x_{i,1}, x_{i,2}} \cap E) = \frac{1}{2k} \mu(B) > \frac{3k-1}{2k} \epsilon_0 \geq \epsilon_0 > \|h\|_1$. This is a contradiction. Therefore, we must have $\mu_{Lmo}(A_{n,k}) = 0$.

This completes the proof of the proposition. $\square$

In the rest of the notes, we will denote the i th unit vector in $\mathbb{R}^m$ by $e_{m,i}$, $\forall m \in \mathbb{N}$, $\forall i \in \{1, \dots, m\}$.

Proposition 12.81 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be an open set, and $F : \Omega \rightarrow \mathcal{Y}$ be continuous on the right (or continuous on the left) and $\mathcal{B}_B(\mathbb{R}^m)$ -measurable. Then, the following statements hold.*

- (i) $\forall i \in \{1, \dots, m\}$, let $D_i F : U_i \rightarrow B(\mathbb{K}, \mathcal{Y})$ denote the partial derivative of F with respect to its i th coordinate in its domain and $U_i := \text{dom}(D_i F)$. Then, $U_i \in \mathcal{B}_B(\mathbb{R}^m)$ and $D_i F$ is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable.
- (ii) Let $F^{(1)} : U \rightarrow B(\mathbb{K}^m, \mathcal{Y})$ be the Fréchet derivative of F , whose domain is $U := \text{dom}(F^{(1)}) \subseteq \bigcap_{i=1}^m U_i \subseteq \Omega$. Then, $U \in \mathcal{B}_B(\mathbb{R}^m)$ and $F^{(1)}$ is $\mathcal{B}_B(\mathbb{R}^m)$ -measurable.

Proof (i) Fix any $i \in \{1, \dots, m\}$. Consider the case that F is continuous on the right. Let $\mathcal{X} := \overline{\mathbb{K}, D := \pi_i(\Omega)} \in \mathcal{B}_B(\mathbb{R})$, which is an open set. $\forall s \in D$, clearly $\text{span}(A_D(s)) = \mathbb{K} = \mathcal{X}$. Then, $D_i F : U_i \rightarrow B(\mathbb{K}, \mathcal{Y})$, where $U_i := \text{dom}(D_i F) \subseteq \Omega$. Clearly, $B(\mathbb{K}, \mathcal{Y})$ is isometrically isomorphic to $\mathcal{Y}$. Then, $D_i F : U_i \rightarrow \mathcal{Y}$. Clearly, $\forall x \in U_i$, $D_i F(x) = \lim_{h \rightarrow 0} (F(x + h e_{m,i}) - F(x))/h \in \mathcal{Y}$.

Claim 12.81.1 $U_i = \bigcap_{n=1}^{\infty} \bigcup_{\substack{\delta \in \mathbb{Q} \\ \delta > 0}} \bigcap_{\substack{h_1 \in \mathbb{Q} \\ 0 < |h_1| < \delta}} \bigcap_{\substack{h_2 \in \mathbb{Q} \\ 0 < |h_2| < \delta}} U_{i,n,h_1,h_2} =: \bar{U}_i \in \mathcal{B}_B(\mathbb{R}^m),$

where $U_{i,n,h_1,h_2} := \{x \in \Omega \mid \|(F(x + h_1 e_{m,i}) - F(x))/h_1 - (F(x + h_2 e_{m,i}) - F(x))/h_2\| < 1/n\}$, $\forall n \in \mathbb{N}$, $\forall h_1, h_2 \in \mathbb{Q}$ with $h_1 \neq 0$ and $h_2 \neq 0$.

Proof of claim: $\forall x \in U_i$, $D_i F(x) = \lim_{h \rightarrow 0} (F(x + h e_{m,i}) - F(x))/h \in \mathcal{Y}$. By Proposition 4.45, $\forall n \in \mathbb{N}$, $\exists \delta \in \mathbb{Q}$ with $\delta > 0$, such that $x + h e_{m,i} \in \Omega$, $\forall h \in \mathcal{B}_{\mathbb{R}}(0, \delta)$, and, $\forall \bar{h}_1, \bar{h}_2 \in \mathcal{B}_{\mathbb{R}}(0, \delta) \setminus \{0\}$, we have $\|(F(x + \bar{h}_1 e_{m,i}) -$

$F(x))/\bar{h}_1 - (F(x + \bar{h}_2 \mathbf{e}_{m,i}) - F(x))/\bar{h}_2\| < 1/n$. Then, $\forall h_1, h_2 \in \mathbb{Q}$ with $0 < |h_1| < \delta$ and $0 < |h_2| < \delta$, we have $\|(F(x + h_1 \mathbf{e}_{m,i}) - F(x))/h_1 - (F(x + h_2 \mathbf{e}_{m,i}) - F(x))/h_2\| < 1/n$. Thus, $x \in U_{i,n,h_1,h_2}$. Hence, $x \in \bar{U}_i$. By the arbitrariness of x , we have $U_i \subseteq \bar{U}_i$.

On the other hand, $\forall x \in \bar{U}_i$. $\forall n \in \mathbb{N}$, $\exists \delta \in \mathbb{Q}$ with $\delta > 0$, such that, $\forall h \in \mathcal{B}_{\mathbf{R}}(0, \delta)$, $x + h \mathbf{e}_{m,i} \in \Omega$, and $\forall h_1, h_2 \in \mathbb{Q}$ with $0 < |h_1| < \delta$ and $0 < |h_2| < \delta$, we have $\|(F(x + h_1 \mathbf{e}_{m,i}) - F(x))/h_1 - (F(x + h_2 \mathbf{e}_{m,i}) - F(x))/h_2\| < 1/n$. $\forall \bar{h}_1, \bar{h}_2 \in \mathcal{B}_{\mathbf{R}}(0, \delta) \setminus \{0\}$, $\exists (h_{j,l})_{l=1}^{\infty} \subseteq (\mathbb{Q} \cap \mathcal{B}_{\mathbf{R}}(0, \delta)) \setminus \{0\}$ such that $\bar{h}_j < h_{j,l}$, $\forall l \in \mathbb{N}$, and $\lim_{l \in \mathbb{N}} h_{j,l} = \bar{h}_j$, $j = 1, 2$. $\forall l \in \mathbb{N}$, $\|(F(x + h_{1,l} \mathbf{e}_{m,i}) - F(x))/h_{1,l} - (F(x + h_{2,l} \mathbf{e}_{m,i}) - F(x))/h_{2,l}\| < 1/n$. Then,

$$\begin{aligned} & \|(F(x + \bar{h}_1 \mathbf{e}_{m,i}) - F(x))/\bar{h}_1 - (F(x + \bar{h}_2 \mathbf{e}_{m,i}) - F(x))/\bar{h}_2\| \\ &= \left\| \lim_{l \in \mathbb{N}} \left(\frac{F(x + h_{1,l} \mathbf{e}_{m,i}) - F(x)}{h_{1,l}} - \frac{F(x + h_{2,l} \mathbf{e}_{m,i}) - F(x)}{h_{2,l}} \right) \right\| \\ &= \lim_{l \in \mathbb{N}} \|(F(x + h_{1,l} \mathbf{e}_{m,i}) - F(x))/h_{1,l} - (F(x + h_{2,l} \mathbf{e}_{m,i}) - F(x))/h_{2,l}\| \\ &\leq 1/n \end{aligned}$$

where the first equality follows from the fact that F is continuous on the right; and the second equality follows from Propositions 3.66, 3.67, and 7.23. By Proposition 4.45, $\lim_{h \rightarrow 0} (F(x + h \mathbf{e}_{m,i}) - F(x))/h \in \mathcal{Y}$. Hence, $x \in \text{dom}(\mathbf{D}_i F) = U_i$. By the arbitrariness of x , we have $\bar{U}_i \subseteq U_i$. Therefore, $U_i = \bar{U}_i$.

By Propositions 7.21, 7.23, 11.38, and 11.39, $U_{i,n,h_1,h_2} \in \mathcal{B}_{\mathbf{B}}(\mathbb{R}^m)$, $\forall n \in \mathbb{N}$, $\forall h_1, h_2 \in \mathbb{Q}$ with $h_1 \neq 0$ and $h_2 \neq 0$. Then, $U_i = \bar{U}_i \in \mathcal{B}_{\mathbf{B}}(\mathbb{R}^m)$. $\square$

$\forall n \in \mathbb{N}$, define $H_n : \Omega \rightarrow \mathcal{Y}$ by, $\forall x \in \Omega$,

$$H_n(x) = \begin{cases} n(F(x + \frac{1}{n} \mathbf{e}_{m,i}) - F(x)) & x \in O := \Omega \cap (\Omega - \frac{1}{n} \mathbf{e}_{m,i}) \\ \vartheta_{\mathcal{Y}} & x \in \Omega \setminus O \end{cases}$$

By Propositions 7.23, 11.38, 11.39, and 11.41, H_n is $\mathcal{B}_{\mathbf{B}}(\mathbb{R}^m)$ -measurable. Then, $H_n|_{U_i}$ is $\mathcal{B}_{\mathbf{B}}(\mathbb{R}^m)$ -measurable. Note that, $\forall x \in U_i$, $\mathbf{D}_i F(x) = \lim_{n \in \mathbb{N}} H_n(x) = \lim_{n \in \mathbb{N}} H_n|_{U_i}(x)$. By Proposition 11.48, $\mathbf{D}_i F$ is $\mathcal{B}_{\mathbf{B}}(\mathbb{R}^m)$ -measurable.

The case when F is continuous on the left can be proved similarly.

(ii) Consider the case that F is continuous on the right. By Proposition 9.9, $\forall x \in U = \text{dom}(F^{(1)})$, we must have $x \in U_i$, $\forall i \in \{1, \dots, m\}$. Then, $U \subseteq \bigcap_{i=1}^m U_i$. $\forall x \in \Omega$, since Ω is open, then $\text{span}(\mathbf{A}_{\Omega}(x)) = \mathbb{K}^m$.

Claim 12.81.2 $U = \bigcap_{n=1}^{\infty} \bigcup_{\substack{\delta \in \mathbb{Q} \\ \delta > 0}} \bigcap_{\substack{h_1 \in \mathbb{Q} \\ |h_1| < \delta}} \cdots \bigcap_{\substack{h_m \in \mathbb{Q} \\ |h_m| < \delta}} V_{n,h_1,\dots,h_m} =: \bar{V} \in \mathcal{B}_{\mathbf{B}}(\mathbb{R}^m)$,

where $V_{n,h_1,\dots,h_m} := \left\{ x \in \bigcap_{i=1}^m U_i \mid \|F(x + \sum_{i=1}^m h_i \mathbf{e}_{m,i}) - F(x) - \sum_{i=1}^m h_i \mathbf{D}_i F(x)\| \leq \frac{1}{n} \sqrt{\sum_{i=1}^m h_i^2} \right\}$, $\forall n \in \mathbb{N}$, $\forall h_1, \dots, h_m \in \mathbb{Q}$.

Proof of claim: $\forall x \in U$, we have $x \in \bigcap_{i=1}^m U_i$. By Definition 9.3 and Proposition 9.9, the candidate for $F^{(1)}(x)$ is $\begin{bmatrix} D_1 F(x) & \cdots & D_m F(x) \end{bmatrix} := L \in \mathcal{B}(\mathbb{K}^m, \mathcal{Y})$. By $x \in \text{dom}(F^{(1)})$, we have $\forall n \in \mathbb{N}$, $\exists \delta \in \mathbb{Q}$ with $\delta > 0$, such that $\mathcal{B}_{\mathbb{R}^m}(x, \sqrt{m}\delta) \subseteq \Omega$, and $\forall \bar{h} \in \mathcal{B}_{\mathbb{R}^m}(\mathbf{0}_m, \sqrt{m}\delta)$, we have $\|F(x + \bar{h}) - F(x) - L\bar{h}\| \leq \frac{1}{n}|\bar{h}|$. Then, $\forall h := (h_1, \dots, h_m) \in \mathbb{Q}^m$ with $|h_i| < \delta$, $i = 1, \dots, m$, we have $\|F(x + h) - F(x) - Lh\| = \|F(x + \sum_{i=1}^m h_i e_{m,i}) - F(x) - \sum_{i=1}^m h_i D_i F(x)\| \leq \frac{1}{n}|h| = \frac{1}{n}\sqrt{\sum_{i=1}^m h_i^2}$. Thus, $x \in V_{n, h_1, \dots, h_m}$. Hence, $x \in \bar{V}$. By the arbitrariness of x , we have $U \subseteq \bar{V}$.

On the other hand, $\forall x \in \bar{V}$. Then, $\forall n \in \mathbb{N}$, $\exists \delta \in \mathbb{Q}$ with $\delta > 0$, such that $\mathcal{B}_{\mathbb{R}^m}(x, \delta) \subseteq \Omega$ and, $\forall h = (h_1, \dots, h_m) \in \mathbb{Q}^m$ with $|h_i| < \delta$, $i = 1, \dots, m$, we have $x \in V_{n, h_1, \dots, h_m} \subseteq \bigcap_{i=1}^m U_i$. $\forall \bar{x} \in \mathcal{B}_{\mathbb{R}^m}(x, \delta) \cap \Omega = \mathcal{B}_{\mathbb{R}^m}(x, \delta)$, let $(\bar{h}_1, \dots, \bar{h}_m) := \bar{x} - x$. Clearly, $|\bar{h}_i| < \delta$, $i = 1, \dots, m$. Then, $\exists (h_{i,l})_{l=1}^\infty \subseteq \mathbb{Q}$ with $|h_{i,l}| < \delta$ such that $\bar{h}_i < h_{i,l}$, $\forall l \in \mathbb{N}$, and $\lim_{l \in \mathbb{N}} h_{i,l} = \bar{h}_i$, $i = 1, \dots, m$. By the assumption,

$$\begin{aligned} & \|F(\bar{x}) - F(x) - L(\bar{x} - x)\| \\ &= \left\| \lim_{l \in \mathbb{N}} \left(F(x + \sum_{i=1}^m h_{i,l} e_{m,i}) - F(x) - \sum_{i=1}^m h_{i,l} D_i F(x) \right) \right\| \\ &= \lim_{l \in \mathbb{N}} \left\| F(x + \sum_{i=1}^m h_{i,l} e_{m,i}) - F(x) - \sum_{i=1}^m h_{i,l} D_i F(x) \right\| \\ &\leq \lim_{l \in \mathbb{N}} \frac{1}{n} \left(\sum_{i=1}^m h_{i,l}^2 \right)^{1/2} = \frac{1}{n} \left(\sum_{i=1}^m \bar{h}_i^2 \right)^{1/2} = \frac{1}{n} |\bar{x} - x| \end{aligned}$$

where the first equality follows from the fact that F is continuous on the right; the second equality follows from Propositions 3.66, 3.67, 7.21, and 7.23; and the first inequality follows from the fact that $x \in V_{n, h_1, l, \dots, h_m, l}$. By Definition 9.3, $F^{(1)}(x) = L$. Hence, $x \in \text{dom}(F^{(1)}) = U$. By the arbitrariness of x , we have $\bar{V} \subseteq U$. Therefore, $U = \bar{V}$.

By Propositions 7.21, 7.23, 11.38, and 11.39, $V_{n, h_1, \dots, h_m} \in \mathcal{B}_{\mathbb{B}}(\mathbb{R}^m)$, $\forall n \in \mathbb{N}$, $\forall h_1, \dots, h_m \in \mathbb{Q}$. Then, $U = \bar{V} \in \mathcal{B}_{\mathbb{B}}(\mathbb{R}^m)$. $\square$

Note that, $\forall x \in U$, $F^{(1)}(x) = \begin{bmatrix} D_1 F(x) & \cdots & D_m F(x) \end{bmatrix} = \begin{bmatrix} (D_1 F)|_U(x) & \cdots & (D_m F)|_U(x) \end{bmatrix}$. By Propositions 11.38, 11.39, 11.41, and 11.139, $F^{(1)}$ is $\mathcal{B}_{\mathbb{B}}(\mathbb{R}^m)$ -measurable.

The case when F is continuous on the left can be proved similarly.

This completes the proof of the proposition. $\square$

Theorem 12.82 *Let $I := [a, b] \subset \mathbb{R}$ with $a, b \in \mathbb{R}$ and $a < b$, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, $f : I \rightarrow \mathcal{Y}$ be absolutely integrable over I with respect to $\mu_{\mathbb{B}}$, and $F : I \rightarrow \mathcal{Y}$ be defined by $F(x) = \int_a^x f(t) dt$, $\forall x \in I$. Assume that f is continuous at $x_0 \in I$. Then, F is Fréchet differentiable at x_0 and $F^{(1)}(x_0) = f(x_0)$. (Note that, when $\mathbb{K} = \mathbb{C}$, I is viewed as a subset of $\mathbb{C}$ in calculations of $F^{(1)}$.)*

Proof Clearly, F is well-defined by Proposition 11.92. Note that $\text{span}(\mathcal{A}_I(x_0)) = \mathbb{K}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the continuity of f at x_0 , $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x \in I \cap \mathcal{B}_{\mathbb{R}}(x_0, \delta)$, we have $\|f(x) - f(x_0)\| < \epsilon$. Then,

$$\begin{aligned}
& \|F(x) - F(x_0) - f(x_0)(x - x_0)\| \\
&= \left\| \int_a^x f(t) dt - \int_a^{x_0} f(t) dt - \int_{x_0}^x f(x_0) dt \right\| \\
&= \left\| \int_{x_0}^x f(t) dt - \int_{x_0}^x f(x_0) dt \right\| = \left\| \int_{x_0}^x (f(t) - f(x_0)) dt \right\| \\
&= \left\| \int_{\mathbb{R}_{x_0 \wedge x, x_0 \vee x}} (f - f(x_0)) d\mu_B \right\| \leq \int_{\mathbb{R}_{x_0 \wedge x, x_0 \vee x}} \mathcal{P} \circ (f - f(x_0)) d\mu_B \\
&\leq \epsilon \mu_B(\mathbb{R}_{x_0 \wedge x, x_0 \vee x}) = \epsilon |x - x_0|
\end{aligned}$$

where the second equality follows from Fact 12.72; the third equality and the first inequality follow from Proposition 11.92; and the second inequality follows from Proposition 11.78. Hence, $F^{(1)}(x_0) = f(x_0)$. This completes the proof of the theorem. $\square$

Theorem 12.83 Let $I := [a, b] \subset \mathbb{R}$ with $a, b \in \mathbb{R}$ and $a < b$, $\mathbb{I} := ((I, |\cdot|), \mathcal{B}, \mu)$ be the finite complete metric measure subspace of $\mathbb{R}$, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, and $F : I \rightarrow \mathcal{Y}$ be $\mathcal{C}_1$. (Note that, when $\mathbb{K} = \mathbb{C}$, I is viewed as a subset of $\mathbb{C}$ in calculations of $F^{(1)}$.) Then, $F^{(1)} : I \rightarrow \mathcal{Y}$ is absolutely integrable over $\mathbb{I}$ and $F(b) - F(a) = \int_a^b F^{(1)}(t) dt = \int_a^b F^{(1)} d\mu_B$.

Proof Since $F^{(1)} : I \rightarrow \mathcal{Y}$ is continuous, then, by Proposition 11.37, $F^{(1)}$ is $\mathcal{B}$ -measurable. Since I is compact, then $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|F^{(1)}(x)\| \leq M$, $\forall x \in I$. By Proposition 11.78, we have $F^{(1)}$ is absolutely integrable over $\mathbb{I}$. Define $F_a : I \rightarrow \mathcal{Y}$ by $F_a(x) = \int_a^x F^{(1)}(t) dt = \int_a^x F^{(1)} d\mu_B$, $\forall x \in I$. By Proposition 11.92, F_a is well defined. By Theorem 12.82, F_a is Fréchet differentiable and $F_a^{(1)} = F^{(1)}$. Then, define $g : I \rightarrow \mathcal{Y}$ by $g = F - F_a$. Then, by Proposition 9.15, g is Fréchet differentiable and $g^{(1)}(x) = \vartheta_{\mathcal{Y}}$, $\forall x \in I$. By Mean Value Theorem 9.23, $\|g(x) - g(a)\| \leq 0$, $\forall x \in I$. Hence, $g(x) = g(a) = F(a)$, $\forall x \in I$. Then, $g(b) = F(a) = F(b) - F_a(b) = F(b) - \int_a^b F^{(1)}(t) dt$. Therefore, $F(b) - F(a) = \int_a^b F^{(1)}(t) dt$. This completes the proof of the theorem. $\square$

Proposition 12.84 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a region, $\mathbb{X} := ((P(\Omega), |\cdot|), \mathcal{B}, \mu)$ be the σ -finite metric measure subspace of $\mathbb{R}^m$, $\mathcal{Y}$ be a finite-dimensional Banach space over $\mathbb{K}$, $f : P(\Omega) \rightarrow \mathcal{Y}$ be $\mathcal{B}_B(\mathbb{R}^m)$ -measurable, $\bar{J} := \{1, \dots, m\}$, and $x_0 := (x_{0,1}, \dots, x_{0,m}) \in \Omega$. Then, we may define the σ -finite $\mathcal{Y}$ -valued measure ν with kernel f over $\mathbb{X}$ according to Proposition 11.116. By Proposition 11.167, f is the unique Radon-Nikodym derivative of ν with respect

to μ . Assume that, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, we have f is bounded over $r_{x_1, x_2} \subseteq P(\Omega)$. Let $F : \Omega \rightarrow \mathcal{Y}$ be a cumulative distribution function of ν . Then, $\forall x := (x_1, \dots, x_m) \in \Omega$ with $\bar{r}_{\hat{x}, \check{x}} \subseteq \Omega$, where $\hat{x} := x_0 \wedge x$ and $\check{x} := x_0 \vee x$, we have

$$F(x) = \int_{x_{0,1}}^{x_1} \cdots \int_{x_{0,m}}^{x_m} f(s_1, \dots, s_m) ds_m \cdots ds_1 - \sum_{J \subseteq \bar{J}, J \neq \emptyset} (-1)^{\text{card}(J)} F(x_J) \quad (12.1)$$

and the order of integration can be arbitrary, where $x_J \in \bar{r}_{\hat{x}, \check{x}} \subseteq \Omega$ is defined

$$\text{by } \pi_i(x_J) = \begin{cases} \pi_i(x_0) = x_{0,i} & i \in J \\ \pi_i(x) = x_i & i \in \bar{J} \setminus J \end{cases}, \forall i \in \bar{J}.$$

Proof $\mathcal{Y}$ is separable and is σ -compact, since it is finite-dimensional. Fix any $x \in \Omega$ with $\bar{r}_{\hat{x}, \check{x}} \subseteq \Omega$. Then, $r_{\hat{x}, \check{x}} \subseteq P(\Omega)$. By Definition 12.42 and Definition 11.166, we have $\Delta_F(r_{\hat{x}, \check{x}}) = \nu(r_{\hat{x}, \check{x}}) = \int_{r_{\hat{x}, \check{x}}} f d\mu = \int_{r_{\hat{x}, \check{x}}} f d\mu_{Bm}$. By the assumption and Fubini's Theorem 12.31, we have

$$\begin{aligned} \int_{r_{\hat{x}, \check{x}}} f d\mu_{Bm} &= \int_{x_{0,1} \wedge x_1}^{x_{0,1} \vee x_1} \cdots \int_{x_{0,m} \wedge x_m}^{x_{0,m} \vee x_m} f(s_1, \dots, s_m) ds_m \cdots ds_1 \\ &= (-1)^{n(x)} \int_{x_{0,1}}^{x_1} \cdots \int_{x_{0,m}}^{x_m} f(s_1, \dots, s_m) ds_m \cdots ds_1 = \Delta_F(r_{\hat{x}, \check{x}}) \end{aligned}$$

and the order of integration can be arbitrary, where $n(x) = \text{card}(\{i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x)\}) =: \text{card}(\bar{J})$. By Definition 12.41,

$$\begin{aligned} \Delta_F(r_{\hat{x}, \check{x}}) &= \sum_{\bar{J} \subseteq \bar{J}} (-1)^{\text{card}(\bar{J})} F(\bar{x}_{\bar{J}}) \\ &= (-1)^{n(x)} \int_{x_{0,1}}^{x_1} \cdots \int_{x_{0,m}}^{x_m} f(s_1, \dots, s_m) ds_m \cdots ds_1 \end{aligned}$$

where $\bar{x}_{\bar{J}} \in \bar{r}_{\hat{x}, \check{x}} \subseteq \mathbb{R}^m$ is defined by, $\forall i \in \bar{J}$,

$$\begin{aligned} \pi_i(\bar{x}_{\bar{J}}) &= \begin{cases} \pi_i(\hat{x}) & \forall i \in \hat{J} \\ \pi_i(\check{x}) & \forall i \in \bar{J} \setminus \hat{J} \end{cases} \\ &= \begin{cases} \pi_i(x) & \forall i \in (\hat{J} \cap \bar{J}) \cup ((\bar{J} \setminus \hat{J}) \cap (\bar{J} \setminus \check{J})) \\ \pi_i(x_0) & \text{otherwise} \end{cases} \\ &= \begin{cases} \pi_i(x) & \forall i \in \bar{J} \setminus (\hat{J} \triangle \check{J}) \\ \pi_i(x_0) & \forall i \in \hat{J} \triangle \check{J} \end{cases} = \pi_i(x_{\hat{J} \triangle \check{J}}) \end{aligned}$$

This leads to

$$\Delta_F(r_{\hat{x}, \check{x}}) = \sum_{\bar{J} \subseteq \bar{J}} (-1)^{\text{card}(\bar{J})} F(x_{\hat{J} \triangle \check{J}})$$

$$\begin{aligned}
&= (-1)^{\text{card}(\bar{J})} \sum_{\hat{J} \subseteq \bar{J}} (-1)^{\text{card}(\hat{J}) + \text{card}(\bar{J})} F(x_{\hat{J} \Delta \bar{J}}) \\
&= (-1)^{n(x)} \sum_{\hat{J} \subseteq \bar{J}} (-1)^{\text{card}(\hat{J} \Delta \bar{J})} F(x_{\hat{J} \Delta \bar{J}}) \\
&= (-1)^{n(x)} \int_{x_{0,1}}^{x_1} \cdots \int_{x_{0,m}}^{x_m} f(s_1, \dots, s_m) ds_m \cdots ds_1
\end{aligned}$$

Then, we can conclude

$$\sum_{J \subseteq \bar{J}} (-1)^{\text{card}(J)} F(x_J) = \int_{x_{0,1}}^{x_1} \cdots \int_{x_{0,m}}^{x_m} f(s_1, \dots, s_m) ds_m \cdots ds_1$$

The above equation is equivalent to (12.1). This completes the proof of the proposition. $\square$

Proposition 12.85 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\Omega)$ be an open rectangle, $\mathbb{X} := ((\Omega, |\cdot|), \mathcal{B}, \mu)$ be the σ -finite metric measure subspace of $\mathbb{R}^m$, $\mathcal{Y}$ be a separable Banach space, $f : \Omega \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $x_0 \in \Omega$, and $\bar{J} := \{1, \dots, m\}$. Assume that f is absolutely integrable over r_{x_a, x_b} with respect to μ , $\forall x_a, x_b \in \Omega$ with $x_a \leq x_b$. Define $\pi_J : \mathbb{R}^m \rightarrow \mathbb{R}^{\text{card}(J)}$ by $\pi_J(x) = (\pi_i(x))_{i \in J}$, $\forall x \in \mathbb{R}^m$, $\forall J \subseteq \bar{J}$; $(\pi_J(\Omega), \mathcal{B}_J, \mu_J)$ be the measure subspace of $\mathbb{R}^{\text{card}(J)}$, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$; $M_J : \mathbb{R}^{\text{card}(J)} \times \mathbb{R}^{m-\text{card}(J)} \rightarrow \mathbb{R}^m$ by $M_J(s, t) \in \mathbb{R}^m$ is such that $\pi_J(M_J(s, t)) = s$ and $\pi_{\bar{J} \setminus J}(M_J(s, t)) = t$, $\forall s \in \mathbb{R}^{\text{card}(J)}$, $\forall t \in \mathbb{R}^{m-\text{card}(J)}$, $\forall J \subseteq \bar{J}$. Then, the following statements hold.*

- (i) $\forall \emptyset \neq J \subseteq \bar{J}$, $\exists U_J \in \mathcal{B}$ with $\mu(\Omega \setminus U_J) = 0$ such that $f(M_J(\cdot, \pi_{\bar{J} \setminus J}(x))) : \pi_J(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $\pi_J(r_{\hat{x}, \check{x}})$ with respect to μ_J , $\forall x \in U_J$; and $F : \Omega \rightarrow \mathcal{Y}$, defined by, $\forall x \in \Omega$, $F(x) =$
- $$\begin{cases} (-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(r_{\hat{x}, \check{x}})} f(M_J(s, \pi_{\bar{J} \setminus J}(x))) d\mu_J(s) & x \in U_J \\ \vartheta_{\mathcal{Y}} & x \in \Omega \setminus U_J \end{cases},$$
- where $\hat{x} := x_0 \wedge x$, $\check{x} := x_0 \vee x$, and $\check{J} := \{i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x)\}$, is $\mathcal{B}$ -measurable.

- (ii) Furthermore, if, in addition, $\mathcal{Y}$ is finite dimensional then $U_J = \{x \in \Omega \mid f(M_J(\cdot, \pi_{\bar{J} \setminus J}(x))) : \pi_J(\Omega) \rightarrow \mathcal{Y} \text{ is absolutely integrable over } \pi_J(r_{\hat{x}, \check{x}}) \text{ with respect to } \mu_J\}$.

Proof First consider the special case when $\mu(\Omega) < \infty$ and f is absolutely integrable over $\mathbb{X}$. Since Ω is a nonempty open rectangle, then Ω is bounded. Define $\tilde{f} : \pi_J(\Omega) \times \Omega \rightarrow \mathcal{Y}$ by $\tilde{f}(s, x) = (-1)^{\text{card}(\bar{J} \cap J)} f(M_J(s, \pi_{\bar{J} \setminus J}(x))) \chi_{\pi_J(r_{\hat{x}, \check{x}}), \pi_J(\Omega)}(s)$, $\forall (s, x) \in \pi_J(\Omega) \times \Omega$. By Propositions 7.23, 11.39, and 11.38, $\tilde{f}$ is $\mathcal{B}_B(\mathbb{R}^{\text{card}(J)+m})$ -measurable. Note

the following derivation,

$$\begin{aligned}
& \int_{\pi_J(\Omega) \times \Omega} \mathcal{P} \circ \bar{f}(s, x) d(\mu_J \times \mu)(s, x) \\
& \leq \int_{\pi_J(\Omega)} \int_{\Omega} \|f(M_J(s, \pi_{\bar{J} \setminus J}(x)))\| d\mu(x) d\mu_J(s) \\
& = \int_{\pi_J(\Omega)} \int_{\pi_{\bar{J} \setminus J}(\Omega)} \int_{\pi_J(\Omega)} \|f(M_J(s, t))\| d\mu_J(\tau) d\mu_{\bar{J} \setminus J}(t) d\mu_J(s) \\
& \leq \mu_J(\pi_J(\Omega)) \int_{\pi_J(\Omega)} \int_{\pi_{\bar{J} \setminus J}(\Omega)} \|f(x)\| d\mu_{\bar{J} \setminus J}(\pi_{\bar{J} \setminus J}(x)) d\mu_J(\pi_J(x)) \\
& = \mu_J(\pi_J(\Omega)) \int_{\Omega} \mathcal{P} \circ f d\mu < \infty
\end{aligned}$$

where the first inequality and the equalities follow from Tonelli's Theorem 12.29; and the second inequality follows from Proposition 11.83. Then, we have that $\bar{f}$ is absolutely integrable over $\pi_J(\Omega) \times \Omega$ with respect to $\mu_J \times \mu$. By Fubini's Theorem 12.30, $\exists p_J : \Omega \rightarrow \mathcal{Y}$ and $U_J \in \mathcal{B}$ such that $\mu(\Omega \setminus U_J) = 0$; $\bar{f}(\cdot, x) : \pi_J(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $(\pi_J(\Omega), \mathcal{B}_J, \mu_J)$, $\forall x \in U_J$; $p_J : \Omega \rightarrow \mathcal{Y}$, defined by, $\forall x \in \Omega$, $p_J(x) = \begin{cases} \int_{\pi_J(\Omega)} \bar{f}(s, x) d\mu_J(s) & x \in U_J \\ \vartheta_{\mathcal{Y}} & x \in \Omega \setminus U_J \end{cases}$, is $\mathcal{B}$ -measurable; p_J is absolutely integrable over $\mathbb{X}$, and $\int_{\pi_J(\Omega) \times \Omega} \bar{f}(s, x) d(\mu_J \times \mu)(s, x) = \int_{\Omega} p_J(x) d\mu(x)$. Clearly, $F(x) = p_J(x)$, $\forall x \in \Omega$. Hence, F is $\mathcal{B}$ -measurable. Then, $f(M_J(\cdot, \pi_{\bar{J} \setminus J}(x))) : \pi_J(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $\pi_J(r_{\hat{x}, \tilde{x}}) \subseteq \pi_J(\Omega)$ with respect to μ_J , $\forall x \in U_J$.

If, in addition, $\mathcal{Y}$ is finite dimensional, let U_J be defined as in the statement (ii). $\forall x \in \Omega$, $f(M_J(\cdot, \pi_{\bar{J} \setminus J}(x))) : \pi_J(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $\pi_J(r_{\hat{x}, \tilde{x}})$ with respect to μ_J if, and only if, $\bar{f}(\cdot, x)$ is absolutely integrable over $(\pi_J(\Omega), \mathcal{B}_J, \mu_J)$. Then, $U_J := \{x \in \Omega \mid \bar{f}(\cdot, x) : \pi_J(\Omega) \rightarrow \mathcal{Y} \text{ is absolutely integrable over } (\pi_J(\Omega), \mathcal{B}_J, \mu_J)\}$. By Fubini's Theorem 12.31, this choice of U_J is admissible for the definition of p_J and p_J is $\mathcal{B}$ -measurable. This proves the result in this special case.

Now consider the general case. Since Ω is an open rectangle and $x_0 \in \Omega$, then there exists $(\hat{x}_j)_{j=1}^{\infty}, (\tilde{x}_j)_{j=1}^{\infty} \subseteq \Omega$ such that $\dots \leq \hat{x}_j \leq \dots \leq \hat{x}_1 \leq x_0 < \tilde{x}_1 < \dots < \tilde{x}_j < \dots$, and $\Omega = \bigcup_{j=1}^{\infty} E_j := \bigcup_{j=1}^{\infty} r_{\hat{x}_j, \tilde{x}_j}^{\circ}$. Fix any $j \in \mathbb{N}$. By the assumption of the proposition, we can apply the result of the special case on E_j . There exists $U_{J,j} \in \mathcal{B}$ with $\mu(E_j \setminus U_{J,j}) = 0$ such that $f(M_J(\cdot, \pi_{\bar{J} \setminus J}(x))) : \pi_J(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $\pi_J(r_{\hat{x}, \tilde{x}})$ with respect to μ_J , $\forall x \in U_{J,j}$; and $F_j : E_j \rightarrow \mathcal{Y}$, defined by, $\forall x \in E_j$,

$$F_j(x) = \begin{cases} (-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(r_{\hat{x}, \tilde{x}})} f(M_J(s, \pi_{\bar{J} \setminus J}(x))) d\mu_J(s) & x \in U_{J,j} \\ \vartheta_{\mathcal{Y}} & x \in E_j \setminus U_{J,j} \end{cases}$$

is $\mathcal{B}$ -measurable. Let $U_J := \bigcup_{j=1}^{\infty} U_{J,j} \in \mathcal{B}$. Then, $0 \leq \mu(\Omega \setminus U_J) = \mu(\bigcup_{j=1}^{\infty} (E_j \setminus U_J)) \leq \mu(\bigcup_{j=1}^{\infty} (E_j \setminus U_{J,j})) \leq 0$. Define F according to the statement of the proposition, we observe that $F(x) = F_j(x)$, $\forall x \in U_{J,j} \setminus (\bigcup_{l=1}^{j-1} U_{J,l})$, if $x \in U_J$; $F(x) = \vartheta_{\mathcal{Y}}$, $\forall x \in \Omega \setminus U_J$. By Proposition 11.41, F is $\mathcal{B}$ -measurable. Clearly, $f(M_J(\cdot, \pi_{J \setminus J}(x))) : \pi_J(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $\pi_J(r_{\hat{x}, \tilde{x}}) \subseteq \pi_J(\Omega)$ with respect to μ_J , $\forall x \in U_J$.

If, in addition, $\mathcal{Y}$ is finite dimensional, then, by the special case, $U_{J,j} = \{x \in E_j \mid f(M_J(\cdot, \pi_{J \setminus J}(x))) : \pi_J(\Omega) \rightarrow \mathcal{Y} \text{ is absolutely integrable over } \pi_J(r_{\hat{x}, \tilde{x}}) \text{ with respect to } \mu_J\}$, $\forall j \in \mathbb{N}$. Then, $U_J = \bigcup_{j=1}^{\infty} U_{J,j} = \{x \in \Omega \mid f(M_J(\cdot, \pi_{J \setminus J}(x))) : \pi_J(\Omega) \rightarrow \mathcal{Y} \text{ is absolutely integrable over } \pi_J(r_{\hat{x}, \tilde{x}}) \text{ with respect to } \mu_J\}$. This completes the proof of the proposition. $\square$

Theorem 12.86 (Fundamental Theorem of Calculus I) *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be an open rectangle, $\mathbb{X} := ((\Omega, |\cdot|), \mathcal{B}, \mu)$ be the σ -finite metric measure subspace of $\mathbb{R}^m$, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, $f : \Omega \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable, $\bar{J} := \{1, \dots, m\}$, and $x_0 \in \Omega$. Assume that f is absolutely integrable over r_{x_a, x_b} with respect to μ , $\forall x_a, x_b \in \Omega$ with $x_a \leq x_b$. Let ν be the σ -finite $\mathcal{Y}$ -valued measure with kernel f over $\mathbb{X}$ according to Proposition 11.116. Define $F : \Omega \rightarrow \mathcal{Y}$ to be the cumulative distribution function of ν with origin at x_0 according to Proposition 12.52. F is absolutely continuous by Proposition 12.75, and $f = \frac{d\nu}{d\mu}$ a.e. in $\mathbb{X}$. Let $D_i F : U_i \rightarrow \mathcal{Y}$, $\forall i \in \bar{J}$, where D_i is the derivative of a function with respect to the i th coordinate variable of Ω . $\forall i \in \bar{J}$, define $\pi_{\bar{i}} : \mathbb{R}^m \rightarrow \mathbb{R}^{m-1}$ by $\pi_{\bar{i}}(x) = (\pi_j(x))_{j \in \bar{J} \setminus \{i\}}$, $\forall x \in \mathbb{R}^m$; also define $M_i : \mathbb{R} \times \mathbb{R}^{m-1} \rightarrow \mathbb{R}^m$ by $\pi_i(M_i(x, y)) = x$ and $\pi_{\bar{i}}(M_i(x, y)) = y$, $\forall x \in \mathbb{R}$, $\forall y \in \mathbb{R}^{m-1}$. Then, the following statements hold.*

$$(i) \quad F(x) = (-1)^{\text{card}(\bar{J})} \int_{r_{\hat{x}, \tilde{x}}} f \, d\mu, \quad \forall x \in \Omega, \text{ where } \hat{x} := x_0 \wedge x, \tilde{x} := x_0 \vee x, \\ \text{and } \bar{J} := \{i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x)\}.$$

$$(ii) \quad U_i \in \mathcal{B}, \mu(\Omega \setminus U_i) = 0, D_i F \text{ is } \mathcal{B}\text{-measurable; } \exists \bar{U}_i \in \mathcal{B} \text{ with } \mu(\Omega \setminus \bar{U}_i) = 0, \forall x \in \bar{U}_i, f(M_i(\pi_i(x), \cdot)) : \pi_{\bar{i}}(\Omega) \rightarrow \mathcal{Y} \text{ is absolutely integrable over } \pi_{\bar{i}}(r_{\hat{x}, \tilde{x}}) \text{ with respect to } \mu_{\bar{i}}, \text{ and } \exists p_i : \Omega \rightarrow \mathcal{Y} \text{ defined by, } \forall x \in \Omega,$$

$$p_i(x) = \begin{cases} (-1)^{\text{card}(\bar{J} \setminus \{i\})} \int_{\pi_{\bar{i}}(r_{\hat{x}, \tilde{x}})} f(M_i(\pi_i(x), s)) \, d\mu_{\bar{i}}(s) & x \in \bar{U}_i \\ \vartheta_{\mathcal{Y}} & x \in \Omega \setminus \bar{U}_i \end{cases}$$

when $m > 1$, $p_i(x) = f(x)$, $\forall x \in \Omega$, $\bar{U}_i = \emptyset$, when $m = 1$; where $(\pi_{\bar{i}}(\Omega), \mathcal{B}_{\bar{i}}, \mu_{\bar{i}})$ is the measure subspace of $\mathbb{R}^{m-1}$; such that p_i is $\mathcal{B}$ -measurable, and $D_i F(x) = p_i(x)$ a.e. in $\mu(x)$, $\forall i \in \bar{J}$.

$$(iii) \quad \text{If, in addition, } \mathcal{Y} \text{ is finite dimensional, then } \bar{U}_i = \{x \in \Omega \mid f(M_i(\pi_i(x), \cdot)) : \pi_{\bar{i}}(\Omega) \rightarrow \mathcal{Y} \text{ is absolutely integrable over } \pi_{\bar{i}}(r_{\hat{x}, \tilde{x}}) \text{ with respect to } \mu_{\bar{i}}\}.$$

Proof By Proposition 12.75, F is absolutely continuous. By Definition 12.41 and Definition 11.166, we have $\Delta_F(r_{\tilde{x}, \tilde{x}}) = \nu(r_{\tilde{x}, \tilde{x}}) = \int_{r_{\tilde{x}, \tilde{x}}} f d\mu$. Since F is the cumulative distribution function of ν with origin at x_0 , then $\Delta_F(r_{\tilde{x}, \tilde{x}}) = (-1)^{\text{card}(\bar{J})} F(x)$. Thus, we have $F(x) = (-1)^{\text{card}(\bar{J})} \int_{r_{\tilde{x}, \tilde{x}}} f d\mu$, $\forall x \in \Omega$. This establishes (i).

(ii) By Proposition 12.81, $D_i F$ is $\mathcal{B}$ -measurable, $\forall i \in \bar{J}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $m = 1$; Case 2: $m > 1$. Case 1: $m = 1$. Then, $i = 1$. $F(x) = (-1)^{\text{card}(\bar{J})} \int_{r_{\tilde{x}, \tilde{x}}} f d\mu = \int_{x_0}^x f d\mu_B$, $\forall x \in \Omega$. Since Ω is a nonempty open interval, then $\exists (E_j)_{j=1}^\infty := ((a_j, b_j))_{j=1}^\infty \subseteq \mathcal{O}_{\mathbb{R}}$ such that $\Omega = \bigcup_{j=1}^\infty E_j$ and $-\infty < \dots < a_n < \dots < a_2 < a_1 < x_0 < b_1 < b_2 < \dots < b_n < \dots < \infty$. Fix any $j \in \mathbb{N}$. Let $\mathcal{E}_j := (E_j, \mathcal{B}_j, \mu_j)$ be the finite measure subspace of $(\Omega, \mathcal{B}, \mu)$. Then, $f \in \bar{L}_1(\mathcal{E}_j, \mathcal{Y})$. Let $c \in (0, 1)$ be a constant. Define $A_j := \{x \in E_j \mid x \text{ is a rectangular Lebesgue point of } f|_{E_j} \text{ with regularity } c\}$. By Proposition 12.80, $A_j \in \mathcal{B}_L$ and $\mu_L(E_j \setminus A_j) = 0$. Fix any $x \in A_j$.

We have $\limsup_{r \rightarrow 0^+} \sup_{\substack{r_{\tilde{x}_1, \tilde{x}_2} \subset \mathbb{R} \text{ with } \frac{1}{2}(\tilde{x}_1 + \tilde{x}_2) = x, \\ |\tilde{x}_2 - \tilde{x}_1| = r, (\tilde{x}_2 - \tilde{x}_1) \geq cr}} \frac{1}{\mu(r_{\tilde{x}_1, \tilde{x}_2} \cap E_j)} \int_{r_{\tilde{x}_1, \tilde{x}_2} \cap E_j} \|f(y) - f(x)\| d\mu(y) = 0$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists r_0 \in (0, \infty) \subset \mathbb{R}$, $\forall r \in (0, r_0) \subset \mathbb{R}$, we have $0 \leq \sup_{\substack{r_{\tilde{x}_1, \tilde{x}_2} \subset \mathbb{R} \text{ with } \frac{1}{2}(\tilde{x}_1 + \tilde{x}_2) = x, \\ |\tilde{x}_2 - \tilde{x}_1| = r, (\tilde{x}_2 - \tilde{x}_1) \geq cr}} \frac{1}{\mu(r_{\tilde{x}_1, \tilde{x}_2} \cap E_j)} \int_{r_{\tilde{x}_1, \tilde{x}_2} \cap E_j} \|f(y) - f(x)\| d\mu(y) < \epsilon$. Let $r_{\tilde{x}_1, \tilde{x}_2} \subseteq \mathbb{R}$ be such that $\frac{1}{2}(\tilde{x}_1 + \tilde{x}_2) = x$, $\tilde{x}_2 - x < r_0/2$, and $\bar{r}_{\tilde{x}_1, \tilde{x}_2} \subseteq E_j$. $\forall \tilde{x} \in r_{\tilde{x}_1, \tilde{x}_2}$ with $\tilde{x} \neq x$, we have, where $h := \tilde{x} - x$,

$$\begin{aligned}
0 &\leq \|F(\tilde{x}) - F(x) - f(x)(\tilde{x} - x)\| \\
&= \left\| \int_x^{x+h} f(y) d\mu(y) - \int_x^{x+h} f(x) d\mu(y) \right\| \\
&= \left\| \int_x^{x+h} (f(y) - f(x)) d\mu(y) \right\| \\
&= \left\| \int_{r_{x \wedge (x+h), x \vee (x+h)}} (f(y) - f(x)) d\mu(y) \right\| \\
&\leq \int_{r_{x-|h|, x+|h|}} \|f(y) - f(x)\| d\mu(y) \\
&= 2|h| \frac{1}{\mu_B(r_{x-|h|, x+|h|})} \int_{r_{x-|h|, x+|h|}} \|f(y) - f(x)\| d\mu_B(y) \\
&\leq 2\epsilon |\tilde{x} - x|
\end{aligned}$$

where the first equality follows from Fact 12.72 and Proposition 11.75; the second equality follows from Proposition 11.92; the third equality follows from Definition 12.71; the second inequality follows from Proposition 11.92; and the last inequality follows from the fact $x \in A_j$ and the choice of r_0 . Then, $x \in \text{dom}(D_1 F) = U_1$ and $D_1 F(x) = f(x)$. By the arbitrariness

of x , we have $A_j \subseteq U_1$. Clearly, $\bigcup_{j=1}^{\infty} A_j \subseteq U_1 \subseteq \Omega$. This implies that $0 \leq \mu(\Omega \setminus U_1) = \mu_B(\Omega \setminus U_1) = \mu_L(\Omega \setminus U_1) \leq \mu_L(\bigcup_{j=1}^{\infty} (E_j \setminus (\bigcup_{l=1}^{\infty} A_l))) \leq \mu_L(\bigcup_{j=1}^{\infty} (E_j \setminus A_j)) \leq 0$. Hence, $\mu(\Omega \setminus U_1) = 0$. By D_1F being $\mathcal{B}$ -measurable and f being $\mathcal{B}$ -measurable, we have $D_1F(x) = f(x)$ a.e. in $\mu(x)$. This case is proved.

Case 2: $m > 1$. Without loss of generality, consider $i = 1$. Let $(\pi_1(\Omega), \bar{\mathcal{B}}, \bar{\mu})$ be the measure subspace of $\mathbb{R}$. By Proposition 12.85, $\exists \bar{U}_1 \in \mathcal{B}$ with $\mu(\Omega \setminus \bar{U}_1) = 0$, and p_1 as defined in the statement of the theorem, such that p_1 is $\mathcal{B}$ -measurable. By Fubini's Theorem 12.30, we have $F(x) = \int_{\pi_1(x_0)}^{\pi_1(x)} p_1(M_1(s, \pi_1(x))) d\bar{\mu}(s)$, $\forall x \in \Omega$. Then, by Case 1, we have $D_1F(x) = p_1(x)$ a.e. in $\bar{\mu}(\pi_1(x))$, $\forall \pi_1(x) \in \pi_1(\Omega)$. Since D_1F and p_1 are $\mathcal{B}$ -measurable and $\mathcal{Y}$ is separable, then, by Propositions 7.23, 11.39, and 11.38, $D_1F - p_1$ is $\mathcal{B}$ -measurable. Then, $\hat{U}_1 := \{x \in \Omega \mid D_1F(x) = p_1(x)\} \in \mathcal{B}$. By Lemma 12.27, $\mu(\Omega \setminus \hat{U}_1) =: \mu(E) = \int_{\pi_1(\Omega)} \bar{\mu}(E_{\langle \pi_1(x) \rangle}) d\mu_1(\pi_1(x)) = \int_{\pi_1(\Omega)} 0 d\mu_1 = 0$, where $E_{\langle \pi_1(x) \rangle} := \{\bar{s} \in \pi_1(\Omega) \mid M_1(\bar{s}, \pi_1(x)) \in E\}$; and $\bar{\mu}(E_{\langle \pi_1(x) \rangle}) = 0$ follows from $D_1F(x) = p_1(x)$ a.e. in $\bar{\mu}(\pi_1(x))$, $\forall \pi_1(x) \in \pi_1(\Omega)$. Hence, $D_1F = p_1$ a.e. in μ , and $U_1 \in \mathcal{B}$ with $\mu(\Omega \setminus U_1) = 0$. This case is also proved.

In both cases, (ii) holds. Hence, (ii) is true.

(iii) This follows immediately from (ii) and Proposition 12.85.

This completes the proof of the theorem. $\square$

Proposition 12.87 *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be region, $\mathcal{Z}$ be a normed linear space, $G : \Omega \rightarrow \mathcal{Z}$, $\bar{J} := \{1, \dots, m\}$, and $x_0 \in \Omega$. $\forall J \subseteq \bar{J}$, define $\pi_J : \mathbb{R}^m \rightarrow \mathbb{R}^{\text{card}(J)}$ by $\pi_J(x) = (\pi_i(x))_{i \in J}$, $\forall x \in \mathbb{R}^m$; define $M_J : \mathbb{R}^{\text{card}(J)} \times \mathbb{R}^{m-\text{card}(J)} \rightarrow \mathbb{R}^m$ by $M_J(s, t) \in \mathbb{R}^m$ is such that $\pi_J(M_J(s, t)) = s$ and $\pi_{\bar{J} \setminus J}(M_J(s, t)) = t$, $\forall s \in \mathbb{R}^{\text{card}(J)}$, $\forall t \in \mathbb{R}^{m-\text{card}(J)}$; also define $G_J : \pi_J(\Omega) \rightarrow \mathcal{Y}$ by $G_J(s) = G(M_J(s, \pi_{\bar{J} \setminus J}(x_0)))$, $\forall s \in \pi_J(\Omega)$. Then, we have, $\forall x \in \Omega$,*

$$\begin{aligned} (-1)^{\text{card}(\bar{J})} \Delta_G(r_{\hat{x}, \tilde{x}}) &= G(x) - G(x_0) \\ &\quad - \sum_{J \subset \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \Delta_{G_J}(\pi_J(r_{\hat{x}, \tilde{x}})) \end{aligned} \quad (12.2)$$

where $\hat{x} := x \wedge x_0$, $\tilde{x} := x \vee x_0$; and $\bar{J} := \{i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x)\}$.

Proof We will prove the result using mathematical induction on m .

1° $m = 1$. This is obvious.

2° Assume the claim holds for $m \leq k \in \mathbb{N}$.

3° Consider the case when $m = k + 1 \in \{2, 3, \dots\}$. Fix any $x \in \Omega$. Define $\bar{G} : \pi_{\bar{J} \setminus \{k+1\}}(\Omega) \rightarrow \mathcal{Z}$ by $\bar{G}(s) = G(M_{\bar{J} \setminus \{k+1\}}(s, \pi_{k+1}(x)))$, $\forall s \in \pi_{\bar{J} \setminus \{k+1\}}(\Omega)$. Define $\bar{G}_J : \pi_J(\Omega) \rightarrow \mathcal{Z}$ in terms of $\bar{G}$ in a similar manner as

G_J in terms of G , $\forall J \subseteq \bar{J} \setminus \{k+1\}$. Then, we have the following sequence of arguments.

$$\begin{aligned}
& (-1)^{\text{card}(\bar{J})} \Delta_G(\mathbf{r}_{\hat{x}, \hat{x}}) \\
&= (-1)^{\text{card}(\bar{J} \setminus \{k+1\})} (\Delta_{\bar{G}}(\pi_{\bar{J} \setminus \{k+1\}}(\mathbf{r}_{\hat{x}, \hat{x}})) - \Delta_G(\pi_{\bar{J} \setminus \{k+1\}}(\mathbf{r}_{\hat{x}, \hat{x}}))) \\
&= \bar{G}(\pi_{\bar{J} \setminus \{k+1\}}(x)) - \bar{G}(\pi_{\bar{J} \setminus \{k+1\}}(x_0)) \\
&\quad - \sum_{J \subset \bar{J} \setminus \{k+1\}, J \neq \emptyset} (-1)^{\text{card}((\bar{J} \setminus \{k+1\}) \cap J)} \Delta_{\bar{G}_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&\quad - (-1)^{\text{card}(\bar{J} \setminus \{k+1\})} \Delta_G(\pi_{\bar{J} \setminus \{k+1\}}(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&= G(x) - G_{\{k+1\}}(\pi_{k+1}(x)) - (-1)^{\text{card}(\bar{J} \setminus \{k+1\})} \Delta_G(\pi_{\bar{J} \setminus \{k+1\}}(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&\quad - \sum_{J \subset \bar{J} \setminus \{k+1\}, J \neq \emptyset} (-1)^{\text{card}((\bar{J} \setminus \{k+1\}) \cap J)} \Delta_{G_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&\quad - \sum_{\substack{J \subset \bar{J} \setminus \{k+1\} \\ J \neq \emptyset}} (-1)^{\text{card}((\bar{J} \setminus \{k+1\}) \cap J)} (\Delta_{\bar{G}_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) - \Delta_{G_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}}))) \\
&= G(x) - G_{\{k+1\}}(\pi_{k+1}(x)) - \sum_{\substack{J \subseteq \bar{J} \setminus \{k+1\} \\ J \neq \emptyset}} (-1)^{\text{card}(\bar{J} \cap J)} \Delta_{G_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&\quad - \sum_{J \subset \bar{J}, J \supset \{k+1\}} (-1)^{\text{card}(\bar{J} \cap J)} \Delta_{G_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&= G(x) - G_{\{k+1\}}(\pi_{k+1}(x)) - \sum_{J \subseteq \bar{J}, k+1 \notin J, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \Delta_{G_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&\quad - \sum_{J \subset \bar{J}, J \supset \{k+1\}} (-1)^{\text{card}(\bar{J} \cap J)} \Delta_{G_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&= G(x) - G_{\{k+1\}}(\pi_{k+1}(x)) - \sum_{J \subseteq \bar{J}, k+1 \notin J, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \Delta_{G_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&\quad - \sum_{J \subset \bar{J}, J \supset \{k+1\}} (-1)^{\text{card}(\bar{J} \cap J)} \Delta_{G_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&\quad + (-1)^{\text{card}(\bar{J} \cap \{k+1\})} \Delta_{G_{\{k+1\}}}(\pi_{\{k+1\}}(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&= G(x) - G_{\{k+1\}}(\pi_{k+1}(x)) - \sum_{J \subset \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \Delta_{G_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&\quad + (-1)^{\text{card}(\bar{J} \cap \{k+1\})} \Delta_{G_{\{k+1\}}}(\pi_{\{k+1\}}(\mathbf{r}_{\hat{x}, \hat{x}})) \\
&= G(x) - G(x_0) - \sum_{J \subset \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \Delta_{G_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}}))
\end{aligned}$$

where the first equality follows from Definition 12.41; the second equality follows from the inductive assumption 2°; the fourth equality follows from Definition 12.41. This establishes the claim in this step, and completes the induction process.

This completes the proof of the claim. $\square$

Theorem 12.88 (Fundamental Theorem of Calculus II) *Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be an open rectangle, $\mathbb{X} := ((\Omega, |\cdot|), \mathcal{B}, \mu)$ be the σ -finite metric measure subspace of $\mathbb{R}^m$, $\mathcal{Y}$ be a separable reflexive Banach space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable, $F : \Omega \rightarrow \mathcal{Y}$ be absolutely continuous, $\bar{J} := \{1, \dots, m\}$, $F^{(1)} : U \rightarrow \mathcal{B}(\mathbb{R}^m, \mathcal{Y})$, $D_i F : U_i \rightarrow \mathcal{Y}$, $\forall i \in \bar{J}$, where D_i is the derivative of a function with respect to the i th coordinate variable of Ω , and $x_0 \in \Omega$. $\forall J \subseteq \bar{J}$, define $\pi_J : \mathbb{R}^m \rightarrow \mathbb{R}^{\text{card}(J)}$ by $\pi_J(x) = (\pi_i(x))_{i \in J}$, $\forall x \in \mathbb{R}^m$; define $M_J : \mathbb{R}^{\text{card}(J)} \times \mathbb{R}^{m-\text{card}(J)} \rightarrow \mathbb{R}^m$ by $M_J(s, t) \in \mathbb{R}^m$ is such that $\pi_J(M_J(s, t)) = s$ and $\pi_{\bar{J} \setminus J}(M_J(s, t)) = t$, $\forall s \in \mathbb{R}^{\text{card}(J)}$, $\forall t \in \mathbb{R}^{m-\text{card}(J)}$; also define $F_J : \pi_J(\Omega) \rightarrow \mathcal{Y}$ by $F_J(s) = F(M_J(s, \pi_{\bar{J} \setminus J}(x_0)))$, $\forall s \in \pi_J(\Omega)$; and let $(\pi_J(\Omega), \mathcal{B}_J, \mu_J)$ be the measure subspace of $\mathbb{R}^{\text{card}(J)}$ when $J \neq \emptyset$. Then, the following statements hold.*

- (i) *There exists a set of functions $f_{J, x_0} : \pi_J(\Omega) \rightarrow \mathcal{Y}$, which is $\mathcal{B}_J$ -measurable, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$, such that f_{J, x_0} 's are absolutely integrable on $\pi_J(\mathbb{r}_{x_a, x_b})$ with respect to μ_J , $\forall x_a, x_b \in \Omega$ with $x_a \leq x_b$, and we have, $\forall x \in \Omega$,*

$$F(x) = \sum_{J \subseteq \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(\mathbb{r}_{\hat{x}, \check{x}})} f_{J, x_0} d\mu_J + F(x_0) \quad (12.3)$$

where $\hat{x} := x \wedge x_0$, $\check{x} := x \vee x_0$; and $\bar{J} := \{i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x)\}$. The set of functions $(f_{J, x_0})_{J \subseteq \bar{J}, J \neq \emptyset}$ is unique in the sense that if $(g_{J, x_0})_{J \subseteq \bar{J}, J \neq \emptyset}$ is another set of such functions satisfying (12.3), then $f_{J, x_0} = g_{J, x_0}$ a.e. in μ_J , $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$.

- (ii) *If we expand absolutely continuous function F at $\bar{x} \in \Omega$, instead of x_0 , then we will have a set of functions $f_{J, \bar{x}} : \pi_J(\Omega) \rightarrow \mathcal{Y}$, which is $\mathcal{B}_J$ -measurable, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$, satisfying the conclusion of (i) with x_0 replaced by $\bar{x}$. The $f_{J, \bar{x}}$'s relate to f_{J, x_0} 's according to the following formula, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$, $\exists \bar{U}_J \in \mathcal{B}_J$ with $\mu_J(\pi_J(\Omega) \setminus \bar{U}_J) = 0$, such that*

$$f_{J, \bar{x}}(s) = \sum_{J \subset \bar{J} \subseteq \bar{J}} (-1)^{\text{card}(\bar{J} \cap \bar{J} \setminus J)} \int_{\pi_{\bar{J} \setminus J}(\mathbb{r}_{\hat{\bar{x}}, \check{\bar{x}}})} f_{\bar{J}, x_0}(\pi_{\bar{J}}(M_J(s, \pi_{\bar{J} \setminus J}(t)))) d\mu_{\bar{J} \setminus J}(\pi_{\bar{J} \setminus J}(t)) + f_{J, x_0}(s), \quad \forall s \in \bar{U}_J \quad (12.4)$$

and $f_{J, \bar{x}}(s) = \vartheta_{\mathcal{Y}}$, $\forall s \in \pi_J(\Omega) \setminus \bar{U}_J$, where $\hat{\bar{x}} := \bar{x} \wedge x_0$, $\check{\bar{x}} := \bar{x} \vee x_0$, and $\bar{J} := \{i \in \bar{J} \mid \pi_i(x_0) > \pi_i(\bar{x})\}$; and all involved integrations are such that the integrands are absolutely integrable over the specified integration domain with respect to the given measures, respectively.

- (iii) $U_i \in \mathcal{B}$, $\mu(\Omega \setminus U_i) = 0$, $D_i F$ is $\mathcal{B}$ -measurable, and, $\forall i \in \bar{J}$, $D_i F(x) = f_{\{i\},x}(\pi_{\{i\}}(x))$ a.e. in $\mu(x)$.
- (iv) $U \subseteq \bigcap_{i=1}^{\infty} U_i$, $U \in \mathcal{B}$, $F^{(1)}$ is $\mathcal{B}$ -measurable, and, $\forall x \in U$, we have $F^{(1)}(x) = [D_1 F(x) \cdots D_m F(x)]$. If, in addition, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$, $\forall x \in \Omega$, $\exists \delta(J, x) \in (0, \infty) \subset \mathbb{R}$, $\exists c(J, x) \in [0, \infty) \subset \mathbb{R}$ such that $\text{ess sup}_{y \in \mathcal{B}_{\mathbb{R}^m}(x, \delta(J, x))} \|f_{J, x_0}(\pi_J(y))\| \leq c(J, x)$ and $\mathcal{B}_{\mathbb{R}^m}(x, \delta(J, x)) \subseteq \Omega$, then $\mu(\Omega \setminus U) = 0$.
- (v) If, in addition, $\mathcal{Y}$ is finite dimensional, then $\bar{U}_J = \left\{ s \in \pi_J(\Omega) \mid f_{\bar{J}, x_0}(\pi_{\bar{J}}(M_J(s, \pi_{\bar{J} \setminus J}(t)))) \right.$
 $\left. f_{\bar{J}, x_0}(\pi_{\bar{J}}(M_J(s, \pi_{\bar{J} \setminus J}(t)))) \right\}$ is absolutely integrable over $\pi_{\bar{J} \setminus J}(\mathbf{r}_{\hat{x}, \hat{x}})$
 with respect to $\mu_{\bar{J} \setminus J}(\pi_{\bar{J} \setminus J}(t))$, $\forall \tilde{J} \subseteq \bar{J}$ with $J \subset \tilde{J}$, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$.

Proof (i) By $F = F_{\bar{J}}$ being absolutely continuous and Definition 12.41, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$, F_J is absolutely continuous. By Proposition 12.73, there exists a unique σ -finite $\mathcal{Y}$ -valued measure ν_J on $(\pi_J(\Omega), \mathcal{B}_J)$ such that F_J is a cumulative distribution function of ν_J . Furthermore, $\mathcal{P} \circ \nu_J \ll \mu_J$. Since $\mathcal{Y}$ is a separable reflexive Banach space with $\mathcal{Y}^*$ being separable, then, by Radon-Nikodym Theorem 11.171, there exists a unique $f_{J, x_0} : \pi_J(\Omega) \rightarrow \mathcal{Y}$, which is $\mathcal{B}_J$ -measurable, such that $f_{J, x_0} = \frac{d\nu_J}{d\mu_J}$ a.e. in μ_J . $\forall x_a, x_b \in \Omega$ with $x_a \leq x_b$, we have $\mathcal{P} \circ \nu_J(\pi_J(\mathbf{r}_{x_a, x_b})) = \int_{\pi_J(\mathbf{r}_{x_a, x_b})} \mathcal{P} \circ f_{J, x_0} d\mu_J = T_{F_J}(\pi_J(\mathbf{r}_{x_a, x_b})) < \infty$, where the first equality follows from Definition 11.166; the second equality follows from Proposition 12.73; and the inequality follows from the fact F_J is absolutely continuous and therefore of locally bounded variation which implies this according to Definition 12.41. Hence, f_{J, x_0} is absolutely integrable over $\pi_J(\mathbf{r}_{x_a, x_b})$ with respect to μ_J . Then, $\forall x \in \Omega$, we have $\Delta_{F_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) = \nu_J(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) = \int_{\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})} f_{J, x_0}(s) d\mu_J(s)$. By Proposition 12.87, we have

$$\begin{aligned} (-1)^{\text{card}(\bar{J})} \Delta_F(\mathbf{r}_{\hat{x}, \hat{x}}) &= F(x) - F(x_0) \\ &\quad - \sum_{J \subset \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \Delta_{F_J}(\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})) \end{aligned}$$

which implies that

$$\begin{aligned} &(-1)^{\text{card}(\bar{J})} \int_{\pi_{\bar{J}}(\mathbf{r}_{\hat{x}, \hat{x}})} f_{\bar{J}, x_0}(s) d\mu_{\bar{J}}(s) \\ &= F(x) - F(x_0) - \sum_{J \subset \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})} f_{J, x_0}(s) d\mu_J(s) \end{aligned}$$

Reorganizing the above formula, we have

$$F(x) = F(x_0) + \sum_{J \subset \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(\mathbf{r}_{\hat{x}, \hat{x}})} f_{J, x_0}(s) d\mu_J(s)$$

This establishes (12.3). The uniqueness of the set of functions $(f_{J,x_0})_{J \subseteq \bar{J}, J \neq \emptyset}$ is clear from the uniqueness of F_J 's and the uniqueness of Radon-Nikodym derivatives. Hence, (i) holds.

(ii) By (i), there exists a set of functions $f_{J,\bar{x}} : \pi_J(\Omega) \rightarrow \mathcal{Y}$, which is $\mathcal{B}$ -measurable, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$, such that $f_{J,\bar{x}}$'s are absolutely integrable on $\pi_J(r_{x_a, x_b})$ with respect to μ_J , $\forall x_a, x_b \in \Omega$ with $x_a \leq x_b$, and we have $F(x) = \sum_{J \subseteq \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(r_{\hat{x}, \bar{x}})} f_{J,\bar{x}} d\mu_J + F(\bar{x})$, where $\hat{x} := \bar{x} \wedge x$, $\check{x} := \bar{x} \vee x$, and $\check{\bar{J}} := \{i \in \bar{J} \mid \pi_i(\bar{x}) > \pi_i(x)\}$. Now, we let $x = \bar{x} + M_J(s, \mathbf{0}_{m-\text{card}(J)})$, $\forall s \in \mathbb{R}^{\text{card}(J)}$, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$. By the preceding formula, we have $F(x) = \sum_{\hat{J} \subseteq J, \hat{J} \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \hat{J})} \int_{\pi_{\hat{J}}(r_{\hat{x}, \bar{x}})} f_{\hat{J}, \bar{x}} d\mu_{\hat{J}} + F(\bar{x})$. This is because $\forall \hat{J} \not\subseteq J$, the domain of integration will be an empty set and that leads to $\vartheta_{\mathcal{Y}}$ for that integral. Now, substitute the formula for $F(x)$ and $F(\bar{x})$ using (12.3), we have

$$\begin{aligned} \sum_{\hat{J} \subseteq J, \hat{J} \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \hat{J})} \int_{\pi_{\hat{J}}(r_{\hat{x}, \bar{x}})} f_{\hat{J}, x_0} d\mu_{\hat{J}} &= \sum_{\hat{J} \subseteq J, \hat{J} \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \hat{J})} \int_{\pi_{\hat{J}}(r_{\hat{x}, \bar{x}})} \\ &\quad f_{\hat{J}, \bar{x}} d\mu_{\hat{J}} + \sum_{\hat{J} \subseteq \bar{J}, \hat{J} \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \hat{J})} \int_{\pi_{\hat{J}}(r_{\bar{x}, \bar{x}})} f_{\hat{J}, x_0} d\mu_{\hat{J}} \end{aligned}$$

Rearrange terms in the above, we obtain

$$\begin{aligned} \sum_{\hat{J} \subseteq J, \hat{J} \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \hat{J})} \int_{\pi_{\hat{J}}(r_{\hat{x}, \bar{x}})} f_{\hat{J}, \bar{x}} d\mu_{\hat{J}} &= \sum_{\hat{J} \subseteq \bar{J}, \hat{J} \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \hat{J})} \int_{\pi_{\hat{J}}(r_{\hat{x}, \bar{x}})} \\ &\quad f_{\hat{J}, x_0} d\mu_{\hat{J}} - \sum_{\hat{J} \subseteq \bar{J}, \hat{J} \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \hat{J})} \int_{\pi_{\hat{J}}(r_{\bar{x}, \bar{x}})} f_{\hat{J}, x_0} d\mu_{\hat{J}} \\ &= \sum_{\substack{\hat{J} \subseteq \bar{J} \\ \hat{J} \neq \emptyset}} \prod_{j \in \hat{J}} \left(\int_{\pi_j(x_0)}^{\pi_j(x)} \right) f_{\hat{J}, x_0}(s_1, \dots, s_{\text{card}(\hat{J})}) d\mu_B(s_{\text{card}(\hat{J})}) \cdots d\mu_B(s_1) \\ &\quad - \sum_{\substack{\hat{J} \subseteq \bar{J} \\ \hat{J} \neq \emptyset}} \prod_{j \in \hat{J}} \left(\int_{\pi_j(x_0)}^{\pi_j(\bar{x})} \right) f_{\hat{J}, x_0}(s_1, \dots, s_{\text{card}(\hat{J})}) d\mu_B(s_{\text{card}(\hat{J})}) \cdots d\mu_B(s_1) \\ &= \sum_{\hat{J} \subseteq \bar{J}, \hat{J} \neq \emptyset} \prod_{j \in \hat{J}} \left(\int_{\pi_j(x_0)}^{\pi_j(\bar{x})} + \int_{\pi_j(\bar{x})}^{\pi_j(x)} \right) f_{\hat{J}, x_0}(s_1, \dots, s_{\text{card}(\hat{J})}) \\ &\quad \cdot d\mu_B(s_{\text{card}(\hat{J})}) \cdots d\mu_B(s_1) - \sum_{\hat{J} \subseteq \bar{J}, \hat{J} \neq \emptyset} \prod_{j \in \hat{J}} \left(\int_{\pi_j(x_0)}^{\pi_j(\bar{x})} \right) \\ &\quad f_{\hat{J}, x_0}(s_1, \dots, s_{\text{card}(\hat{J})}) d\mu_B(s_{\text{card}(\hat{J})}) \cdots d\mu_B(s_1) \end{aligned}$$

$$\begin{aligned}
&= \sum_{\hat{J} \subseteq \bar{J}, \hat{J} \neq \emptyset} \sum_{\bar{J} \subseteq \hat{J}, \bar{J} \neq \emptyset} \prod_{j \in \bar{J}} \left(\int_{\pi_j(\bar{x})}^{\pi_j(x)} \right) \prod_{l \in \hat{J} \setminus \bar{J}} \left(\int_{\pi_l(x_0)}^{\pi_l(\bar{x})} \right) \\
&\quad f_{\hat{J}, x_0}(s_1, \dots, s_{\text{card}(\hat{J})}) d\mu_B(s_{\text{card}(\hat{J})}) \cdots d\mu_B(s_1) \\
&= \sum_{\substack{\hat{J} \subseteq \bar{J} \\ \bar{J} \neq \emptyset}} \sum_{\substack{\bar{J} \subseteq \hat{J} \\ \bar{J} \neq \emptyset}} (-1)^{\text{card}(\bar{J} \cap \hat{J}) + \text{card}(\bar{J} \cap \hat{J} \setminus \bar{J})} \int_{\pi_{\hat{J}}(\mathbf{r}_{\hat{x}_{\bar{J}, x, \bar{x}, x_0}, \bar{x}_{\bar{J}, x, \bar{x}, x_0}})} f_{\hat{J}, x_0} d\mu_{\hat{J}} \\
&= \sum_{\hat{J} \subseteq \bar{J}, \hat{J} \neq \emptyset} \sum_{\bar{J} \subseteq \hat{J} \cap \bar{J}, \bar{J} \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \hat{J}) + \text{card}(\bar{J} \cap \hat{J} \setminus \bar{J})} \int \pi_{\hat{J}}(\mathbf{r}_{\hat{x}_{\bar{J}, x, \bar{x}, x_0}, \bar{x}_{\bar{J}, x, \bar{x}, x_0}}) \\
&\quad f_{\hat{J}, x_0} d\mu_{\hat{J}} \\
&= \sum_{\bar{J} \subseteq \bar{J}, \bar{J} \neq \emptyset} \sum_{\hat{J} \subseteq \bar{J} \cap \bar{J}, \hat{J} \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \hat{J}) + \text{card}(\bar{J} \cap \hat{J} \setminus \bar{J})} \int \pi_{\bar{J}}(\mathbf{r}_{\hat{x}_{\bar{J}, x, \bar{x}, x_0}, \bar{x}_{\bar{J}, x, \bar{x}, x_0}}) \\
&\quad f_{\bar{J}, x_0} d\mu_{\bar{J}} \\
&= \sum_{\substack{\hat{J} \subseteq J \\ \hat{J} \neq \emptyset}} \sum_{\substack{\bar{J} \subseteq \bar{J} \\ \bar{J} \neq \emptyset}} (-1)^{\text{card}(\bar{J} \cap \hat{J}) + \text{card}(\bar{J} \cap \hat{J} \setminus \bar{J})} \int \pi_{\bar{J}}(\mathbf{r}_{\hat{x}_{\bar{J}, x, \bar{x}, x_0}, \bar{x}_{\bar{J}, x, \bar{x}, x_0}}) f_{\bar{J}, x_0} d\mu_{\bar{J}} \\
&= \sum_{\hat{J} \subseteq J, \hat{J} \neq \emptyset} \left(\sum_{\bar{J} \subseteq \bar{J}, \bar{J} \supset \hat{J}} (-1)^{\text{card}(\bar{J} \cap \hat{J}) + \text{card}(\bar{J} \cap \hat{J} \setminus \bar{J})} \int \pi_{\bar{J}}(\mathbf{r}_{\hat{x}, \bar{x}}) p_{\bar{J}, \hat{J}} d\mu_{\bar{J}} \right. \\
&\quad \left. + (-1)^{\text{card}(\bar{J} \cap \hat{J})} \int \pi_{\hat{J}}(\mathbf{r}_{\hat{x}, \bar{x}}) f_{\hat{J}, x_0} d\mu_{\hat{J}} \right) \\
&= \sum_{\substack{\hat{J} \subseteq J \\ \hat{J} \neq \emptyset}} (-1)^{\text{card}(\bar{J} \cap \hat{J})} \int \pi_{\hat{J}}(\mathbf{r}_{\hat{x}, \bar{x}}) \left(\sum_{\substack{\bar{J} \subseteq \bar{J} \\ \bar{J} \supset \hat{J}}} (-1)^{\text{card}(\bar{J} \cap \hat{J} \setminus \bar{J})} p_{\bar{J}, \hat{J}} + f_{\hat{J}, x_0} \right) d\mu_{\hat{J}}
\end{aligned}$$

where $\hat{x}_{\bar{J}, x, \bar{x}, x_0} := x_{1, \bar{J}, x, \bar{x}, x_0} \wedge x_{2, \bar{J}, x, \bar{x}, x_0}$, $\bar{x}_{\bar{J}, x, \bar{x}, x_0} := x_{1, \bar{J}, x, \bar{x}, x_0} \vee x_{2, \bar{J}, x, \bar{x}, x_0}$; $x_{1, \bar{J}, x, \bar{x}, x_0} := M_{\bar{J}}(\pi_{\bar{J}}(x), \pi_{\bar{J} \setminus \bar{J}}(\bar{x}))$, $x_{2, \bar{J}, x, \bar{x}, x_0} := M_{\bar{J}}(\pi_{\bar{J}}(\bar{x}), \pi_{\bar{J} \setminus \bar{J}}(x_0))$; the second to fifth equalities follow from intuition, we have not used rigorous notation in these equations in expressing the joint integrals as repeated integrals and the other way around, this should not cause concern since from the first equality to the fifth, all that we have done is breaking down $(-1)^{\text{card}(\bar{J} \cap \hat{J})} \int_{\pi_{\hat{J}}(\mathbf{r}_{\hat{x}, \bar{x}})} f_{\hat{J}, x_0} d\mu_{\hat{J}}$ into pieces along the intermediate point $\bar{x}$; the sixth equality follows from the idea that the corresponding integral will be zero if its domain of integration is an empty set, which will happen since $\pi_i(x) = \pi_i(\bar{x})$, $\forall i \in \bar{J} \setminus J$; the seventh equality follows from the interchange of summing variables $\hat{J}$ and $\bar{J}$; the ninth equality follows from Fubini's Theorem 12.30, and $\bar{U}_{\bar{J}, \hat{J}} \in \mathcal{B}_J$ with $\mu_{\hat{J}}(\pi_{\hat{J}}(\Omega) \setminus \bar{U}_{\bar{J}, \hat{J}}) = 0$, and $p_{\bar{J}, \hat{J}} : \pi_{\hat{J}}(\Omega) \rightarrow \mathcal{Y}$, $\forall \hat{J} \subseteq \bar{J}$, defined by $p_{\bar{J}, \hat{J}}(s) := \begin{cases} \int_{\pi_{\bar{J} \setminus \hat{J}}(\mathbf{r}_{\hat{x}, \bar{x}})} f_{\bar{J}, x_0}(\pi_{\bar{J}}(M_{\hat{J}}(s, \pi_{\bar{J} \setminus \hat{J}}(t)))) d\mu_{\bar{J} \setminus \hat{J}}(\pi_{\bar{J} \setminus \hat{J}}(t)) & s \in \bar{U}_{\bar{J}, \hat{J}} \\ \vartheta_{\mathcal{Y}} & s \in \pi_{\hat{J}}(\Omega) \setminus \bar{U}_{\bar{J}, \hat{J}} \end{cases}$

is $\mathcal{B}_{\bar{J}}$ -measurable; and the tenth equality follows from Proposition 11.92. Let $\bar{U}_{\bar{J}} := \bigcap_{\bar{J} \subsetneq \bar{J}' \subseteq \bar{J}} \bar{U}_{\bar{J}', \bar{J}}$, $\forall \hat{J} \subseteq \bar{J}$ with $\hat{J} \neq \emptyset$. Clearly, $\bar{U}_{\bar{J}} \in \mathcal{B}_{\bar{J}}$ and $\mu_{\bar{J}}(\pi_{\bar{J}}(\Omega) \setminus \bar{U}_{\bar{J}}) = 0$. The definition of $f_{J, \bar{x}}$'s (12.4) satisfies the above equalities, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$. Then, by (i), this set of functions $(f_{J, \bar{x}})_{J \subseteq \bar{J}, J \neq \emptyset}$ is the unique one we seek.

(iii) Fix any $i \in \bar{J}$. Without loss of generality, let $i = 1$. By Proposition 12.81, $U_1 \in \mathcal{B}$ and $D_1 F$ is $\mathcal{B}$ -measurable. By Fundamental Theorem of Calculus 12.86, every term on the right-hand-side of (12.3) is differentiable with respect to the i th coordinate variable of Ω almost everywhere in $\mathbb{X}$. $\forall J \subseteq \bar{J}$ with $1 \in J$, $\exists U_{J,1} \in \mathcal{B}$ with $\mu(\Omega \setminus U_{J,1}) = 0$, $D_1 \left((-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(r_{\bar{x}, \bar{x}})} f_{J, x_0} d\mu_J \right) : U_{J,1} \rightarrow \mathcal{Y}$ is $\mathcal{B}$ -measurable, $\exists \bar{U}_{J,1} \in \mathcal{B}$ with $\mu(\Omega \setminus \bar{U}_{J,1}) = 0$, and $\exists p_{J,1} : \Omega \rightarrow \mathcal{Y}$ defined by $p_{J,1} = \begin{cases} (-1)^{\text{card}(\bar{J} \cap J \setminus \{1\})} \int_{\pi_{J \setminus \{1\}}(r_{\bar{x}, \bar{x}})} f_{J, x_0} (\pi_J(M_{\{1\}}(\pi_1(x)), \pi_{\bar{J} \setminus \{1\}}(s))) d\mu_{J \setminus \{1\}}(\pi_{J \setminus \{1\}}(s)) & x \in \bar{U}_{J,1} \\ \vartheta_{\mathcal{Y}} & x \in \Omega \setminus \bar{U}_{J,1} \end{cases}, \forall x \in \Omega$, when $J \supset \{1\}$, $p_{J,1} = f_{J, x_0}(\pi_J(x))$, $\forall x \in \Omega$, $\bar{U}_{J,1} = \Omega$, when $J = \{1\}$, such that $p_{J,1}$ is $\mathcal{B}$ -measurable and $D_1 \left((-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(r_{\bar{x}, \bar{x}})} f_{J, x_0} d\mu_J \right) = p_{J,1}(x)$ a.e. in $\mu(x)$. $\forall J \subseteq \bar{J}$ with $1 \notin J$, we have $D_1 \left((-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(r_{\bar{x}, \bar{x}})} f_{J, x_0} d\mu_J \right) = \vartheta_{\mathcal{Y}}$, $\forall x \in \Omega$. In this case, we have $\bar{U}_{J,1} = \Omega$. Let $\bar{U}_1 := \bigcap_{J \subseteq \bar{J}, J \neq \emptyset} \bar{U}_{J,1}$. Then, $\bar{U}_1 \in \mathcal{B}$ and $\mu(\Omega \setminus \bar{U}_1) = 0$. Hence, $D_1 F(x) = \sum_{J \subseteq \bar{J}, J \neq \emptyset} p_{J,1}(x) = f_{\{1\}, x}(\pi_{\{1\}}(x))$ a.e. in $\mu(x)$. This implies that $\mu(\Omega \setminus U_1) = 0$ and (iii) holds.

(iv) By Proposition 12.81, $U \subseteq \bigcap_{i=1}^m U_i$, $F^{(1)}$ is $\mathcal{B}$ -measurable, and $\forall x \in U$, we have $F^{(1)}(x) = [D_1 F(x) \cdots D_m F(x)]$. All we need to show is that $\mu(\Omega \setminus U) = 0$ under the additional assumption that f_{J, x_0} 's are locally bounded.

$\forall i \in \bar{J}$, $\forall J \subseteq \bar{J}$ with $i \in J$, by the proof of (iii), $\exists \tilde{U}_{J,i} \in \mathcal{B}$ with $\mu(\Omega \setminus \tilde{U}_{J,i}) = 0$, such that $D_i H_J(x) := D_i \left((-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(r_{\bar{x}, \bar{x}})} f_{J, x_0} d\mu_J \right) (x) = p_{J,i}(x)$, $\forall x \in \tilde{U}_{J,i}$. Then, let $\hat{U}_{J,i} := U_{J,i} \cap \bar{U}_{J,i} \cap \tilde{U}_{J,i} \in \mathcal{B}$, we have $\mu(\Omega \setminus \hat{U}_{J,i}) = 0$ and $D_i H_J(x) = (-1)^{\text{card}(\bar{J} \cap J \setminus \{i\})} \cdot \int_{\pi_{J \setminus \{i\}}(r_{\bar{x}, \bar{x}})} f_{J, x_0} (\pi_J(M_{\{i\}}(\pi_i(x), \pi_{\bar{J} \setminus \{i\}}(s))) d\mu_{J \setminus \{i\}}(\pi_{J \setminus \{i\}}(s)) =: H_{J,i}(x)$, $\forall x \in \hat{U}_{J,i}$.

Let $\check{U} := \left(\bigcap_{i=1}^m \bigcap_{i \in J \subseteq \bar{J}} \hat{U}_{J,i} \right) \cap \left(\bigcap_{J \subseteq \bar{J}, J \neq \emptyset} \bar{U}_J \right) \in \mathcal{B}$. Clearly, $\mu(\Omega \setminus \check{U}) = 0$. Fix any $\bar{x} \in \check{U}$. Then, by (ii), there exists a unique set of functions $(f_{J, \bar{x}})_{J \subseteq \bar{J}, J \neq \emptyset}$ such that $F(x) = \sum_{J \subseteq \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \cdot \int_{\pi_J(r_{\bar{x}, \bar{x}})} f_{J, \bar{x}} d\mu_J + F(\bar{x})$, $\forall x \in \Omega$. Furthermore, $f_{J, \bar{x}}$'s are related to f_{J, x_0} 's by (12.4). Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. By f_{J, x_0} 's being locally bounded on their domain of definition and (12.4), we can conclude that $f_{J, \bar{x}}$'s are also locally bounded on their domain of definition. $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$, $\forall x \in \Omega$, $\exists \bar{\delta}(J, x) \in (0, \infty) \subset \mathbb{R}$, $\exists \bar{c}(J, x) \in [0, \infty) \subset \mathbb{R}$ such that

$\text{ess sup}_{y \in \mathcal{B}_{\mathbb{R}^m}(x, \bar{\delta}(J, x))} \|f_{J, \bar{x}}(\pi_J(y))\| \leq \bar{c}(J, x)$ and $\mathcal{B}_{\mathbb{R}^m}(x, \bar{\delta}(J, x)) \subseteq \Omega$.
 Let $\bar{\delta}_0 := \min_{J \subseteq \bar{J}, J \neq \emptyset} \bar{\delta}(J, \bar{x}) \in (0, \infty) \subset \mathbb{R}$ and $\bar{c}_0 := \max_{J \subseteq \bar{J}, J \neq \emptyset} \bar{c}(J, \bar{x}) \in [0, \infty) \subset \mathbb{R}$. Then, $\text{ess sup}_{y \in \mathcal{B}_{\mathbb{R}^m}(\bar{x}, \bar{\delta}_0)} \|f_{J, \bar{x}}(\pi_J(y))\| \leq \bar{c}_0$ and $\mathcal{B}_{\mathbb{R}^m}(\bar{x}, \bar{\delta}_0) \subseteq \Omega$. $\forall i \in \bar{J}$, we have

$$\begin{aligned} D_i F(\bar{x}) &= f_{\{i\}, x_0}(\pi_i(\bar{x})) + \sum_{J \subseteq \bar{J}, \{i\} \subset J} (-1)^{\text{card}(\bar{J} \cap J \setminus \{i\})} \int_{\pi_{J \setminus \{i\}}(\Gamma_{\bar{x}, \bar{x}})} \\ &\quad f_{J, x_0}(\pi_J(M_{\{i\}}(\pi_i(\bar{x}), \pi_{J \setminus \{i\}}(s)))) d\mu_{J \setminus \{i\}}(\pi_{J \setminus \{i\}}(s)) \\ &= f_{\{i\}, \bar{x}}(\pi_i(\bar{x})) \end{aligned}$$

where the first equality follows from the preceding discussion and the second equality follows from (12.4). $\forall i \in \bar{J}, \forall J \subseteq \bar{J}$ with $i \in J, \exists \delta_{J, i}(\bar{x}) \in (0, \infty) \subset \mathbb{R}$ such that $\|H_J(\bar{x} + h_i \mathbf{e}_{m, i}) - H_J(\bar{x}) - H_{J, i}(\bar{x})h_i\| \leq m^{-1}2^{-m}\epsilon |h_i|$. Let $\delta_1 := (\min_{i \in \bar{J}} \min_{J \subseteq \bar{J}} \delta_{J, i}(\bar{x})) \wedge \bar{\delta}_0 \wedge \frac{\epsilon}{2(1+\bar{c}_0)(2^m-1-m)} \wedge 1 \in (0, \infty) \subset \mathbb{R}$. This leads to, $\forall x \in \mathcal{B}_{\mathbb{R}^m}(\bar{x}, \delta_1)$,

$$\begin{aligned} &\left\| F(x) - F(\bar{x}) - \sum_{i=1}^m D_i F(\bar{x}) \pi_i(x - \bar{x}) \right\| \\ &= \left\| \sum_{J \subseteq \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(\Gamma_{\bar{x}, \bar{x}})} f_{J, \bar{x}} d\mu_J - \sum_{i=1}^m f_{\{i\}, \bar{x}}(\pi_i(\bar{x})) \pi_i(x - \bar{x}) \right. \\ &\quad \left. - \sum_{J \subseteq \bar{J}, \text{card}(J) \geq 2} (-1)^{\text{card}(\bar{J} \cap J)} \int_{\pi_J(\Gamma_{\bar{x}, \bar{x}})} f_{J, \bar{x}} d\mu_J \right. \\ &\quad \left. + \sum_{i=1}^m (-1)^{\text{card}(\bar{J} \cap \{i\})} \int_{\pi_{\{i\}}(\Gamma_{\bar{x}, \bar{x}})} f_{\{i\}, \bar{x}} d\mu_{\{i\}} \right. \\ &\quad \left. - \sum_{i=1}^m (-1)^{\text{card}(\bar{J} \cap \{i\})} \int_{\pi_{\{i\}}(\Gamma_{\bar{x}, \bar{x}})} f_{\{i\}, \bar{x}}(\pi_i(\bar{x})) d\mu_{\{i\}}(t) \right\| \\ &\leq \sum_{J \subseteq \bar{J}, \text{card}(J) \geq 2} \int_{\pi_J(\Gamma_{\bar{x}, \bar{x}})} \mathcal{P} \circ f_{J, \bar{x}} d\mu_J + \left\| \sum_{i=1}^m (-1)^{\text{card}(\bar{J} \cap \{i\})} \int_{\pi_{\{i\}}(\Gamma_{\bar{x}, \bar{x}})} \right. \\ &\quad \left. (f_{\{i\}, \bar{x}}(t) - f_{\{i\}, \bar{x}}(\pi_i(\bar{x}))) d\mu_{\{i\}}(t) \right\| \\ &\leq \sum_{J \subseteq \bar{J}, \text{card}(J) \geq 2} \int_{\pi_J(\Gamma_{\bar{x}, \bar{x}})} \mathcal{P} \circ f_{J, \bar{x}} d\mu_J \\ &\quad + \sum_{i=1}^m \sum_{J \subseteq \bar{J}} \|H_J(\bar{x} + \pi_i(x - \bar{x}) \mathbf{e}_{m, i}) - H_J(\bar{x}) - H_{J, i}(\bar{x}) \pi_i(x - \bar{x})\| \\ &\leq \sum_{J \subseteq \bar{J}, \text{card}(J) \geq 2} \bar{c}_0 \mu_J(\pi_J(\Gamma_{\bar{x}, \bar{x}})) + \sum_{i=1}^m \sum_{J \subseteq \bar{J}} m^{-1} 2^{-m} \epsilon |\pi_i(x - \bar{x})| \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{J \subseteq \bar{J}, \text{card}(J) \geq 2} \bar{c}_0 \prod_{j \in J} |\pi_j(x - \bar{x})| + \frac{\epsilon}{2} |x - \bar{x}| \\
&\leq (2^m - 1 - m) \bar{c}_0 \delta_1^{\text{card}(J)-1} |x - \bar{x}| + \frac{\epsilon}{2} |x - \bar{x}| \leq \epsilon |x - \bar{x}|
\end{aligned}$$

where the first equality follows from the above discussion; the first inequality follows from Proposition 11.92; the second inequality follows from Proposition 11.92; and the third inequality follows from choice of δ_1 and the preceding discussion. Then, $\bar{x} \in \text{dom}(F^{(1)})$. By the arbitrariness of $\bar{x}$, we have $\tilde{U} \subseteq \text{dom}(F^{(1)}) = U$. Hence, $\mu(\Omega \setminus U) = 0$.

(v) This is a direct consequence of Fubini's Theorem 12.31 and the proof of (ii).

This completes the proof of the theorem. $\square$

Theorem 12.89 (Integration by Parts) *Let $I \subseteq \mathbb{R}$ be a nonempty open or closed interval, $\mathbb{I} := ((P(I), |\cdot|), \mathcal{B}, \mu)$ be the σ -finite metric measure subspace of $\mathbb{R}$, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, $\mathcal{Z}$ be a separable reflexive Banach space over $\mathbb{K}$ with $\mathcal{Z}^*$ being separable, and $\mathcal{W} := B(\mathcal{Y}, \mathcal{Z})$ be separable, $a : P(I) \rightarrow \mathcal{W}$ be $\mathcal{B}$ -measurable and absolutely integrable over $\mathbb{I}$, ν_A be the $\mathcal{W}$ -valued measure with kernel a over $\mathbb{I}$, and $A : I \rightarrow \mathcal{W}$ be a cumulative distribution function for ν_A , $f : P(I) \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable and absolutely integrable over $\mathbb{I}$, ν_F be the $\mathcal{Y}$ -valued measure with kernel f over $\mathbb{I}$, and $F : I \rightarrow \mathcal{Y}$ be a cumulative distribution function for ν_F . Then, $H : I \rightarrow \mathcal{Z}$ defined by $H(x) = A(x)F(x)$, $\forall x \in I$, is absolutely continuous and is a cumulative distribution function of ν_H which is the $\mathcal{Z}$ -valued measure with kernel $h : P(I) \rightarrow \mathcal{Z}$ over $\mathbb{I}$, where h is defined by $h(x) = A(x)f(x) + a(x)F(x)$, $\forall x \in P(I)$, and Af , aF , and h are absolutely integrable over $\mathbb{I}$. As a consequence, $\forall b, c \in I$ with $b \leq c$, we have*

$$\int_b^c A(x)f(x) dx + \int_b^c a(x)F(x) dx = A(c)F(c) - A(b)F(b)$$

Proof By Proposition 7.66, $\mathcal{W}$ is a separable Banach space. Fix any $x_0 \in I$. Let $F_0 : I \rightarrow \mathcal{Y}$ be the cumulative distribution function of ν_F with origin x_0 , and $A_0 : I \rightarrow \mathcal{W}$ be the cumulative distribution function of ν_A with origin x_0 . By Proposition 12.75, F_0 and A_0 are absolutely continuous and therefore continuous. Since f and a are absolutely integrable over $\mathbb{I}$, then $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|F_0(x)\|_{\mathcal{Y}} \leq M$ and $\|A_0(x)\|_{\mathcal{W}} \leq M$, $\forall x \in I$. By the assumption, we have $F(x) = F(x_0) + F_0(x)$ and $A(x) = A(x_0) + A_0(x)$, $\forall x \in I$. Then, F and A are absolutely continuous, therefore continuous, and $\exists \bar{M} \in [0, \infty) \subset \mathbb{R}$ such that $\|F(x)\|_{\mathcal{Y}} \leq \bar{M}$ and $\|A(x)\|_{\mathcal{W}} \leq \bar{M}$, $\forall x \in I$. By Proposition 11.37, F and A are $\mathcal{B}$ -measurable. By Propositions 7.23, 7.65, 11.38, and 11.39, Af , aF and h are $\mathcal{B}$ -measurable. Then, Af , aF are absolutely integrable over $\mathbb{I}$. Hence, h is absolutely integrable over $\mathbb{I}$ by Proposition 11.83. By Proposition 12.68, H is absolutely continuous.

We will first consider the case that I is an open interval. By Fundamental Theorem of Calculus I 12.86, $F^{(1)} = D_1 F : U_F \rightarrow \mathcal{Y}$ and $A^{(1)} = D_1 A : U_A \rightarrow \mathcal{W}$ are such that $U_F, U_A \in \mathcal{B}$, $\mu(I \setminus U_F) = 0 = \mu(I \setminus U_A)$, $F^{(1)}$ and $A^{(1)}$ are $\mathcal{B}$ -measurable, $F^{(1)} = f$ a.e. in $\mathbb{I}$, and $A^{(1)} = a$ a.e. in $\mathbb{I}$. Then, $\exists \hat{U}_F, \hat{U}_A \in \mathcal{B}$ such that $\mu(I \setminus \hat{U}_F) = 0 = \mu(I \setminus \hat{U}_A)$ and $F^{(1)}(x) = f(x)$, $\forall x \in \hat{U}_F$, and $A^{(1)}(x) = a(x)$, $\forall x \in \hat{U}_A$.

$\forall x \in \hat{U}_F \cap \hat{U}_A$, by Propositions 9.17 and 9.19 and the Chain Rule, Theorem 9.18, we have $H^{(1)}(x) = A(x)f(x) + a(x)F(x) = h(x)$. By Fundamental Theorem of Calculus II 12.88, $H^{(1)} = D_1 H : U_H \rightarrow \mathcal{Z}$ is such that $U_H \in \mathcal{B}$, $\mu(I \setminus U_H) = 0$, and $H^{(1)}$ is $\mathcal{B}$ -measurable. Then, $E := (I \setminus U_H) \cup \{x \in U_H \mid H^{(1)}(x) \neq h(x)\} \in \mathcal{B}$ and $E \subseteq (I \setminus \hat{U}_F) \cup (I \setminus \hat{U}_A)$. Hence, $\mu(E) = 0$ and $h = H^{(1)}$ a.e. in $\mathbb{I}$. By Fundamental Theorem of Calculus II 12.88, $H(x) = \int_{x_0}^x \bar{h} d\mu$, $\forall x \in I$, $\bar{h}$ is $\mathcal{B}$ -measurable and absolutely integrable over r_{x_a, x_b} , $\forall x_a, x_b \in I$ with $x_a \leq x_b$, and $H^{(1)}(x) = \bar{h}(x)$ a.e. in $\mu(x)$. Thus, by Lemma 11.44, $h = \bar{h}$ a.e. in μ . Then, by Proposition 11.92, we have and $H(x) - H(x_0) = \int_{x_0}^x h(s) ds$, $\forall x \in I$. By Proposition 11.92 and Fact 12.72, $\forall b, c \in I$ with $b \leq c$, we have $A(c)F(c) - A(b)F(b) = H(c) - H(x_0) - (H(b) - H(x_0)) = \int_{x_0}^c h(x) dx - \int_{x_0}^b h(x) dx = \int_b^c (A(x)f(x) + a(x)F(x)) dx = \int_b^c A(x)f(x) dx + \int_b^c a(x)F(x) dx$.

Define μ_H to be the $\mathcal{Z}$ -valued measure with kernel h over $\mathbb{I}$. By the preceding paragraph, we have H is the cumulative distribution function of μ_H . This completes the proof for this case.

Next, we consider the case when I is a closed interval. We will define $\hat{I}$ to be an open interval such that $I \subseteq \hat{I}$. Define $\hat{f} : \hat{I} \rightarrow \mathcal{Y}$ by $\hat{f}(x) = \begin{cases} f(x) & x \in P(I) \\ \vartheta_{\mathcal{Y}} & x \in \hat{I} \setminus P(I) \end{cases}$. Similarly define $\hat{a}$ in terms of a . Define $\hat{F} : \hat{I} \rightarrow \mathcal{Y}$ by $\hat{F}(x) = \begin{cases} F(x) & x \in I \\ F(\min_{s \in I} s) & x < \min_{s \in I} s \\ F(\max_{s \in I} s) & x > \max_{s \in I} s \end{cases}$. Similarly define $\hat{A}$ in terms of A .

Then, the result can be applied for $\hat{\mathbb{I}} := ((\hat{I}, |\cdot|), \hat{\mathcal{B}}, \hat{\mu})$ for the functions $\hat{f}, \hat{F}, \hat{a}, \hat{A}$, to yield $\hat{H} : \hat{I} \rightarrow \mathcal{Z}$ and $\hat{h} : \hat{I} = P(\hat{I}) \rightarrow \mathcal{Z}$. Restrict these functions to I to yield the desired result.

This completes the proof of the theorem. $\square$

12.7 Surface measures

12.8 Representation of $(\mathcal{C}_k(\Omega, \mathcal{Y}))^*$

Theorem 12.90 (Riesz Representation Theorem) *Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $\mathcal{Y}$ be Banach space over $\mathbb{K}$, $\Omega \subseteq \mathcal{X}$ be compact with the subset topology $\mathcal{O}$, and $k \in \mathbb{N}$. Assume that, $\forall x \in \Omega$, $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$ such that $\Omega \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x) - x$ is a conic segment and $\overline{\text{span}(\mathcal{A}_{\Omega}(x))} = \mathcal{X}$. Let $\mathcal{Z} := \mathcal{C}_k(\Omega, \mathcal{Y})$ be the Banach space over $\mathbb{K}$ as prescribed in Example 9.70.*

$\forall f \in \mathcal{Z}^*, \exists \mu := (\mu_0, \dots, \mu_k) \in \prod_{i=0}^k (\mathcal{C}(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y})))^*$ such that $f(z) = \sum_{i=0}^k \langle \mu_i, z^{(i)} \rangle, \forall z \in \mathcal{Z}$, with $\|\mu\| = \left(\sum_{i=0}^k \|\mu_i\|^2 \right)^{1/2} = \|f\|_{\mathcal{Z}^*}$. Furthermore, $\forall \nu := (\nu_0, \dots, \nu_k) \in \prod_{i=0}^k (\mathcal{C}(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y})))^*$, define $h : \mathcal{Z} \rightarrow \mathbb{K}$ by $h(z) = \sum_{i=0}^k \langle \nu_i, z^{(i)} \rangle, \forall z \in \mathcal{Z}$. Then, $h \in \mathcal{Z}^*$ and $\|h\|_{\mathcal{Z}^*} \leq \|\nu\|$.

Proof Let $\mathcal{W} := \prod_{i=0}^k \mathcal{C}(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y})) =: \prod_{i=0}^k \mathcal{W}_i$. By Proposition 7.66 and Example 7.32, $\mathcal{W}_i$ is a Banach space over $\mathbb{K}$, $\forall i \in \{0, \dots, k\}$. Then, by Propositions 7.22 and 4.31, $\mathcal{W}$ is a Banach space over $\mathbb{K}$. $\forall z \in \mathcal{Z}$, $\forall i \in \{0, \dots, k\}$, we have $z^{(i)} \in \mathcal{W}_i$. Then, we may define a mapping $T : \mathcal{Z} \rightarrow \mathcal{W}$ by $T(z) = (z^{(0)}, \dots, z^{(k)})$, $\forall z \in \mathcal{Z}$. Clearly, T is linear and norm-preserving. Hence, $T \in \mathcal{B}(\mathcal{Z}, \mathcal{W})$ and is injective.

Claim 12.90.1 $\mathcal{R}(T) \subseteq \mathcal{W}$ is closed.

Proof of claim: Clearly, we have $\mathcal{R}(T) \subseteq \overline{\mathcal{R}(T)} \subseteq \mathcal{W}$. $\forall w \in \overline{\mathcal{R}(T)}$, $\exists (z_n)_{n=1}^\infty \subseteq \mathcal{Z}$ such that $w = \lim_{n \in \mathbb{N}} T(z_n)$ in $\mathcal{W}$. Then, $(T(z_n))_{n=1}^\infty \subseteq \mathcal{W}$ is a Cauchy sequence. Since T is norm-preserving, then $(z_n)_{n=1}^\infty \subseteq \mathcal{Z}$ is a Cauchy sequence. By the completeness of $\mathcal{Z}$, $\exists z_0 \in \mathcal{Z}$ such that $\lim_{n \in \mathbb{N}} z_n = z_0$ in $\mathcal{Z}$. By the continuity of T , we have $w = \lim_{n \in \mathbb{N}} T(z_n) = T(z_0) \in \mathcal{W}$. Hence, $w \in \mathcal{R}(T)$. By the arbitrariness of w , we have $\overline{\mathcal{R}(T)} \subseteq \mathcal{R}(T)$. Hence, $\mathcal{R}(T)$ is closed. This completes the proof of the claim. $\square$

By Claim 12.90.1 and Proposition 4.39, $\mathcal{R}(T) \subseteq \mathcal{W}$ is a Banach space over $\mathbb{K}$. By Open Mapping Theorem 7.103, $T_{\text{inv}} : \mathcal{R}(T) \rightarrow \mathcal{Z}$ and $T_{\text{inv}} \in \mathcal{B}(\mathcal{R}(T), \mathcal{Z})$.

$\forall f \in \mathcal{Z}^*$, define $g : \mathcal{R}(T) \rightarrow \mathbb{K}$ by $g(w) = f(T_{\text{inv}}(w))$, $\forall w \in \mathcal{R}(T)$. Clearly, g is a bounded linear functional on $\mathcal{R}(T)$. By Hahn-Banach Extension Theorem 7.83, $\exists \mu \in \mathcal{W}^*$ such that $\mu(w) = g(w)$, $\forall w \in \mathcal{R}(T)$, and $\|\mu\|_{\mathcal{W}^*} = \|g\|_{\mathcal{R}(T)} \leq \|f\|_{\mathcal{Z}^*} \|T_{\text{inv}}\|_{\mathcal{B}(\mathcal{R}(T), \mathcal{Z})} \leq \|f\|_{\mathcal{Z}^*}$, where the last inequality follows from the fact that T is norm-preserving and so is T_{inv} . By Example 7.76, $\mathcal{W}^* = \prod_{i=0}^k \mathcal{W}_i^* = \prod_{i=0}^k (\mathcal{C}(\Omega, \mathcal{B}_i(\mathcal{X}, \mathcal{Y})))^*$. Then, $\mu = (\mu_0, \dots, \mu_k) \in \mathcal{W}^*$ and $f(z) = g(T(z)) = \langle \mu, T(z) \rangle = \sum_{i=0}^k \langle \mu_i, z^{(i)} \rangle$, $\forall z \in \mathcal{Z}$.

$\forall \nu := (\nu_0, \dots, \nu_k) \in \mathcal{W}^*$, define $h : \mathcal{Z} \rightarrow \mathbb{K}$ by $h(z) = \sum_{i=0}^k \langle \nu_i, z^{(i)} \rangle$, $\forall z \in \mathcal{Z}$. Then, $h(z) = \langle \nu, T(z) \rangle$, $\forall z \in \mathcal{Z}$. Then, $h \in \mathcal{Z}^*$ and $\|h\|_{\mathcal{Z}^*} \leq \|\nu\|_{\mathcal{W}^*} \|T\|_{\mathcal{B}(\mathcal{Z}, \mathcal{W})} \leq \|\nu\|_{\mathcal{W}^*}$, where the last inequality follows from the fact that T is norm-preserving.

Then, $\|f\|_{\mathcal{Z}^*} \leq \|\mu\|_{\mathcal{W}^*} \leq \|f\|_{\mathcal{Z}^*}$. Hence, $\|f\|_{\mathcal{Z}^*} = \|\mu\|_{\mathcal{W}^*}$. This completes the proof of the theorem. $\square$

Theorem 12.91 (Riesz Representation Theorem) Let $m \in \mathbb{Z}_+$, $\mathcal{X} := \mathbb{K}^m$ be the Euclidean space, $I_1, \dots, I_m \subset \mathbb{R}$ be nondegenerate compact intervals, $\Omega := \prod_{j=1}^m I_j \subset \mathcal{X}$ with the subset topology $\mathcal{O}$, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $k \in \mathbb{N}$, and $\mathcal{Z} := \mathcal{C}_k(\Omega, \mathcal{Y})$ be the Banach space over $\mathbb{K}$ as prescribed in Example 9.70. $\forall f \in \mathcal{Z}^*, \exists \mu := (\mu_0, \dots, \mu_k) \in \prod_{i=0}^k \mathcal{M}_{ft}(\Omega, (\mathcal{B}_i(\mathcal{X}, \mathcal{Y}))^*) =$

$\prod_{i=0}^k \mathcal{M}_f(\Omega, \mathcal{B}_B(\Omega), (B_i(\mathcal{X}, \mathcal{Y}))^*)$ such that $f(z) = \sum_{i=0}^k \int_{\Omega} \langle \langle d\mu_i, z^{(i)} \rangle \rangle$, $\forall z \in \mathcal{Z}$, with $\|\mu\| = \left(\sum_{i=0}^k (\mathcal{P} \circ \mu_i(\Omega))^2 \right)^{1/2} = \|f\|_{\mathcal{Z}^*}$. Furthermore, $\forall \nu := (\nu_0, \dots, \nu_k) \in \prod_{i=0}^k \mathcal{M}_{ft}(\Omega, (B_i(\mathcal{X}, \mathcal{Y}))^*)$, define $h : \mathcal{Z} \rightarrow \mathbb{K}$ by $h(z) = \sum_{i=0}^k \int_{\Omega} \langle \langle d\nu_i, z^{(i)} \rangle \rangle$, $\forall z \in \mathcal{Z}$. Then, $h \in \mathcal{Z}^*$ and $\|h\|_{\mathcal{Z}^*} \leq \|\nu\|$.

Proof Let $\mathcal{W} := \prod_{i=0}^k \mathcal{C}(\Omega, B_i(\mathcal{X}, \mathcal{Y})) =: \prod_{i=0}^k \mathcal{W}_i$. By Proposition 7.66 and Example 9.70, $\mathcal{W}_i$ is a Banach space over $\mathbb{K}$, $\forall i \in \{0, \dots, k\}$. It is easy to show that $\Omega \subseteq \mathcal{X}$ satisfies the assumption of Theorem 12.90. By Riesz Representation Theorem 11.204, we have $\mathcal{W}^* = \prod_{i=0}^k (\mathcal{C}(\Omega, B_i(\mathcal{X}, \mathcal{Y})))^* = \prod_{i=0}^k \mathcal{M}_{ft}(\Omega, (B_i(\mathcal{X}, \mathcal{Y}))^*) = \prod_{i=0}^k \mathcal{M}_f(\Omega, \mathcal{B}_B(\Omega), (B_i(\mathcal{X}, \mathcal{Y}))^*)$. Then, the result follows immediately from Theorem 12.90. This completes the proof of the theorem. $\square$

Theorem 12.92 (Riesz Representation Theorem) Let $\mathcal{X}$ and $\mathcal{Y}$ be finite-dimensional Banach spaces over $\mathbb{K}$, $\Omega \subseteq \mathcal{X}$ be compact with the subset topology $\mathcal{O}$, and $k \in \mathbb{N}$. Assume that, $\forall x \in \Omega$, $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$ such that $\Omega \cap \mathcal{B}_{\mathcal{X}}(x, \delta_x) - x$ is a conic segment and $\text{span}(\overline{A_{\Omega}(x)}) = \mathcal{X}$. Let $\mathcal{Z} := \mathcal{C}_k(\Omega, \mathcal{Y})$ be the Banach space over $\mathbb{K}$ as prescribed in Example 9.70. $\forall f \in \mathcal{Z}^*$, $\exists \mu := (\mu_0, \dots, \mu_k) \in \prod_{i=0}^k \mathcal{M}_{ft}(\Omega, (B_i(\mathcal{X}, \mathcal{Y}))^*) = \prod_{i=0}^k \mathcal{M}_f(\Omega, \mathcal{B}_B(\Omega), (B_i(\mathcal{X}, \mathcal{Y}))^*)$ such that $f(z) = \sum_{i=0}^k \int_{\Omega} \langle \langle d\mu_i, z^{(i)} \rangle \rangle$, $\forall z \in \mathcal{Z}$, with $\|\mu\| = \left(\sum_{i=0}^k (\mathcal{P} \circ \mu_i(\Omega))^2 \right)^{1/2} = \|f\|_{\mathcal{Z}^*}$. Furthermore, $\forall \nu := (\nu_0, \dots, \nu_k) \in \prod_{i=0}^k \mathcal{M}_{ft}(\Omega, (B_i(\mathcal{X}, \mathcal{Y}))^*)$, define $h : \mathcal{Z} \rightarrow \mathbb{K}$ by $h(z) = \sum_{i=0}^k \int_{\Omega} \langle \langle d\nu_i, z^{(i)} \rangle \rangle$, $\forall z \in \mathcal{Z}$. Then, $h \in \mathcal{Z}^*$ and $\|h\|_{\mathcal{Z}^*} \leq \|\nu\|$.

Proof Let $\mathcal{W} := \prod_{i=0}^k \mathcal{C}(\Omega, B_i(\mathcal{X}, \mathcal{Y})) =: \prod_{i=0}^k \mathcal{W}_i$. The result follows immediately from Theorem 12.90 if we can prove that $\mathcal{W}^* = \prod_{i=0}^k \mathcal{M}_{ft}(\Omega, (B_i(\mathcal{X}, \mathcal{Y}))^*) = \prod_{i=0}^k \mathcal{M}_f(\Omega, \mathcal{B}_B(\Omega), (B_i(\mathcal{X}, \mathcal{Y}))^*)$. The first equality follows from Example 7.76 and Riesz Representation Theorem 11.205. Since $\Omega \subseteq \mathcal{X}$ is compact then it is locally compact. Since $\mathcal{X}$ is a finite dimensional Banach space, then it is separable by Proposition 11.190. By Propositions 4.37 and 4.38, Ω is separable metric space. By Theorem 11.198, we have $\mathcal{M}_{ft}(\Omega, (B_i(\mathcal{X}, \mathcal{Y}))^*) = \mathcal{M}_f(\Omega, \mathcal{B}_B(\Omega), (B_i(\mathcal{X}, \mathcal{Y}))^*)$, $\forall i \in \{0, \dots, k\}$. This completes the proof of the theorem. $\square$

Chapter 13

Hilbert Spaces

Hilbert spaces are vector spaces equipped with inner products. They possess a wealth of structural properties generalizing basic geometrical insights. The concepts of orthonormal basis, Fourier series, and least-square minimization all have natural setting in Hilbertspaces.

13.1 Fundamental Notions

Definition 13.1 A pre-Hilbert space $\mathcal{X}$ over $\mathbb{K}$ is a vector space $\mathcal{X}$ over $\mathbb{K}$, together with an inner product $\langle \cdot, \cdot \rangle : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{K}$ satisfying, $\forall x, y, z \in \mathcal{X}$, $\forall \lambda \in \mathbb{K}$,

- (i) $\langle x, y \rangle = \overline{\langle y, x \rangle}$;
- (ii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$;
- (iii) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$;
- (iv) $\langle x, x \rangle \in [0, \infty) \subset \mathbb{R}$ and $\langle x, x \rangle = 0$ if, and only if, $x = \vartheta_{\mathcal{X}}$.

Proposition 13.2 Let $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$. Define $\|\cdot\| : \mathcal{X} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\|x\| = \sqrt{\langle x, x \rangle}$, $\forall x \in \mathcal{X}$. Then, $\forall x, y, z \in \mathcal{X}$, $\forall \alpha, \beta \in \mathbb{K}$, we have

- (i) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$;
- (ii) $\langle z, \alpha x + \beta y \rangle = \overline{\alpha} \langle z, x \rangle + \overline{\beta} \langle z, y \rangle$;
- (iii) $\langle \vartheta_{\mathcal{X}}, x \rangle = 0 = \langle x, \vartheta_{\mathcal{X}} \rangle$;
- (iv) $|\langle x, y \rangle| \leq \|x\| \|y\|$, where equality holds if, and only if, $y = \vartheta_{\mathcal{X}}$ or $x = \lambda y$ for some $\lambda \in \mathbb{K}$; (Cauchy-Schwarz Inequality)
- (v) $\|x + y\| \leq \|x\| + \|y\|$, where equality holds if, and only if, $y = \vartheta_{\mathcal{X}}$ or $x = \lambda y$, for some $\lambda \in [0, \infty) \subset \mathbb{R}$.

(vi) $\|\cdot\|$ is a norm of $\mathcal{X}$. (Induced norm)

Proof (i) $\langle \alpha x + \beta y, z \rangle = \langle \alpha x, z \rangle + \langle \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$, where the first equality follows from (ii) of Definition 13.1; and the second equality follows from (iii) of Definition 13.1.

(ii) $\langle z, \alpha x + \beta y \rangle = \overline{\langle \alpha x + \beta y, z \rangle} = \overline{\alpha \langle x, z \rangle + \beta \langle y, z \rangle} = \overline{\alpha} \overline{\langle x, z \rangle} + \overline{\beta} \overline{\langle y, z \rangle} = \overline{\alpha} \langle z, x \rangle + \overline{\beta} \langle z, y \rangle$, where the first equality follows from (i) of Definition 13.1; the second equality follows from (i); and the fourth equality follows from (i) of Definition 13.1.

(iii) $\langle \vartheta_{\mathcal{X}}, x \rangle = \langle y - y, x \rangle = \langle y, x \rangle - \langle y, x \rangle = 0$, where the second equality follows from (i). Then, $\langle x, \vartheta_{\mathcal{X}} \rangle = \overline{\langle \vartheta_{\mathcal{X}}, x \rangle} = 0$.

(iv) We will distinguish two exhaustive and mutually exclusives cases: Case 1: $y = \vartheta_{\mathcal{X}}$; Case 2: $y \neq \vartheta_{\mathcal{X}}$. Case 1: $y = \vartheta_{\mathcal{X}}$. Then, $|\langle x, y \rangle| = 0 = \|x\| \cdot 0 = \|x\| \|y\|$. This case is proved.

Case 2: $y \neq \vartheta_{\mathcal{X}}$. By (iv) of Definition 13.1, we have $\langle y, y \rangle > 0$. $\forall \lambda \in \mathbb{K}$, we have $0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x - \lambda y \rangle - \lambda \langle y, x - \lambda y \rangle = \langle x, x \rangle - \overline{\lambda} \langle x, y \rangle - \lambda \langle y, x \rangle + \lambda \overline{\lambda} \langle y, y \rangle$, where the inequality follows from (iv) of Definition 13.1; the first equality follows from (i); and the second equality follows from (ii). Take $\lambda := \langle x, y \rangle / \langle y, y \rangle$. Then, $0 \leq \langle x, x \rangle - \overline{\lambda} \langle x, y \rangle - \lambda \overline{\langle x, y \rangle} + |\lambda|^2 \langle y, y \rangle = \langle x, x \rangle - \langle x, y \rangle \overline{\langle x, y \rangle} / \langle y, y \rangle = \langle x, x \rangle - |\langle x, y \rangle|^2 / \langle y, y \rangle$, where the equality follows (i) of Definition 13.1. Hence, $|\langle x, y \rangle| \leq \|x\| \|y\|$. Equality holds if, and only if, $x - \lambda y = \vartheta_{\mathcal{X}}$. This case is also proved. Hence, (iv) holds.

(v) Note that $\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + 2 \operatorname{Re}(\langle x, y \rangle) + \|y\|^2 \leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$, where the first equality follows from (i) and (ii); the second equality follows from (i) of Definition 13.1; the second inequality follows from (iv). Hence, $\|x + y\| \leq \|x\| + \|y\|$. Equality holds if, and only if, $\operatorname{Re}(\langle x, y \rangle) = |\langle x, y \rangle| = \|x\| \|y\|$. By (iv), have equality holds if, and only if, $y = \vartheta_{\mathcal{X}}$ or $x = \lambda y$ for some $\lambda \in [0, \infty) \subset \mathbb{R}$.

(vi) Clearly, $\|x\| \in [0, \infty) \subset \mathbb{R}$. $\|x\| = 0 \Leftrightarrow \langle x, x \rangle = 0 \Leftrightarrow x = \vartheta_{\mathcal{X}}$. Hence, (i) of Definition 7.1 holds. By (v), (ii) of Definition 7.1 holds. Note also that $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{\alpha \overline{\alpha} \langle x, x \rangle} = |\alpha| \|x\|$, where the second equality follows from (i) and (ii); and the third equality follows from the fact that $\alpha \overline{\alpha} = |\alpha|^2$. Hence, (iii) of Definition 7.1 holds. Therefore, $\|\cdot\|$ defines a norm of $\mathcal{X}$. $\square$

Lemma 13.3 (Parallelogram Law) *Let $\mathcal{X}$ be a pre-Hilbert space. $\forall x, y \in \mathcal{X}$, we have $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$.*

Proof This is straightforward, and is therefore omitted. $\square$

Pre-Hilbert space is a special kind of normed linear space. The concept of convergence, continuity, topology, and completeness, etc, apply in pre-Hilbert spaces.

Proposition 13.4 *Let $\mathcal{X}$ be a pre-Hilbert space. Then, the inner product is a continuous function of the product space $\mathcal{X} \times \mathcal{X}$.*

Proof We will show that the inner product is continuous at any $(x_0, y_0) \in \mathcal{X} \times \mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $M := \max\{\|x_0\|, \|y_0\|\} + 1 \in (0, \infty) \subset \mathbb{R}$ and $\delta := \min\{\epsilon, 1\}/M \in (0, 1] \subset \mathbb{R}$. Fix any $(x, y) \in \mathcal{B}_{\mathcal{X} \times \mathcal{X}}((x_0, y_0), \delta)$, we have $|\langle x, y \rangle - \langle x_0, y_0 \rangle| = |\langle x - x_0, y \rangle + \langle x_0, y - y_0 \rangle| \leq |\langle x - x_0, y \rangle| + |\langle x_0, y - y_0 \rangle| \leq \|x - x_0\| \|y\| + \|x_0\| \|y - y_0\| \leq \|x - x_0\| (\|y_0\| + \|y - y_0\|) + M \|y - y_0\| \leq M (\|x - x_0\| + \|y - y_0\|) < M\delta \leq \epsilon$, where the second inequality follows from Proposition 13.2. Hence, the inner product is continuous at (x_0, y_0) . By the arbitrariness of (x_0, y_0) , we have the inner product is continuous. This completes the proof of the proposition. $\square$

Definition 13.5 *A complete pre-Hilbert space is called a Hilbert space.*

Definition 13.6 *Let $\mathcal{X}$ be a pre-Hilbert space, $x, y \in \mathcal{X}$, and $S \subseteq \mathcal{X}$. The vectors x and y are said to be orthogonal if $\langle x, y \rangle = 0$. We will then write $x \perp y$. The vector x is said to be orthogonal to the set S if $x \perp s$, $\forall s \in S$. We will then write $x \perp S$.*

Theorem 13.7 (Pythagorean) *Let $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$ and $x, y \in \mathcal{X}$. If $x \perp y$, then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. Furthermore, if $\mathbb{K} = \mathbb{R}$ and $\|x + y\|^2 = \|x\|^2 + \|y\|^2$, then $x \perp y$.*

Proof Let $x \perp y$, then $\langle x, y \rangle = 0$. This implies that $\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2$, where the second equality follows from Proposition 13.2; and the third equality follows from Definition 13.1.

Let $\mathbb{K} = \mathbb{R}$ and $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. Then, we have $0 = \langle x, y \rangle + \langle y, x \rangle = 2\langle x, y \rangle$, where the first equality follows from Proposition 13.2; and the second equality follows from Definition 13.1 and $\mathbb{K} = \mathbb{R}$. Hence, $x \perp y$. This completes the proof of the theorem. $\square$

Example 13.8 Let $n \in \mathbb{Z}_+$ and $\mathcal{X} = \mathbb{K}^n$. Define the inner product by $\langle x, y \rangle = \sum_{i=1}^n \xi_i \overline{\eta_i} \in \mathbb{K}$, $\forall x := (\xi_1, \dots, \xi_n) \in \mathbb{K}^n$ and $\forall y := (\eta_1, \dots, \eta_n) \in \mathbb{K}^n$. It is easy to check that the pair $\mathcal{X} := (\mathcal{X}, \langle \cdot, \cdot \rangle)$ is a pre-Hilbert space over $\mathbb{K}$. The induced norm is clearly the Euclidean norm. By Example 7.29, $\mathcal{X}$ is complete. Therefore, $\mathcal{X}$ is a Hilbert space, which is to be denoted simply by $\mathbb{K}^n$. $\diamond$

Example 13.9 Let $\mathcal{X}_i$ be a Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle_i$, $i = 1, 2$. Let $\mathcal{Z} := \mathcal{X}_1 \times \mathcal{X}_2$. Define inner product $\langle \cdot, \cdot \rangle_{\mathcal{Z}} : \mathcal{Z} \times \mathcal{Z} \rightarrow \mathbb{K}$ by $\langle (x_1, x_2), (y_1, y_2) \rangle_{\mathcal{Z}} = \langle x_1, y_1 \rangle_1 + \langle x_2, y_2 \rangle_2$, $\forall (x_1, x_2), (y_1, y_2) \in \mathcal{Z}$. It is easy to show that $\mathcal{Z}$ with this inner product is a pre-Hilbertspace over $\mathbb{K}$. It is also easy to see that the induced norm of $\mathcal{Z}$ is equal to the norm of $\mathcal{Z}$ as defined in Proposition 7.22. By Proposition 4.31, $\mathcal{Z}$ is complete. Therefore, $\mathcal{Z}$ is a Hilbert space over $\mathbb{K}$. $\diamond$

The above example can be easily generalized to $\prod_{i=1}^n \mathcal{X}_i$ case, where $n \in \mathbb{Z}_+$ and $\mathcal{X}_i$ is a Hilbert space over $\mathbb{K}$, $\forall i \in \{1, \dots, n\}$.

Example 13.10 Let $\mathcal{X}$ be a Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{X}}$ and $\mathcal{Z} := l_2(\mathcal{X})$. We will show that $\mathcal{Z}$ is a Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ defined by $\langle x, y \rangle_{\mathcal{Z}} = \sum_{i=1}^{\infty} \langle \xi_i, \eta_i \rangle$, $\forall x := (\xi_1, \xi_2, \dots) \in \mathcal{Z}$ and $\forall y := (\eta_1, \eta_2, \dots) \in \mathcal{Z}$. To see that $\langle x, y \rangle_{\mathcal{Z}} \in \mathbb{K}$, we note that $x \in \mathcal{Z}$ implies that $\sum_{i=1}^{\infty} \|\xi_i\|^2 < \infty$. Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. Then, $\exists n_1 \in \mathbb{N}$ such that $\sum_{i=n_1}^{\infty} \|\xi_i\|^2 < \epsilon$. By $y \in \mathcal{Z}$, $\exists n_2 \in \mathbb{N}$ such that $\sum_{i=n_2}^{\infty} \|\eta_i\|^2 < \epsilon$. Let $n_0 := \max\{n_1, n_2\} \in \mathbb{N}$. $\forall n, m \in \mathbb{N}$ with $n_0 \leq n \leq m$, we have $|\sum_{i=n}^m \langle \xi_i, \eta_i \rangle| \leq (\sum_{i=n}^m \|\xi_i\|^2)^{1/2} (\sum_{i=n}^m \|\eta_i\|^2)^{1/2} \leq (\sum_{i=n_1}^{\infty} \|\xi_i\|^2)^{1/2} (\sum_{i=n_2}^{\infty} \|\eta_i\|^2)^{1/2} < \epsilon$, where the first inequality follows from Hölder's Inequality 7.8. Hence, $(\sum_{i=1}^n \langle \xi_i, \eta_i \rangle)_{n=1}^{\infty} \subseteq \mathbb{K}$ is a Cauchy sequence, which converges since $\mathbb{K}$ is complete. Hence, $\langle x, y \rangle_{\mathcal{Z}} \in \mathbb{K}$. It is easy to show that (i) – (iv) of Definition 13.1 are satisfied. Hence, $\mathcal{Z}$ is a pre-Hilbert space. It is easy to see that the induced norm is equal to the norm of $\mathcal{Z}$ as defined in Example 7.10. By Example 7.33, we have $\mathcal{Z}$ is complete and therefore a Hilbert space. $\diamond$

Example 13.11 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a σ -finite measure space, $\mathcal{Y}$ be a separable Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{Y}}$ and $\mathcal{Z} := L_2(\mathcal{X}, \mathcal{Y})$. We will show that $\mathcal{Z}$ is a Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ defined by $\langle [z_1], [z_2] \rangle_{\mathcal{Z}} = \int_X \langle z_1(x), z_2(x) \rangle_{\mathcal{Y}} d\mu(x)$, $\forall z_1, z_2 \in \bar{L}_2(\mathcal{X}, \mathcal{Y})$.

First, we show that $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ is well-defined. $\forall z_1, z_2 \in \bar{L}_2(\mathcal{X}, \mathcal{Y})$, we have z_1 and z_2 are $\mathcal{B}$ -measurable. By Propositions 11.39, 11.38, and 13.4, the function $h : \mathcal{X} \rightarrow \mathbb{K}$ defined by $h(x) = \langle z_1(x), z_2(x) \rangle_{\mathcal{Y}}$, $\forall x \in \mathcal{X}$, is $\mathcal{B}$ -measurable. $\mathcal{P} \circ h(x) = |h(x)| \leq \|z_1(x)\| \|z_2(x)\| = \mathcal{P} \circ z_1(x) \mathcal{P} \circ z_2(x)$, $\forall x \in \mathcal{X}$, where the inequality follows from Proposition 13.2. Then, $\int_X \mathcal{P} \circ h d\mu \leq \int_X (\mathcal{P} \circ z_1 \mathcal{P} \circ z_2) d\mu \leq (\int_X \mathcal{P}_2 \circ z_1 d\mu)^{1/2} (\int_X \mathcal{P}_2 \circ z_2 d\mu)^{1/2} = \|z_1\|_2 \|z_2\|_2 < \infty$, where the first inequality follows from Definition 11.79; and the second inequality follows from Hölder's Inequality 11.178. Hence, h is absolute integrable over $\mathcal{X}$. By Proposition 11.92, h is integrable over $\mathcal{X}$ and $\int_X \langle z_1(x), z_2(x) \rangle_{\mathcal{Y}} d\mu(x) \in \mathbb{K}$. $\forall \bar{z}_1, \bar{z}_2 \in \bar{L}_2(\mathcal{X}, \mathcal{Y})$ with $[z_1] = [\bar{z}_1]$ and $[z_2] = [\bar{z}_2]$, by Example 11.173, we have $z_1 = \bar{z}_1$ a.e. in $\mathcal{X}$ and $z_2 = \bar{z}_2$ a.e. in $\mathcal{X}$. Define $\bar{h} : \mathcal{X} \rightarrow \mathbb{K}$ by $\bar{h}(x) = \langle \bar{z}_1(x), \bar{z}_2(x) \rangle_{\mathcal{Y}}$, $\forall x \in \mathcal{X}$. By Lemmas 11.45 and 11.46, $h = \bar{h}$ a.e. in $\mathcal{X}$. By Proposition 11.92, $\int_X h d\mu = \int_X \bar{h} d\mu$. Hence, $\langle [z_1], [z_2] \rangle_{\mathcal{Z}} \in \mathbb{K}$ is well-defined.

Next, we show that $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ satisfies (i) – (iv) of Definition 13.1. Fix any $z_1, z_2, z_3 \in \bar{L}_2(\mathcal{X}, \mathcal{Y})$ and any $\lambda \in \mathbb{K}$. (i) $\langle [z_1], [z_2] \rangle_{\mathcal{Z}} = \int_X \langle z_1(x), z_2(x) \rangle_{\mathcal{Y}} d\mu(x) = \int_X \langle z_2(x), z_1(x) \rangle_{\mathcal{Y}} d\mu(x) = \int_X \langle z_2(x), z_1(x) \rangle_{\mathcal{Y}} d\mu(x) = \langle [z_2], [z_1] \rangle_{\mathcal{Z}}$, where the second equality follows from the fact that $\mathcal{Y}$ is a Hilbert space over $\mathbb{K}$; and the third equality follows from Proposition 11.92. (ii) $\langle [z_1] + [z_2], [z_3] \rangle_{\mathcal{Z}} = \langle [z_1 + z_2], [z_3] \rangle_{\mathcal{Z}} = \int_X \langle z_1(x) + z_2(x), z_3(x) \rangle_{\mathcal{Y}} d\mu(x) = \int_X (\langle z_1(x), z_3(x) \rangle_{\mathcal{Y}} +$

$\langle z_2(x), z_3(x) \rangle_{\mathcal{Y}} d\mu(x) = \int_X \langle z_1(x), z_3(x) \rangle_{\mathcal{Y}} d\mu(x) + \int_X \langle z_2(x), z_3(x) \rangle_{\mathcal{Y}} d\mu(x) = \langle [z_1], [z_3] \rangle_{\mathcal{Z}} + \langle [z_2], [z_3] \rangle_{\mathcal{Z}}$, where the first equality follows from Proposition 7.43; the third equality follows from the fact that $\mathcal{Y}$ is a Hilbert space; and the fourth equality follows from Proposition 11.92. (iii) $\langle \lambda[z_1], [z_2] \rangle_{\mathcal{Z}} = \langle [\lambda z_1], [z_2] \rangle_{\mathcal{Z}} = \int_X \langle \lambda z_1(x), z_2(x) \rangle_{\mathcal{Y}} d\mu(x) = \int_X \lambda \langle z_1(x), z_2(x) \rangle_{\mathcal{Y}} d\mu(x) = \lambda \int_X \langle z_1(x), z_2(x) \rangle_{\mathcal{Y}} d\mu(x) = \lambda \langle [z_1], [z_2] \rangle_{\mathcal{Z}}$, where the first equality follows from Proposition 7.43; the third equality follows from the fact that $\mathcal{Y}$ is a Hilbert space; and the fourth equality follows from Proposition 11.92. (iv) $\langle [z_1], [z_1] \rangle_{\mathcal{Z}} = \int_X \langle z_1(x), z_1(x) \rangle_{\mathcal{Y}} d\mu(x) = \int_X \|z_1(x)\|_{\mathcal{Y}}^2 d\mu(x) = \|z_1\|_2^2 \in [0, \infty) \subset \mathbb{R}$. Clearly, $\langle [z_1], [z_1] \rangle_{\mathcal{Z}} = 0 \Leftrightarrow \|z_1\|_2 = 0 \Leftrightarrow [z_1] = \vartheta_{\mathcal{Z}}$. Hence, $\mathcal{Z}$ is a pre-Hilbert space over $\mathbb{K}$.

Note that the induced norm of $\mathcal{Z}$ is equal to the norm of $\mathcal{Z}$ as defined in Example 11.173. By Example 11.179, $\mathcal{Z}$ is complete. Hence, $\mathcal{Z}$ is a Hilbert space over $\mathbb{K}$. $\diamond$

13.2 Projection Theorems

The basic concept of the Projection Theorem is illustrated in the Figure 13.1.

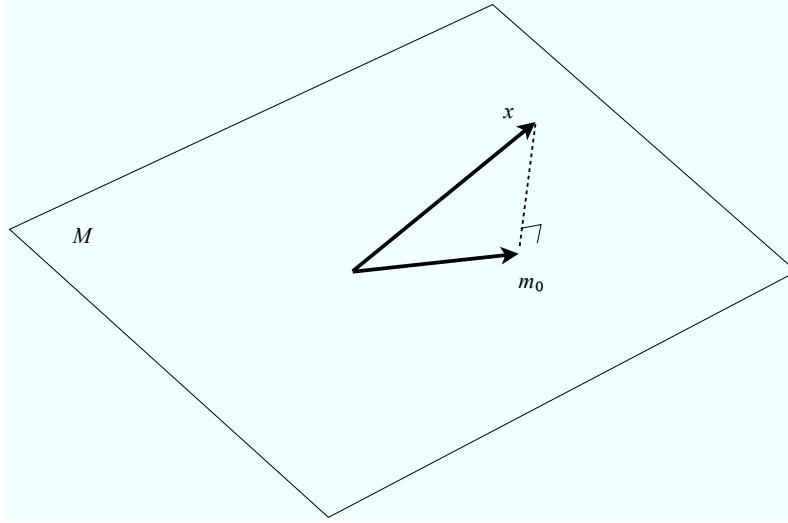


Figure 13.1: Projection onto a subspace.

Theorem 13.12 *Let $\mathcal{X}$ be a pre-Hilbert space, $M \subseteq \mathcal{X}$ be a subspace, and $x \in \mathcal{X}$. Consider the problem $\min_{m \in M} \|x - m\|$. If there exists $m_0 \in M$ such that $\|x - m_0\| \leq \|x - m\|$, $\forall m \in M$, then m_0 is the unique vector*

in M that minimizes $\|x - m\|$. A necessary and sufficient condition for $m_0 \in M$ being the unique minimizing vector is $(x - m_0) \perp M$.

Proof We first show that if $m_0 \in M$ is a minimizing vector then $(x - m_0) \perp M$ by an argument of contradiction. Suppose that $(x - m_0) \not\perp M$. Then, $\exists m \in M$ such that $\langle x - m_0, m \rangle =: \delta \neq 0$. Without loss of generality, assume $\|m\| = 1$. Define $m_1 := m_0 + \delta m \in M$. Then, by Proposition 13.2, $\|x - m_1\|^2 = \langle x - m_0 - \delta m, x - m_0 - \delta m \rangle = \|x - m_0\|^2 - \langle x - m_0, \delta m \rangle - \langle \delta m, x - m_0 \rangle + \|\delta m\|^2 = \|x - m_0\|^2 - \delta \bar{\delta} - \delta \bar{\delta} + |\delta|^2 \|m\|^2 = \|x - m_0\|^2 - |\delta|^2 < \|x - m_0\|^2$. This contradicts the assumption that m_0 is a minimizing vector. Hence, $(x - m_0) \perp M$.

Next, we show that if $(x - m_0) \perp M$, where $m_0 \in M$, then m_0 is the unique minimizing vector. $\forall m \in M$, by Pythagorean Theorem 13.7, we have $\|x - m\|^2 = \|x - m_0 - (m - m_0)\|^2 = \|x - m_0\|^2 + \|m - m_0\|^2$. Then, $\|x - m\| > \|x - m_0\|$ if $m \neq m_0$. This completes the proof of the theorem. $\square$

Theorem 13.13 (The Classical Projection Theorem) *Let $\mathcal{X}$ be a Hilbert space and $M \subseteq \mathcal{X}$ be a closed subspace. $\forall x \in \mathcal{X}$, there exists a unique vector $m_0 \in M$ such that $\|x - m_0\| = \min_{m \in M} \|x - m\|$. Furthermore, a necessary and sufficient condition for $m_0 \in M$ being the unique minimizing vector is $(x - m_0) \perp M$.*

Proof The uniqueness and orthogonality are immediate consequences of Theorem 13.12. We are only required to establish the existence of m_0 . Let $\delta := \inf_{m \in M} \|x - m\| \in [0, \infty) \subset \mathbb{R}$. Then, $\exists (m_n)_{n=1}^\infty \subseteq M$ such that $\lim_{n \in \mathbb{N}} \|x - m_n\| = \delta$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n_0 \leq n$, we have $\delta \leq \|x - m_n\| < \sqrt{\delta^2 + \epsilon^2/4}$. By Parallelogram Law 13.3, we have, $\forall i, j \in \mathbb{N}$ with $n_0 \leq i$ and $n_0 \leq j$, $\|m_i - x + x - m_j\|^2 + \|x - m_i + x - m_j\|^2 = 2\|x - m_j\|^2 + 2\|x - m_i\|^2$. Then, $\|m_i - m_j\|^2 = 2\|x - m_j\|^2 + 2\|x - m_i\|^2 - 4\|x - (m_i + m_j)/2\|^2 < 4(\delta^2 + \epsilon^2/4) - 4\delta^2 = \epsilon^2$ and $\|m_i - m_j\| < \epsilon$. This implies that $(m_n)_{n=1}^\infty$ is a Cauchy sequence, which must converge to $m_0 \in M$, since M is a complete subspace by Proposition 4.39. By Proposition 7.21, we have $\|x - m_0\| = \lim_{n \in \mathbb{N}} \|x - m_n\| = \delta = \inf_{m \in M} \|x - m\|$. This completes the proof of the theorem. $\square$

In the above proof, we observe that the key to the existence of the minimizing vector is that M is complete. Hence, we have the following modified version of the projection theorem.

Theorem 13.14 *Let $\mathcal{X}$ be a pre-Hilbert space and $M \subseteq \mathcal{X}$ be a complete subspace. $\forall x \in \mathcal{X}$, there exists a unique vector $m_0 \in M$ such that $\|x - m_0\| = \min_{m \in M} \|x - m\|$. Furthermore, a necessary and sufficient condition for $m_0 \in M$ being the unique minimizing vector is $(x - m_0) \perp M$.*

Proof This is straightforward by the proof of the Projection Theorem 13.13, and is therefore omitted. $\square$

13.3 Dual of Hilbert Spaces

Theorem 13.15 (Riesz-Fréchet) *Let $\mathcal{X}$ be a Hilbert space over $\mathbb{K}$. Then, the following statements hold.*

- (i) $\forall f \in \mathcal{X}^*$, there exists a unique $y_0 \in \mathcal{X}$ such that $f(x) = \langle x, y_0 \rangle$, $\forall x \in \mathcal{X}$, and $\|f\|_{\mathcal{X}^*} = \|y_0\|_{\mathcal{X}}$. Therefore, we may define a mapping $\Phi : \mathcal{X}^* \rightarrow \mathcal{X}$ by $\Phi(f) = y_0$.
- (ii) $\forall y \in \mathcal{X}$, define $g : \mathcal{X} \rightarrow \mathbb{K}$ by $g(x) = \langle x, y \rangle$, $\forall x \in \mathcal{X}$, then $g \in \mathcal{X}^*$.
- (iii) The mapping Φ is bijective, norm preserving, and $\Phi(\alpha f_1 + \beta f_2) = \alpha \Phi(f_1) + \beta \Phi(f_2)$, $\forall f_1, f_2 \in \mathcal{X}^*$, $\forall \alpha, \beta \in \mathbb{K}$.
- (iv) If $\mathbb{K} = \mathbb{R}$, then Φ is a isometrical isomorphism between $\mathcal{X}^*$ and $\mathcal{X}$.
- (v) If $\mathbb{K} = \mathbb{C}$, let $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ be the natural mapping as defined in Remark 7.88, then ϕ is surjective and $\mathcal{X}$ is reflexive.

Henceforth, we will denote $\Phi_{\text{inv}}(x) =: x^*$, $\forall x \in \mathcal{X}$. Furthermore, the following statement hold.

- (vi) When $\mathbb{K} = \mathbb{C}$, let $\Phi_* : \mathcal{X}^{**} = \mathcal{X} \rightarrow \mathcal{X}^*$ be the mapping of Φ if $\mathcal{X}$ is replaced by $\mathcal{X}^*$. Then, $\Phi_* = \Phi_{\text{inv}}$. This leads to the identity $(x^*)^* = x$, $\forall x \in \mathcal{X}$.

Proof We will first show that, $\forall y \in \mathcal{X}$, the functional g defined in the statement of the theorem is a bounded linear functional with $\|g\|_{\mathcal{X}^*} = \|y\|_{\mathcal{X}}$. By Definition 13.1, g is a linear functional. By Proposition 13.2, we have $|g(x)| = |\langle x, y \rangle| \leq \|x\|_{\mathcal{X}} \|y\|_{\mathcal{X}}$. Hence, $g \in \mathcal{X}^*$ and $\|g\|_{\mathcal{X}^*} \leq \|y\|_{\mathcal{X}}$. Note that $|g(y)| = |\langle y, y \rangle| = \|y\|_{\mathcal{X}}^2$. Hence, $\|g\|_{\mathcal{X}^*} \geq \|y\|_{\mathcal{X}}$. Hence, $\|g\|_{\mathcal{X}^*} = \|y\|_{\mathcal{X}}$.

Next, we will show that, for any $f \in \mathcal{X}^*$, there exists a $y_0 \in \mathcal{X}$ such that $f(x) = \langle x, y_0 \rangle$, $\forall x \in \mathcal{X}$. Let $M := \mathcal{N}(f)$. We will distinguish two exhaustive and mutually exclusives cases: Case 1: $M = \mathcal{X}$; Case 2: $M \subset \mathcal{X}$. Case 1: $M = \mathcal{X}$. Take $y = \vartheta_{\mathcal{X}}$. Then, $f(x) = 0 = \langle x, y \rangle$, $\forall x \in \mathcal{X}$. This case is proved.

Case 2: $M \subset \mathcal{X}$. Then, $\exists x_0 \in \mathcal{X} \setminus M$. Since $f \in \mathcal{X}^*$, then M is a closed subspace of $\mathcal{X}$ by Proposition 7.68. By Projection Theorem 13.13, $\exists m_0 \in M$ such that $(x_0 - m_0) \perp M$. Let $z_0 := x_0 - m_0 \in \mathcal{X} \setminus M$. Then, $f(z_0) \neq 0$. Without loss of generality, we may assume that $f(z_0) = 1$. $\forall x \in \mathcal{X}$, by the linearity of f , we have $f(x - f(x)z_0) = f(x) - f(x)f(z_0) = 0$ and $x - f(x)z_0 \in M$. Since $z_0 \perp M$, we have $0 = \langle x - f(x)z_0, z_0 \rangle = \langle x, z_0 \rangle - f(x)\|z_0\|^2$. Then, $f(x) = \langle x, z_0 \rangle / \|z_0\|^2 = \langle x, z_0 / \|z_0\|^2 \rangle$. Let $y_0 := z_0 / \|z_0\|^2 \in \mathcal{X}$, we have $f(x) = \langle x, y_0 \rangle$, $\forall x \in \mathcal{X}$. This case is also proved.

Next, we will show that, for any $f \in \mathcal{X}^*$, the vector $y_0 \in \mathcal{X}$ is unique. Let $y \in \mathcal{X}$ be another vector such that $f(x) = \langle x, y \rangle$, $\forall x \in \mathcal{X}$. Then,

$0 = f(x) - f(x) = \langle x, y_0 \rangle - \langle x, y \rangle = \langle x, y_0 - y \rangle, \forall x \in \mathcal{X}$. Then, $0 = \langle y_0 - y, y_0 - y \rangle = \|y_0 - y\|^2$. Thus, $y = y_0$. Hence, y_0 is unique.

Based on the above, we have (i) and (ii) holds.

(iii) Note that, $\forall f_1, f_2 \in \mathcal{X}^*, \forall \alpha, \beta \in \mathbb{K}, (\alpha f_1 + \beta f_2)(x) = \alpha f_1(x) + \beta f_2(x) = \alpha \langle x, \Phi(f_1) \rangle + \beta \langle x, \Phi(f_2) \rangle = \langle x, \overline{\alpha} \Phi(f_1) + \overline{\beta} \Phi(f_2) \rangle, \forall x \in \mathcal{X}$, where the second equality follows from (i); and the third equality follows from Proposition 13.2. Then, by (i), we have $\Phi(\alpha f_1 + \beta f_2) = \overline{\alpha} \Phi(f_1) + \overline{\beta} \Phi(f_2)$. By (i), $\|\Phi(f)\|_{\mathcal{X}} = \|f\|_{\mathcal{X}^*}, \forall f \in \mathcal{X}^*$. Therefore, Φ is norm preserving. By (ii), Φ is surjective. $\forall f_1, f_2 \in \mathcal{X}^*, \Phi(f_1) = \Phi(f_2)$ implies that $0 = \|\Phi(f_1) - \Phi(f_2)\|_{\mathcal{X}} = \|\Phi(f_1 - f_2)\|_{\mathcal{X}} = \|f_1 - f_2\|_{\mathcal{X}^*}$, which further implies that $f_1 = f_2$. Hence, Φ is injective. Hence, Φ is bijective.

(iv) Let $\mathbb{K} = \mathbb{R}$. By (iii), Φ is linear. Then, by Definition 7.24, Φ is an isometrical isomorphism.

(v) Let $\mathbb{K} = \mathbb{C}$. By Remark 7.88, $\forall x_* \in \mathcal{X}^*, \forall x \in \mathcal{X}$, we have $\langle x_*, x \rangle = \langle \langle \phi(x), x_* \rangle \rangle$. Then, $\forall y \in \mathcal{X}$, we have $\langle x, y \rangle = \langle \langle \Phi_{\text{inv}}(y), x \rangle \rangle = \langle \langle \phi(x), \Phi_{\text{inv}}(y) \rangle \rangle$. We will show that ϕ is surjective, which then implies that ϕ is isometrical isomorphism between $\mathcal{X}$ and $\mathcal{X}^{**}$, and $\mathcal{X}$ is reflexive. $\forall x_{**} \in \mathcal{X}^{**}, \langle \langle x_{**}, x_* \rangle \rangle = \langle \langle x_{**}, \Phi_{\text{inv}}(\Phi(x_*)) \rangle \rangle$. Define $G : \mathcal{X} \rightarrow \mathbb{K}$ by $G(y) = \langle \langle x_{**}, \Phi_{\text{inv}}(y) \rangle \rangle, \forall y \in \mathcal{X}$. By (iii), it is easy to show that $G \in \mathcal{X}^*$. By (i), $G(y) = \langle y, \Phi(G) \rangle, \forall y \in \mathcal{X}$. Then, $\langle \langle x_{**}, x_* \rangle \rangle = \langle \langle x_{**}, \Phi_{\text{inv}}(\Phi(x_*)) \rangle \rangle = \langle \langle \Phi(x_*), \Phi(G) \rangle \rangle = \langle \langle \Phi(x_*), \Phi(G) \rangle \rangle = \langle \langle \phi(\Phi(G)), x_* \rangle \rangle, \forall x_* \in \mathcal{X}^*$. Thus, we have $x_{**} = \phi(\Phi(G)) \in \mathcal{R}(\phi)$. By the arbitrariness of x_{**} , we have ϕ is surjective.

(vi) Let $\mathbb{K} = \mathbb{C}$. $\forall x \in \mathcal{X}$. Let $f = \Phi_*(x)$. Then, $f(x) = \langle x, \Phi(f) \rangle = (\phi_{\mathcal{X}}(x))(f) = \langle f, \Phi_*(\phi_{\mathcal{X}}(x)) \rangle_{\mathcal{X}^*} = \langle f, \Phi_*(x) \rangle_{\mathcal{X}^*} = \langle f, f \rangle_{\mathcal{X}^*} = \|f\|_{\mathcal{X}^*}^2$, where the first equality follows from (i); the second equality follows from the definition for natural mapping $\phi_{\mathcal{X}} : \mathcal{X} \rightarrow \mathcal{X}^{**} = \mathcal{X}$; the third equality follows from (i); and the fourth equality follows from (v). By (iii), $\|f\|_{\mathcal{X}^*} = \|x\|_{\mathcal{X}}$ and $\|f\|_{\mathcal{X}^*} = \|\Phi(f)\|_{\mathcal{X}}$. This implies that $\langle x, \Phi(f) \rangle = \|x\|_{\mathcal{X}} \|\Phi(f)\|_{\mathcal{X}}$. By (iv) of Proposition 13.2, $\Phi(f) = \vartheta_{\mathcal{X}}$ or $x = \lambda \Phi(f)$, for some $\lambda \in \mathbb{C}$.

If $\Phi(f) = \vartheta_{\mathcal{X}}$, by (iii), $f = \vartheta_{\mathcal{X}^*}$. Thus, by (iii), $x = \vartheta_{\mathcal{X}} = \Phi(f)$ since $f = \Phi_*(x)$.

If $\Phi(f) \neq \vartheta_{\mathcal{X}}$, we must have $x = \lambda \Phi(f)$, for some $\lambda \in \mathbb{C}$. Then, $\|f\|_{\mathcal{X}^*}^2 = \langle x, \Phi(f) \rangle = \lambda \langle \Phi(f), \Phi(f) \rangle = \lambda \|\Phi(f)\|_{\mathcal{X}}^2 = \lambda \|f\|_{\mathcal{X}^*}^2$. Thus, $\lambda = 1$. This implies that $x = \Phi(f)$.

In both cases, we have $x = \Phi(f) = \Phi(\Phi_*(x))$. By the arbitrariness of x , we have $\Phi \circ \Phi_* = \text{id}_{\mathcal{X}}$. By (iii) and Proposition 2.4, we have $\Phi_* = \Phi_{\text{inv}}$.

This completes the proof of the theorem. $\square$

Definition 13.16 Let $\mathcal{X}$ be a Hilbert space over $\mathbb{K}$, $S \subseteq \mathcal{X}$. By Definition 7.95, $S^{\perp} \subseteq \mathcal{X}^*$ is the set $S^{\perp} = \{x_* \in \mathcal{X}^* \mid \langle x_*, x \rangle = 0, \forall x \in S\}$. We will abuse the notation and denote $S^{\perp} := \{y \in \mathcal{X} \mid \langle x, y \rangle = \langle y^*, x \rangle = 0, \forall x \in S\}$, which is said to be the orthogonal complement of S .

Proposition 13.17 Let $\mathcal{X}$ be a Hilbert space over $\mathbb{K}$ and $S, T \subseteq \mathcal{X}$. Then,

- (i) $S^\perp \subseteq \mathcal{X}$ is a closed subspace;
- (ii) if $S \subseteq T$, then $T^\perp \subseteq S^\perp$;
- (iii) $S^{\perp\perp\perp} = S^\perp$;
- (iv) $S^{\perp\perp} = \overline{\text{span}(S)}$.

Proof (i) and (ii) follow immediately from Proposition 7.98.

(iv) By Riesz-Fréchet Theorem 13.15, $\mathcal{X}$ is reflexive. Then, $\mathcal{X} = \mathcal{X}^{**}$. By Definition 7.95, ${}^\perp S \subseteq \mathcal{X}^*$ is well-defined and ${}^\perp S = S^\perp$. Then, by Proposition 7.98, we have $S^{\perp\perp} = \overline{{}^\perp(S^\perp)} = \overline{\text{span}(S)}$.

(iii) By (iv), we have $S \subseteq \overline{\text{span}(S)} = S^{\perp\perp}$. By (ii), we have $S^\perp \supseteq S^{\perp\perp\perp}$. Again by (iv), we have $S^\perp \subseteq \overline{\text{span}(S^\perp)} = S^{\perp\perp\perp}$. Therefore, $S^\perp = S^{\perp\perp\perp}$. This completes the proof of the proposition. $\square$

Definition 13.18 Let $\mathcal{X}$ be a vector space over the field $\mathcal{F}$ and $M, N \subseteq \mathcal{X}$ are subspaces. We say that $\mathcal{X}$ is the direct sum of M and N if $\forall x \in \mathcal{X}$, $\exists! m \in M$ and $\exists! n \in N$, such that $x = m + n$. In this case, we will write $\mathcal{X} = M \oplus N$.

Theorem 13.19 Let M be a closed subspace of a Hilbert space $\mathcal{X}$. Then, $\mathcal{X} = M \oplus M^\perp$.

Proof $\forall x \in \mathcal{X}$, by Projection Theorem 13.13, $\exists! m_0 \in M$ such that $n_0 = (x - m_0) \perp M$. Hence, $x = m_0 + n_0$ with $m_0 \in M$ and $n_0 \in M^\perp$. Clearly, m_0 and n_0 is unique since m_0 is unique. Hence, $\mathcal{X} = M \oplus M^\perp$. This completes the proof of the theorem. $\square$

Proposition 13.20 Let $\mathcal{X}$ be a real Hilbert space, $f : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ be given by $f(x, y) = \langle x, y \rangle$, $\forall (x, y) \in \mathcal{X} \times \mathcal{X}$. Then, f is $\mathcal{C}_\infty$, $f^{(1)}(x, y)(h_x, h_y) = \langle h_x, y \rangle + \langle x, h_y \rangle$, $f^{(2)}(x, y)(h_{x2}, h_{y2})(h_{x1}, h_{y1}) = \langle h_{x2}, h_{y1} \rangle + \langle h_{x1}, h_{y2} \rangle$, and $f^{(i+2)}(x, y) = \vartheta_{\text{BS}_{i+2}(\mathcal{X} \times \mathcal{X}, \mathbb{R})}$, $\forall (x, y) \in \mathcal{X} \times \mathcal{X}$, $\forall i \in \mathbb{N}$, $\forall (h_x, h_y) \in \mathcal{X} \times \mathcal{X}$, $\forall (h_{x1}, h_{y1}) \in \mathcal{X} \times \mathcal{X}$, $\forall (h_{x2}, h_{y2}) \in \mathcal{X} \times \mathcal{X}$.

Proof This follows immediately from Propositions 9.41 and 9.27 and Riesz-Fréchet Theorem 13.15. $\square$

13.4 Hermitian Adjoints

Definition 13.21 Let $\mathcal{X}$ and $\mathcal{Y}$ be Hilbert spaces over $\mathbb{K}$, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, and $\Phi_{\mathcal{X}} : \mathcal{X}^* \rightarrow \mathcal{X}$ and $\Phi_{\mathcal{Y}} : \mathcal{Y}^* \rightarrow \mathcal{Y}$ as defined in Riesz-Fréchet Theorem 13.15. Define the Hermitian adjoint of A by $A^* := \Phi_{\mathcal{X}} \circ A' \circ \Phi_{\mathcal{Y} \text{ inv}} \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$. A is said to be Hermitian if $A^* = A$.

Proposition 13.22 Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be Hilbert spaces over $\mathbb{K}$, $A, B \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, $C \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$, and $\lambda \in \mathbb{K}$. Then,

- (i) $\text{id}_{\mathcal{X}}^* = \text{id}_{\mathcal{X}}$;
- (ii) $(A + B)^* = A^* + B^*$;
- (iii) $(\lambda A)^* = \bar{\lambda} A^*$;
- (iv) $(CA)^* = A^* C^*$;
- (v) if A is bijective, then $(A^*)^{-1} = (A^{-1})^* =: A^{-*}$;
- (vi) $\langle Ax, y \rangle_{\mathcal{Y}} = \langle x, A^* y \rangle_{\mathcal{X}}$;
- (vii) $\langle x, Ay \rangle_{\mathcal{X}} = \langle A^* x, y \rangle_{\mathcal{Y}}$;
- (viii) $(A^*)^* = A$;
- (ix) $\|A^*\|_{\mathcal{B}(\mathcal{Y}, \mathcal{X})} = \|A\|_{\mathcal{B}(\mathcal{X}, \mathcal{Y})}$;
- (x) $(\mathcal{R}(A))^{\perp} = \mathcal{N}(A^*)$ and $(\mathcal{R}(A^*))^{\perp} = \mathcal{N}(A)$;
- (xi) if $\mathcal{R}(A) \subseteq \mathcal{X}$ is closed, then $\mathcal{R}(A) = (\mathcal{N}(A^*))^{\perp}$ and $\mathcal{R}(A^*) = (\mathcal{N}(A))^{\perp}$.

Proof Let $\Phi_{\mathcal{X}} : \mathcal{X}^* \rightarrow \mathcal{X}$, $\Phi_{\mathcal{Y}} : \mathcal{Y}^* \rightarrow \mathcal{Y}$, and $\Phi_{\mathcal{Z}} : \mathcal{Z}^* \rightarrow \mathcal{Z}$ be as defined in Riesz-Fréchet Theorem 13.15.

(i) $\text{id}_{\mathcal{X}}^* = \Phi_{\mathcal{X}} \circ \text{id}_{\mathcal{X}}' \circ \Phi_{\mathcal{X} \text{ inv}} = \Phi_{\mathcal{X}} \circ \text{id}_{\mathcal{X}^*} \circ \Phi_{\mathcal{X} \text{ inv}} = \text{id}_{\mathcal{X}}$, where the first equality follows from Definition 13.21; and the second equality follows from Proposition 7.110.

(ii) $(A + B)^* = \Phi_{\mathcal{X}} \circ (A + B)' \circ \Phi_{\mathcal{Y} \text{ inv}} = \Phi_{\mathcal{X}} \circ (A' + B') \circ \Phi_{\mathcal{Y} \text{ inv}} = \Phi_{\mathcal{X}} \circ A' \circ \Phi_{\mathcal{Y} \text{ inv}} + \Phi_{\mathcal{X}} \circ B' \circ \Phi_{\mathcal{Y} \text{ inv}} = A^* + B^*$, where the first equality follows from Definition 13.21; the second equality follows from Proposition 7.110; and the third equality follows from Riesz-Fréchet Theorem 13.15.

(iii) $(\lambda A)^* = \Phi_{\mathcal{X}} \circ (\lambda A)' \circ \Phi_{\mathcal{Y} \text{ inv}} = \Phi_{\mathcal{X}} \circ (\lambda A') \circ \Phi_{\mathcal{Y} \text{ inv}} = \bar{\lambda} \Phi_{\mathcal{X}} \circ A' \circ \Phi_{\mathcal{Y} \text{ inv}} = \bar{\lambda} A^*$, where the first equality follows from Definition 13.21; the second equality follows from Proposition 7.110; and the third equality follows from Riesz-Fréchet Theorem 13.15.

(iv) $(CA)^* = \Phi_{\mathcal{X}} \circ (CA)' \circ \Phi_{\mathcal{Z} \text{ inv}} = \Phi_{\mathcal{X}} \circ (A' C') \circ \Phi_{\mathcal{Z} \text{ inv}} = \Phi_{\mathcal{X}} \circ A' \circ C' \circ \Phi_{\mathcal{Z} \text{ inv}} = \Phi_{\mathcal{X}} \circ A' \circ \Phi_{\mathcal{Y} \text{ inv}} \circ \Phi_{\mathcal{Y}} \circ C' \circ \Phi_{\mathcal{Z} \text{ inv}} = A^* \circ C^* = A^* C^*$, where the first equality follows from Definition 13.21; and the second equality follows from Proposition 7.110.

(v) $(A^{-1})^* = \Phi_{\mathcal{Y}} \circ (A^{-1})' \circ \Phi_{\mathcal{X} \text{ inv}} = \Phi_{\mathcal{Y}} \circ (A')^{-1} \circ \Phi_{\mathcal{X} \text{ inv}} = (\Phi_{\mathcal{X}} \circ A' \circ \Phi_{\mathcal{Y} \text{ inv}})^{-1} = (A^*)^{-1}$, where the first equality follows from Definition 13.21; and the second equality follows from Proposition 7.110.

(vi) $\langle Ax, y \rangle_{\mathcal{Y}} = \langle \langle y^*, Ax \rangle \rangle_{\mathcal{Y}} = \langle \langle A' y^*, x \rangle \rangle_{\mathcal{X}} = \langle x, \Phi_{\mathcal{X}}(A' y^*) \rangle_{\mathcal{X}} = \langle x, \Phi_{\mathcal{X}}(A'(\Phi_{\mathcal{Y} \text{ inv}}(y))) \rangle_{\mathcal{X}} = \langle x, A^* y \rangle_{\mathcal{X}}$, where the first equality follows from Riesz-Fréchet Theorem 13.15; the second equality follows from Proposition 7.109; the third and fourth equalities follow from Riesz-Fréchet Theorem 13.15; and the last equality follows from Definition 13.21.

(vii) $\langle x, Ay \rangle_{\mathcal{X}} = \overline{\langle Ay, x \rangle_{\mathcal{X}}} = \overline{\langle y, A^*x \rangle_{\mathcal{Y}}} = \langle A^*x, y \rangle_{\mathcal{Y}}$, where the first equality follows from Definition 13.1; the second equality follows from (vi); and the third equality follows from Definition 13.1.

(viii) $\langle x, Ay \rangle_{\mathcal{X}} = \langle A^*x, y \rangle_{\mathcal{Y}} = \langle x, (A^*)^*y \rangle_{\mathcal{X}}$, where the first equality follows from (vii); and the second equality follows from (vi). By the arbitrariness of x and y , we have $A = (A^*)^*$.

(ix) $\|A^*\|_{B(\mathcal{Y}, \mathcal{X})} = \sup_{y \in \mathcal{Y}, \|y\|_{\mathcal{Y}} \leq 1} \|A^*y\|_{\mathcal{X}} = \sup_{y \in \mathcal{Y}, \|y\|_{\mathcal{Y}} \leq 1} \|\Phi_{\mathcal{X}} \circ A' \circ \Phi_{\mathcal{Y} \text{ inv}}(y)\|_{\mathcal{X}} = \|A'\|_{B(\mathcal{Y}^*, \mathcal{X}^*)} = \|A\|_{B(\mathcal{X}, \mathcal{Y})}$, where the first equality follows from Proposition 7.63; the second equality follows from Definition 13.21; the third equality follows from the fact that Φ is bijective and norm preserving; and the fourth equality follows from Proposition 7.109.

(x) By Proposition 7.112 and Definition 13.16, we have $(\mathcal{R}(A))^{\perp} = \Phi_{\mathcal{Y}}(\mathcal{N}(A')) = \mathcal{N}(A^*)$. Then, $(\mathcal{R}(A^*))^{\perp} = \mathcal{N}((A^*)^*) = \mathcal{N}(A)$, where the first equality follows from the previous sentence; and the second equality follows from (viii).

(xi) By Proposition 7.114 and Definition 13.16, we have $\mathcal{R}(A) = {}^{\perp}(\mathcal{N}(A')) = {}^{\perp}(\Phi_{\mathcal{Y} \text{ inv}}(\Phi_{\mathcal{Y}}(\mathcal{N}(A')))) = (\Phi_{\mathcal{Y}}(\mathcal{N}(A')))^{\perp} = (\mathcal{N}(A^*))^{\perp}$. Furthermore, $\mathcal{R}(A^*) = (\mathcal{N}((A^*)^*))^{\perp} = (\mathcal{N}(A))^{\perp}$, where the first equality follows from the previous sentence; the second equality follows from (viii).

This completes the proof of the proposition. $\square$

13.5 Approximation in Hilbert Spaces

Definition 13.23 Let $\mathcal{X}$ be a pre-Hilbert space and $S \subseteq \mathcal{X}$. We will say that S is an orthogonal set if $x \perp y, \forall x, y \in S$ with $x \neq y$. The set S is said to be orthonormal if S is an orthogonal set and $\|x\| = 1, \forall x \in S$.

Proposition 13.24 Let $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$ and $S \subseteq \mathcal{X}$ be an orthogonal set. Assume that $\vartheta_{\mathcal{X}} \notin S$. Then, S is a linearly independent set.

Proof Suppose S is not a linearly independent set. Then, $\exists x_0 \in S$ such that $x_0 \in \text{span}(S \setminus \{x_0\})$. Then, $\exists n \in \mathbb{Z}_+, \exists x_1, \dots, x_n \in S \setminus \{x_0\}$ with $x_i \neq x_j, \forall 1 \leq i < j \leq n, \exists \alpha_1, \dots, \alpha_n \in \mathbb{K}$, such that $x_0 = \sum_{i=1}^n \alpha_i x_i$. Then, $0 = \|x_0 - \sum_{i=1}^n \alpha_i x_i\|^2 = \langle x_0, x_0 \rangle + \sum_{i=1}^n \langle \alpha_i x_i, \alpha_i x_i \rangle = \|x_0\|^2 + \sum_{i=1}^n |\alpha_i|^2 \|x_i\|^2 > 0$, where the second equality follows from the fact that S is an orthogonal set; and the inequality follows from the fact that $x_0 \neq \vartheta_{\mathcal{X}}$. This is a contradiction. Hence, S is a linearly independent set. This completes the proof of the proposition. $\square$

In Hilbert space, orthonormal sets are greatly favored over other linearly independent sets, and it is reassuring to know that they exist and can be easily constructed. The next theorem presents the Gram-Schmidt orthogonalization procedure.

Theorem 13.25 (Gram-Schmidt Procedure) *Let $n \in \mathbb{N} \cup \{\infty\}$, $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$, and $(x_i)_{i=1}^n \subseteq \mathcal{X}$ be a linearly independent set and vectors in the sequence are distinct. Then, there is an orthonormal sequence $(e_i)_{i=1}^n \subseteq \mathcal{X}$ such that $\text{span}(\{x_1, \dots, x_k\}) = \text{span}(\{e_1, \dots, e_k\})$, $\forall k \in \mathbb{N}$ with $k \leq n$, and $\text{span}((x_i)_{i=1}^n) = \text{span}((e_i)_{i=1}^n)$.*

Proof We construct $(e_i)_{i=1}^n$ recursively as follows.

1° For $i = 1$, clearly $x_1 \neq \vartheta_{\mathcal{X}}$ since $(x_i)_{i=1}^n$ is a linearly independent set. Let $e_1 := x_1 / \|x_1\| \in \mathcal{X}$. Then, $\text{span}(\{e_1\}) = \text{span}(\{x_1\})$ and the sequence $(e_i)_{i=1}^1$ is orthonormal. This completes the first step of the construction.

2° Assume that we have constructed $(e_i)_{i=1}^k$, where $k \in \mathbb{N}$ and $k < n$, such that the sequence is orthonormal and $\text{span}(\{e_1, \dots, e_k\}) = \text{span}(\{x_1, \dots, x_k\})$.

3° Consider $k + 1$. Let $z_{k+1} := x_{k+1} - \sum_{i=1}^k \langle x_{k+1}, e_i \rangle e_i \in \mathcal{X}$. It is easy to show that $\langle z_{k+1}, e_i \rangle = 0$, $\forall i \in \{1, \dots, k\}$. Then, $z_{k+1} \perp (e_i)_{i=1}^k$. By the assumption and Lemma 6.49, the vectors $x_1, \dots, x_{k+1}$ are linearly independent. Then, $z_{k+1} \neq \vartheta_{\mathcal{X}}$. Define $e_{k+1} := z_{k+1} / \|z_{k+1}\| \in \mathcal{X}$. The inductive assumption and the construction imply that $(e_i)_{i=1}^{k+1}$ is orthonormal and $\text{span}(\{e_1, \dots, e_{k+1}\}) = \text{span}(\{x_1, \dots, x_{k+1}\})$.

This completes the induction process. Then, there exists an orthonormal $(e_i)_{i=1}^n \subseteq \mathcal{X}$ such that $\text{span}(\{e_1, \dots, e_k\}) = \text{span}(\{x_1, \dots, x_k\})$, $\forall k \in \mathbb{N}$ with $k \leq n$.

If $n \in \mathbb{N}$, the theorem is proved. Consider the case when $n = \infty$. We need to show that $\text{span}((e_i)_{i=1}^{\infty}) = \text{span}((x_i)_{i=1}^{\infty})$. $\forall x \in \text{span}((x_i)_{i=1}^{\infty})$, $\exists m \in \mathbb{N}$, $\exists \alpha_1, \dots, \alpha_m \in \mathbb{K}$ such that $x = \sum_{i=1}^m \alpha_i x_i$. Then, $x \in \text{span}(\{x_1, \dots, x_m\}) = \text{span}(\{e_1, \dots, e_m\}) \subseteq \text{span}((e_i)_{i=1}^{\infty})$. Hence, we have $\text{span}((e_i)_{i=1}^{\infty}) \subseteq \text{span}((x_i)_{i=1}^{\infty})$. By symmetry, we have $\text{span}((x_i)_{i=1}^{\infty}) \subseteq \text{span}((e_i)_{i=1}^{\infty})$. Hence, $\text{span}((e_i)_{i=1}^{\infty}) = \text{span}((x_i)_{i=1}^{\infty})$. This completes the proof of the theorem. $\square$

Definition 13.26 *Let $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$, $n \in \mathbb{N}$, $y_1, \dots, y_n \in \mathcal{X}$. The Gram matrix of $y_1, \dots, y_n$ is*

$$\text{Gram}(y_1, \dots, y_n) := \begin{bmatrix} \langle y_1, y_1 \rangle & \langle y_1, y_2 \rangle & \cdots & \langle y_1, y_n \rangle \\ \langle y_2, y_1 \rangle & \langle y_2, y_2 \rangle & \cdots & \langle y_2, y_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle y_n, y_1 \rangle & \langle y_n, y_2 \rangle & \cdots & \langle y_n, y_n \rangle \end{bmatrix}$$

The Gram determinant of $y_1, \dots, y_n$ is $\text{gram}(y_1, \dots, y_n) := \det(\text{Gram}(y_1, \dots, y_n))$.

Proposition 13.27 *Let $n \in \mathbb{N}$, $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$, $y_1, \dots, y_n \in \mathcal{X}$. Then, $\text{gram}(y_1, \dots, y_n) \neq 0$ if, and only if, $y_1, \dots, y_n$ are linearly independent.*

Proof We will prove the equivalent statement that $\text{gram}(y_1, \dots, y_n) = 0$ if, and only if, $y_1, \dots, y_n$ are linearly dependent. “Sufficiency” Let $y_1, \dots, y_n$ be linearly dependent. Then, $\exists \alpha_1, \dots, \alpha_n \in \mathbb{K}$, which are not all zeros, such that $\sum_{i=1}^n \alpha_i y_i = \vartheta$. Then,

$$\begin{aligned} \mathbf{0}_{1 \times n} &= \left[\langle \sum_{i=1}^n \alpha_i y_i, y_1 \rangle \quad \cdots \quad \langle \sum_{i=1}^n \alpha_i y_i, y_n \rangle \right] \\ &= \left[\alpha_1 \quad \cdots \quad \alpha_n \right] \text{Gram}(y_1, \dots, y_n) \end{aligned}$$

This shows that the $\text{rank}(\text{Gram}(y_1, \dots, y_n)) < n$. Therefore, $\text{gram}(y_1, \dots, y_n) = 0$.

“Necessity” $\text{gram}(y_1, \dots, y_n) = 0$ implies that $\text{rank}(\text{Gram}(y_1, \dots, y_n)) < n$. Then, the row vectors of the Gram matrix is linearly dependent. This implies that $\exists \alpha_1, \dots, \alpha_n \in \mathbb{K}$, which are not all zeros, such that

$$\begin{aligned} \mathbf{0}_{1 \times n} &= \left[\alpha_1 \quad \cdots \quad \alpha_n \right] \text{Gram}(y_1, \dots, y_n) \\ &= \left[\langle \sum_{i=1}^n \alpha_i y_i, y_1 \rangle \quad \cdots \quad \langle \sum_{i=1}^n \alpha_i y_i, y_n \rangle \right] \end{aligned}$$

Then,

$$\begin{aligned} 0 &= \left[\langle \sum_{i=1}^n \alpha_i y_i, y_1 \rangle \quad \cdots \quad \langle \sum_{i=1}^n \alpha_i y_i, y_n \rangle \right] \begin{bmatrix} \overline{\alpha_1} \\ \vdots \\ \overline{\alpha_n} \end{bmatrix} \\ &= \sum_{j=1}^n \overline{\alpha_j} \left\langle \sum_{i=1}^n \alpha_i y_i, y_j \right\rangle = \left\langle \sum_{i=1}^n \alpha_i y_i, \sum_{j=1}^n \alpha_j y_j \right\rangle \\ &= \left\| \sum_{i=1}^n \alpha_i y_i \right\|^2 \end{aligned}$$

Hence, $\sum_{i=1}^n \alpha_i y_i = \vartheta$ and $y_1, \dots, y_n$ are linearly dependent.

This completes the proof of the proposition. $\square$

Theorem 13.28 Let $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$, $n \in \mathbb{N}$, $y_1, \dots, y_n \in \mathcal{X}$, $M = \text{span}(\{y_1, \dots, y_n\})$, $x \in \mathcal{X}$, and $\delta := \min_{m \in M} \|x - m\|$. Then, the following statements hold.

- (i) There exists a unique $m_0 \in M$ such that $\delta = \|x - m_0\|$ and $m_0 = \sum_{i=1}^n \alpha_i y_i$, where $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ are solutions to the following normal equation

$$(\text{Gram}(y_1, \dots, y_n))' \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \langle x, y_1 \rangle \\ \vdots \\ \langle x, y_n \rangle \end{bmatrix} \quad (13.1)$$

- (ii) If the vectors $y_1, \dots, y_n$ are linearly independent, then the Gram matrix $\text{Gram}(y_1, \dots, y_n)$ is invertible and $\delta^2 = \frac{\text{gram}(y_1, \dots, y_n, x)}{\text{gram}(y_1, \dots, y_n)}$.

Proof (i) Clearly, M is a finite-dimensional subspace of $\mathcal{X}$. By Theorem 7.36, M is complete. By Theorem 13.14, $\exists! m_0 \in M$ such that $\delta = \|x - m_0\|$ and $(x - m_0) \perp M$. Let $m_0 = \sum_{i=1}^n \alpha_i y_i$ with $\alpha_1, \dots, \alpha_n \in \mathbb{K}$. Then, $\langle x - \sum_{i=1}^n \alpha_i y_i, y_j \rangle = 0$, $j = 1, \dots, n$. By Proposition 13.2, we have $\sum_{i=1}^n \alpha_i \langle y_i, y_j \rangle = \langle x, y_j \rangle$, $j = 1, \dots, n$. Thus, (13.1) holds.

(ii) If vectors $y_1, \dots, y_n$ are linearly independent, then the Gram determinant $\text{gram}(y_1, \dots, y_n) \neq 0$, by Proposition 13.27, which further implies that the Gram matrix $\text{Gram}(y_1, \dots, y_n)$ is invertible. Note that $\delta^2 = \|x - m_0\|^2 = \langle x - m_0, x - m_0 \rangle = \langle x - m_0, x \rangle = \langle x, x \rangle - \sum_{i=1}^n \alpha_i \langle y_i, x \rangle$, where the third equality follows from the fact that $(x - m_0) \perp m_0 \in M$. Equivalently, we have $\sum_{i=1}^n \alpha_i \langle y_i, x \rangle + \delta^2 = \langle x, x \rangle$. Combining this equation with the normal equation, we have the following matrix equation.

$$\left[\begin{array}{ccc|c} (\text{Gram}(y_1, \dots, y_n))' & & & \mathbf{0}_{n \times 1} \\ \hline \langle y_1, x \rangle & \cdots & \langle y_n, x \rangle & 1 \end{array} \right] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ \frac{\alpha_n}{\delta^2} \end{bmatrix} = \begin{bmatrix} \langle x, y_1 \rangle \\ \vdots \\ \langle x, y_n \rangle \\ \langle x, x \rangle \end{bmatrix}$$

By Cramer's Rule, we have $\delta^2 = \frac{\det(G_1)}{\det(G)}$, where

$$G = \begin{bmatrix} \langle y_1, y_1 \rangle & \cdots & \langle y_n, y_1 \rangle & 0 \\ \vdots & & \vdots & \vdots \\ \langle y_1, y_n \rangle & \cdots & \langle y_n, y_n \rangle & 0 \\ \langle y_1, x \rangle & \cdots & \langle y_n, x \rangle & 1 \end{bmatrix}; \quad G_1 = (\text{Gram}(y_1, \dots, y_n, x))'$$

By Laplace Expansion Theorem, we have $\det(G) = \det((\text{Gram}(y_1, \dots, y_n))') = \det(\text{Gram}(y_1, \dots, y_n)) = \text{gram}(y_1, \dots, y_n)$. Clearly, $\det(G_1) = \text{gram}(y_1, \dots, y_n, x)$. Hence, $\delta^2 = \frac{\text{gram}(y_1, \dots, y_n, x)}{\text{gram}(y_1, \dots, y_n)}$.

This completes the proof of the theorem. $\square$

In Gram Theorem 13.28, when $y_1, \dots, y_n$ are orthonormal, we denote y_i by e_i , $i = 1, \dots, n$. Then, $\alpha_i = \langle x, e_i \rangle$, $i = 1, \dots, n$, $m_0 = \sum_{i=1}^n \langle x, e_i \rangle e_i$, and $\delta = \sqrt{\langle x, x \rangle - \sum_{i=1}^n |\langle x, e_i \rangle|^2}$.

Proposition 13.29 Let $\mathcal{X}$ be Hilbert space over $\mathbb{K}$, $(e_i)_{i=1}^\infty \subseteq \mathcal{X}$ be orthonormal, and $(\xi_i)_{i=1}^\infty \subseteq \mathbb{K}$. Then, $x = \sum_{i=1}^\infty \xi_i e_i \in \mathcal{X}$ if, and only if, $\sum_{i=1}^\infty |\xi_i|^2 < \infty$. In this case, $\xi_i = \langle x, e_i \rangle$, $\forall i \in \mathbb{N}$.

Proof Define the partial sum : $s_i := \sum_{j=1}^i \xi_j e_j \in \mathcal{X}$, $\forall i \in \mathbb{N}$.

"Sufficiency" Let $\sum_{i=1}^\infty |\xi_i|^2 < \infty$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\sum_{i=n_0+1}^\infty |\xi_i|^2 < \epsilon^2$. $\forall n, m \in \mathbb{N}$ with $n_0 \leq n \leq m$, $\|s_n - s_m\|^2 = \langle \sum_{i=n+1}^m \xi_i e_i, \sum_{i=n+1}^m \xi_i e_i \rangle = \sum_{i=n+1}^m |\xi_i|^2 < \epsilon^2$, where the second equality follows from the assumption that $(e_i)_{i=1}^\infty$ is orthonormal. Hence, $\|s_n - s_m\| < \epsilon$ and $(s_i)_{i=1}^\infty \subseteq \mathcal{X}$ is a Cauchy sequence. By the completeness of $\mathcal{X}$, $\sum_{i=1}^\infty \xi_i e_i = \lim_{n \in \mathbb{N}} s_n =: x \in \mathcal{X}$.

“Necessity” Let $x = \sum_{i=1}^{\infty} \xi_i e_i \in \mathcal{X}$. Then, $x = \lim_{n \in \mathbb{N}} s_n$ and $(s_n)_{n=1}^{\infty} \subseteq \mathcal{X}$ is a Cauchy sequence. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$, $\forall n, m \in \mathbb{N}$ with $n_0 \leq n \leq m$, we have $\|s_n - s_m\| < \sqrt{\epsilon}$. Then, $\sum_{i=n+1}^m |\xi_i|^2 = \langle \sum_{i=n+1}^m \xi_i e_i, \sum_{i=n+1}^m \xi_i e_i \rangle = \|s_n - s_m\|^2 < \epsilon$. Hence, $\sum_{i=1}^{\infty} |\xi_i|^2 < \infty$.

When $x = \sum_{i=1}^{\infty} \xi_i e_i \in \mathcal{X}$, by Propositions 13.4 and 3.66, we have $\langle x, e_i \rangle = \lim_{n \in \mathbb{N}} \langle s_n, e_i \rangle = \xi_i$, $\forall i \in \mathbb{N}$. This completes the proof of the proposition. $\square$

Proposition 13.30 (Bessel’s Inequality) *Let $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$, $x \in \mathcal{X}$, and $(e_i)_{i=1}^{\infty} \subseteq \mathcal{X}$ be orthonormal. Then, $\sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \leq \|x\|^2$.*

Proof Let $\alpha_i := \langle x, e_i \rangle \in \mathbb{K}$, $\forall i \in \mathbb{N}$. $\forall n \in \mathbb{N}$, $0 \leq \|x - \sum_{i=1}^n \alpha_i e_i\|^2 = \|x\|^2 - \sum_{i=1}^n \alpha_i \langle e_i, x \rangle - \sum_{i=1}^n \overline{\alpha_i} \langle x, e_i \rangle + \sum_{i=1}^n |\alpha_i|^2 = \|x\|^2 - \sum_{i=1}^n |\alpha_i|^2$. Hence, $\sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \leq \|x\|^2$. $\square$

Proposition 13.31 *Let $\mathcal{X}$ be a Hilbert space over $\mathbb{K}$, $x \in \mathcal{X}$, and $(e_i)_{i=1}^{\infty} \subseteq \mathcal{X}$ be orthonormal. Then, $m_0 := \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \in \overline{\text{span}((e_i)_{i=1}^{\infty})}$, $(x - m_0) \perp M$, and $\|x - m_0\| = \min_{m \in M} \|x - m\|$.*

Proof By Bessel’s Inequality 13.30, $\sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \leq \|x\|^2 < \infty$. By Proposition 13.29, $m_0 := \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \in \mathcal{X}$. Since $\sum_{i=1}^n \langle x, e_i \rangle e_i \in M$, $\forall n \in \mathbb{N}$, then $m_0 \in M$ by Proposition 4.13. By Propositions 13.4 and 3.66, $\langle x - m_0, e_j \rangle = \lim_{n \in \mathbb{N}} \langle x - \sum_{i=1}^n \langle x, e_i \rangle e_i, e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0$, $\forall j \in \mathbb{N}$. Hence, $(x - m_0) \in ((e_i)_{i=1}^{\infty})^{\perp} = ((e_i)_{i=1}^{\infty})^{\perp \perp \perp} = \left(\overline{\text{span}((e_i)_{i=1}^{\infty})} \right)^{\perp} = M^{\perp}$, where the first and the second equalities follow from Proposition 13.17. By Projection Theorem 13.13, $\|x - m_0\| = \min_{m \in M} \|x - m\|$. This completes the proof of the proposition. $\square$

Based on Proposition 13.31, it is apparent that if $M = \mathcal{X}$, then, $\forall x \in \mathcal{X}$ can be expanded as a infinite series of e_i ’s with coefficients equal to the Fourier coefficients $\langle x, e_i \rangle$. This motivates the following definition and proposition.

Definition 13.32 *Let $\mathcal{X}$ be a Hilbert space, $n \in \mathbb{Z}_+ \cup \{\infty\}$, and $(e_i)_{i=1}^n \subseteq \mathcal{X}$ be orthonormal. The sequence $(e_i)_{i=1}^n$ is said to be complete if $\overline{\text{span}((e_i)_{i=1}^n)} = \mathcal{X}$.*

Proposition 13.33 *Let $\mathcal{X}$ be a Hilbert space, $n \in \mathbb{Z}_+ \cup \{\infty\}$, and $(e_i)_{i=1}^n \subseteq \mathcal{X}$ be orthonormal. The sequence is complete if, and only if, $((e_i)_{i=1}^n)^{\perp} = \{\vartheta\}$.*

Proof By Proposition 13.17, $((e_i)_{i=1}^n)^{\perp} = \{\vartheta\} \Leftrightarrow ((e_i)_{i=1}^n)^{\perp \perp \perp} = \{\vartheta\} \Leftrightarrow \left(\overline{\text{span}((e_i)_{i=1}^n)} \right)^{\perp} = \{\vartheta\} \Leftrightarrow \overline{\text{span}((e_i)_{i=1}^n)} = \mathcal{X}$. This completes the proof of the proposition. $\square$

Example 13.34 Let $I := [-1, 1] \subset \mathbb{R}$, $\mathbb{I} := ((I, |\cdot|), \mathcal{B}, \mu)$ be the complete metric measure subspace of $\mathbb{R}$, $\mathcal{Z} := L_2(\mathbb{I}, \mathbb{K})$, $z_i \in \bar{L}_2(\mathbb{I}, \mathbb{K}) =: \tilde{\mathcal{Z}}$ be defined by $z_i(t) = t^i$, $\forall t \in I$, $\forall i \in \mathbb{Z}_+$. It is easy to show that $([z_i])_{i=0}^\infty \subseteq \mathcal{Z}$ is a linearly independent set. Then, the Gram-Schmidt Procedure 13.25 can be applied to this sequence to produce an orthonormal sequence $([e_i])_{i=0}^\infty \subseteq \mathcal{Z}$, where $e_i \in \tilde{\mathcal{Z}}$, $\forall i \in \mathbb{Z}_+$, are i -order polynomials. Let $p_i \in \tilde{\mathcal{Z}}$ be defined by $p_i(t) = \frac{(-1)^i}{2^i i!} (1 - t^2)^i$, $\forall t \in I$, $\forall i \in \mathbb{Z}_+$, and $P_i \in \tilde{\mathcal{Z}}$ be defined by $P_i(t) = p_i^{(i)}(t)$, $\forall t \in I$, $\forall i \in \mathbb{Z}_+$. Then, we will show that $e_i = \sqrt{\frac{2i+1}{2}} P_i$, $\forall i \in \mathbb{Z}_+$, and $([e_i])_{i=0}^\infty \subseteq \mathcal{Z}$ is a complete orthonormal sequence. As defined, $(P_i)_{i=0}^\infty \subseteq \mathcal{Z}$ is the well-known Legendre polynomials.

We will first show that $([P_i])_{i=0}^\infty$ is an orthogonal set. $\forall n, m \in \mathbb{Z}_+$, we have

$$\begin{aligned} \langle [P_n], [P_m] \rangle &= \int_{-1}^1 P_n(t) \overline{P_m(t)} dt = \int_{-1}^1 p_n^{(n)}(t) p_m^{(m)}(t) dt \\ &= p_n^{(n-1)}(t) p_m^{(m)}(t) \Big|_{-1}^1 - \int_{-1}^1 p_n^{(n-1)}(t) p_m^{(m+1)}(t) dt \\ &= - \int_{-1}^1 p_n^{(n-1)}(t) p_m^{(m+1)}(t) dt \\ &\vdots \\ &= (-1)^n \int_{-1}^1 p_n^{(0)}(t) p_m^{(m+n)}(t) dt \end{aligned}$$

where the first equality follows from Example 13.8; the third equality follows from Integration by Parts Theorem 12.89; the fourth equality follows the fact $p_n^{(i)}(x) = 0$, $\forall i \in \mathbb{Z}_+$ with $0 \leq i < n$, $x \in \{1, -1\}$; and the last equality follows from the recursive application of the Integration by Parts Theorem 12.89. Thus, if $m < n$, we have $p_m^{(m+n)} = \vartheta_{\tilde{\mathcal{Z}}}$ since p_m is a polynomial of order $2m < m+n$. Then, $\langle [P_n], [P_m] \rangle = 0$. Hence, $([P_i])_{i=0}^\infty$ is an orthogonal set.

Next, we show that $\|[P_i]\|_2 = \sqrt{\frac{2}{2i+1}}$, $\forall i \in \mathbb{Z}_+$. Clearly, $P_0(t) = 1$, $\forall t \in I$ and $\|[P_0]\|_2^2 = \langle [P_0], [P_0] \rangle = 2 = \frac{2}{2 \cdot 0 + 1}$. Recursively, assume that $\|[P_{n-1}]\|_2 = \sqrt{\frac{2}{2n-1}}$, for some $(n-1) \in \mathbb{Z}_+$. Then,

$$\begin{aligned} \|[P_n]\|_2^2 &= \langle [P_n], [P_n] \rangle = (-1)^n \int_{-1}^1 p_n^{(0)}(t) p_n^{(2n)}(t) dt \\ &= \frac{(2n)!}{2^{2n} n! n!} \int_{-1}^1 (1 - t^2)^n dt \\ &= \frac{(2n)!}{2^{2n} n! n!} (1 - t^2)^n t \Big|_{-1}^1 + \frac{2n(2n)!}{2^{2n} n! n!} \int_{-1}^1 t^2 (1 - t^2)^{n-1} dt \end{aligned}$$

$$\begin{aligned}
&= \frac{2n(2n)!}{2^{2n}n!n!} \left(\int_{-1}^1 (1-t^2)^{n-1} dt - \int_{-1}^1 (1-t^2)^n dt \right) \\
&= (2n-1) \| [P_{n-1}] \|_2^2 - 2n \| [P_n] \|_2^2
\end{aligned}$$

where the second equality follows from the previous paragraph; the third equality follows from the definition of $p_n \in \tilde{\mathcal{Z}}$; the fourth equality follows from Integration by Parts Theorem 12.89; the fifth equality follows from Proposition 11.92; the last equality follows from the third equality. Therefore, we have $\| [P_n] \|_2^2 = \frac{2}{2n+1}$. This completes the induction process.

Clearly, P_i is an i th order polynomial with positive leading coefficient, then e_i is also an i th order polynomial with positive leading coefficient. By the above, we have $([e_i])_{i=0}^\infty$ is orthonormal and $\text{span}(\{e_0, \dots, e_n\}) = \text{span}(\{z_0, \dots, z_n\})$, $\forall n \in \mathbb{Z}_+$. Then, $([e_i])_{i=0}^\infty$ is the sequence generated by the Gram-Schmidt Procedure on $([z_i])_{i=0}^\infty$.

Finally, we will show that $([e_i])_{i=0}^\infty \subseteq \mathcal{Z}$ is complete by Proposition 13.33. All we need to show is that $\forall z \in \tilde{\mathcal{Z}}$ with $\langle [z], [z_i] \rangle = 0$, $\forall i \in \mathbb{Z}_+$, we have $[z] = \vartheta_{\mathcal{Z}}$. Since $\mu(I) = 2 < \infty$, by Cauchy-Schwarz Inequality, we have $\int_{-1}^1 \mathcal{P} \circ z dt \leq (\int_{-1}^1 \mathcal{P}_2 \circ z dt)^{1/2} (\int_{-1}^1 1 dt)^{1/2} = \sqrt{2} \| z \|_2 < \infty$. Then, $z \in \bar{\mathcal{L}}_1(\mathbb{I}, \mathbb{K})$. Define $F : I \rightarrow \mathbb{K}$ by $F(t) = \int_{-1}^t z(\tau) d\tau$, $\forall t \in I$. By Proposition 12.75, F is absolutely continuous and therefore continuous. Since $[z] \perp [z_0]$, then $F(-1) = 0 = F(1) = \langle [z], [z_0] \rangle$. By Proposition 11.37, F is $\mathcal{B}$ -measurable. Then, $F \in \tilde{\mathcal{Z}}$. $\forall i \in \mathbb{Z}_+$, we have

$$\begin{aligned}
\langle [F], [z_i] \rangle &= \int_{-1}^1 F(t) t^i dt \\
&= \frac{t^{i+1}}{i+1} F(t) \Big|_{-1}^1 - \int_{-1}^1 \frac{t^{i+1}}{i+1} z(t) dt = 0
\end{aligned}$$

where the second equality follows from Integration by Parts Theorem 12.89 and Fundamental Theorem of Calculus I 12.88. Hence, $[F] \perp ([z_i])_{i=0}^\infty$. Since F is continuous, by Bernstein Approximation Theorem 11.203, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists$ polynomial $Q : I \rightarrow \mathbb{K}$, such that $\| F - Q \|_{\mathcal{C}(I, \mathbb{K})} < \epsilon$. Then,

$$\begin{aligned}
\| [F] \|_2^2 &= \int_{-1}^1 |F(t)|^2 dt = \int_{-1}^1 F(t) \overline{(F(t) - Q(t))} dt \\
&\leq \| [F] \|_2 \| [F(t) - Q(t)] \|_2 \leq \| [F] \|_2 \epsilon
\end{aligned}$$

where the second equality follows from the fact that $[F] \perp ([z_i])_{i=0}^\infty$; and the first inequality follows from Cauchy-Schwarz Inequality. Then, $\| [F] \|_2 \leq \epsilon$. By the arbitrariness of ϵ , we have $\| [F] \|_2 = 0$. Then $[F] = \vartheta_{\mathcal{Z}}$. Since F is continuous, then $F(t) = 0$, $\forall t \in I$. Then, by Fundamental Theorem of Calculus I 12.88, $z = F^{(1)}$ a.e. in $\mathbb{I}$. Then, $[z] = \vartheta_{\mathcal{Z}}$.

Then, by Proposition 13.33, $([e_i])_{i=0}^\infty$ is complete. $\diamond$

Proposition 13.35 *Let $S_1 = [0, 2\pi] \subset \mathbb{R}$ and $z \in \mathcal{C}(S_1, \mathbb{C}) =: \mathcal{Z}$. Assume that $z(0) = z(2\pi)$. Let $M = \{\bar{z} \in \mathcal{Z} \mid \bar{z}(x) = \cos(nx), \forall x \in S_1 \text{ or } \bar{z}(x) = \sin(nx), \forall x \in S_1, \text{ where } n \in \mathbb{Z}_+\} \subseteq \mathcal{Z}$. Let $\mathcal{M} = \text{span}(M) \subseteq \mathcal{Z}$. Then, $z \in \overline{\mathcal{M}}$.*

Proof Let $S_2 \subset \mathbb{R}^2$ be the unit circle: $S_2 := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1^2 + y_2^2 = 1\}$. Define a mapping $\Psi : S_1 \rightarrow S_2$ by $\Psi(x) = (\cos(x), \sin(x))$, $\forall x \in S_1$. Clearly, Ψ is surjective and continuous. Note that S_1 is compact and S_2 is Hausdorff. It is obvious that we may define a function $\Phi : S_2 \rightarrow \mathbb{C}$ such that $\Phi \circ \Psi = z$. By Proposition 5.18, we have Φ is continuous. Note that S_2 is closed and bounded in $\mathbb{R}^2$, then S_2 is compact by Proposition 5.40. Hence, $\Phi \in \mathcal{C}(S_2, \mathbb{C})$. Let $\Phi_r := \text{Re} \circ \Phi \in \mathcal{C}(S_2, \mathbb{R})$ and $\Phi_i := \text{Im} \circ \Phi \in \mathcal{C}(S_2, \mathbb{R})$. By Tietze Extension Theorem 3.57, there exists $\bar{\Phi}_r \in \mathcal{C}(I^2, \mathbb{R})$ and $\bar{\Phi}_i \in \mathcal{C}(I^2, \mathbb{R})$, where $I = [-1, 1] \subset \mathbb{R}$, such that $\bar{\Phi}_r|_{S_2} = \Phi_r$ and $\bar{\Phi}_i|_{S_2} = \Phi_i$. Let $\bar{\Phi} := \bar{\Phi}_r + i\bar{\Phi}_i \in \mathcal{C}(I^2, \mathbb{C})$. Then, $\bar{\Phi}|_{S_2} = \Phi$. By Bernstein Approximation Theorem 11.203, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists a polynomial $P : I^2 \rightarrow \mathbb{C}$ such that $|\bar{\Phi}(y_1, y_2) - P(y_1, y_2)| < \epsilon$, $\forall y_1, y_2 \in I^2$. Then, $|z(x) - P \circ \Psi(x)| = |\bar{\Phi} \circ \Psi(x) - P \circ \Psi(x)| < \epsilon$, $\forall x \in S_1$. Note that $P \circ \Psi \in \mathcal{M}$, since, $\forall \gamma, \theta \in \mathbb{R}$,

$$\begin{aligned} (\sin(\theta))^2 &= \frac{1}{2} (1 - \cos(2\theta)); & (\cos(\theta))^2 &= \frac{1}{2} (1 + \cos(2\theta)); \\ \sin(\gamma) \cos(\theta) &= \frac{1}{2} (\sin(\gamma + \theta) + \sin(\gamma - \theta)); \\ \sin(\gamma) \sin(\theta) &= \frac{1}{2} (\cos(\gamma - \theta) - \cos(\gamma + \theta)); \\ \cos(\gamma) \cos(\theta) &= \frac{1}{2} (\cos(\gamma - \theta) + \cos(\gamma + \theta)) \end{aligned}$$

Hence, $z \in \overline{\mathcal{M}}$, by the arbitrariness of ϵ . This completes the proof of the proposition. $\square$

Given independent vectors $y_1, \dots, y_n$ generating a subspace M of a Hilbert space $\mathcal{X}$, we wish to find $m_0 \in M$ that minimizes $\|x - m_0\|$, for some given $x \in \mathcal{X}$. Rather than solving the normal equation (13.1), we may employ the Gram-Schmidt orthogonalization procedure together with Proposition 13.31. First use the Gram-Schmidt Procedure to obtain orthonormal sequence $(e_i)_{i=1}^n$. Then, $m_0 = \sum_{i=1}^n \langle x, e_i \rangle e_i$. Thus, the optimization problem is easily solved once the vectors y_i 's are orthogonalized.

Since solution to the approximation problem is equivalent to solution to the normal equation. It is clear that the Gram-Schmidt procedure can be interpreted as a procedure for inverting the Gram matrix.

It is interesting to note that the Gram-Schmidt procedure may itself be viewed as an approximation problem.

$$e_1 = \frac{y_1}{\|y_1\|}; \quad e_k = \frac{y_k - \sum_{i=1}^{k-1} \langle y_k, e_i \rangle e_i}{\left\| y_k - \sum_{i=1}^{k-1} \langle y_k, e_i \rangle e_i \right\|}; \quad \forall k > 1$$

The vector $y_k - \sum_{i=1}^{k-1} \langle y_k, e_i \rangle e_i$ is the optimal error for the problem $\min_{m \in \text{span}(\{y_1, \dots, y_{k-1}\})} \|y_k - m\|$. Thus, the Gram-Schmidt procedure consists of solving a series of minimum norm approximation problems by use of the Projection Theorem. Alternatively, the minimum norm approximation of x can be found by applying the Gram-Schmidt procedure to the sequence $(y_1, \dots, y_n, x)$. The optimal error $x - m_0$ is found at the last step.

13.6 Other Minimum Norm Problems

In approximation problems, finite dimensionality of the subspace allows the reduction of the problem into a finite-dimensional normal equation, which leads to a feasible computation procedure. In many important and interesting practical problems, the subspace is not finite-dimensional. In such problems, it is generally not possible to reduce the problem to a finite dimensional normal equation. However, there is an important class of such problems that can be reduced by the projection theorem to finite-dimensional equations. The motivation for this class of problems is illustrated in Figure 13.2. In this class of problems, the subspace $M^\perp$ is finite-dimensional, which leads to feasible computation procedures.

Next, we state a modified version of the Projection Theorem 13.13.

Theorem 13.36 (Restatement of Projection Theorem) *Let $\mathcal{X}$ be a Hilbert space, $M \subseteq \mathcal{X}$ be a closed subspace, $x, y \in \mathcal{X}$, $V = y + M$ be a closed linear variety. Then, there is a unique vector $v_0 \in V$ such that $\|x - v_0\| = \min_{v \in V} \|x - v\|$. A necessary and sufficient condition that $v_0 \in V$ is the unique minimizing vector is that $(x - v_0) \perp M$.*

Proof This is straightforward, and is therefore omitted. $\square$

A point of caution is necessary here: $(x - v_0) \perp M$ but not $(x - v_0) \perp V$. The concept of the above theorem is illustrated in Figure 13.3.

A special type of linear variety is $V = \{v \in \mathcal{X} \mid \langle v, y_i \rangle = c_i, i = 1, \dots, n\}$, where $y_1, \dots, y_n \in \mathcal{X}$ are fixed vectors and $c_1, \dots, c_n \in \mathbb{K}$ are fixed scalars. Let $M = \text{span}(\{y_1, \dots, y_n\})$. When $c_1 = \dots = c_n = 0$, then $V = M^\perp$. For arbitrary c_i 's, V is equal to $y + M^\perp$ assuming $V \neq \emptyset$. A linear variety of this form is said to be of finite *co-dimension*, since $M^{\perp\perp} = M$ is finite-dimensional by Proposition 13.17.

Theorem 13.37 *Let $\mathcal{X}$ be a Hilbert space over $\mathbb{K}$, $n \in \mathbb{N}$, $y_1, \dots, y_n \in \mathcal{X}$, $c_1, \dots, c_n \in \mathbb{K}$, $V = \{v \in \mathcal{X} \mid \langle v, y_i \rangle = c_i, i = 1, \dots, n\} \neq \emptyset$, $x \in \mathcal{X}$. Then, the following equation*

$$(\text{Gram}(y_1, \dots, y_n))' \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} \langle x, y_1 \rangle - c_1 \\ \vdots \\ \langle x, y_n \rangle - c_n \end{bmatrix} \quad (13.2)$$

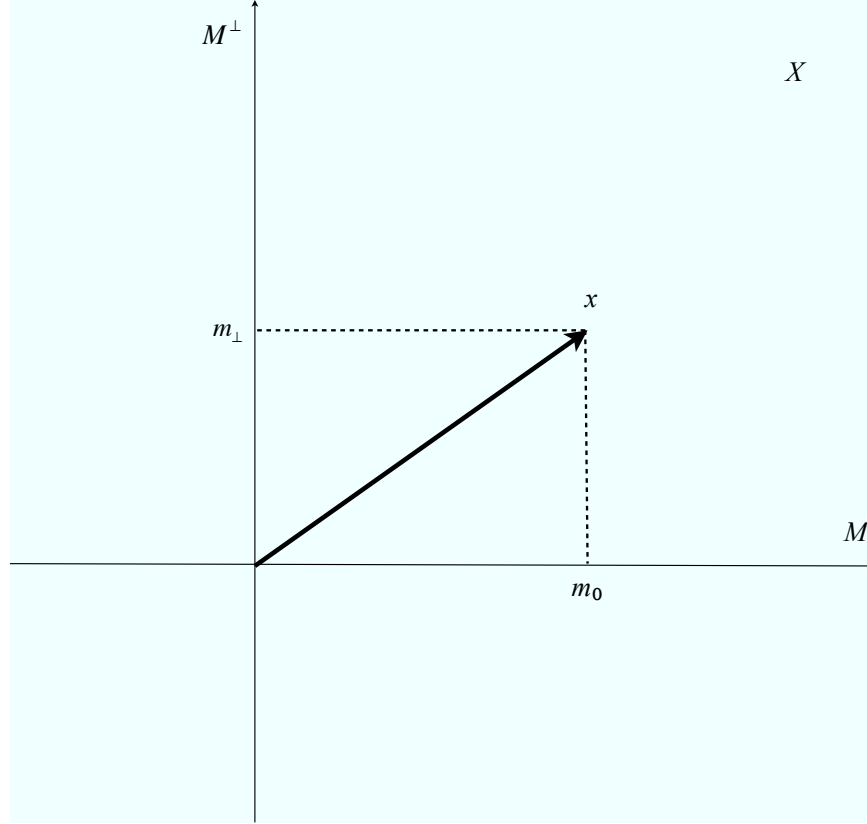


Figure 13.2: Dual projection problems.

admits at least one solution $(\beta_1, \dots, \beta_n) \in \mathbb{K}^n$. Any such solution implies that $v_0 = x - \sum_{i=1}^n \beta_i y_i \in V$ is the unique minimizing vector for $\min_{v \in V} \|x - v\|$.

Proof Let $M := \text{span}(\{y_1, \dots, y_n\})$. By Theorem 7.36 and Proposition 4.39, M is a closed subspace of $\mathcal{X}$. Then, $M^\perp = \{\bar{x} \in \mathcal{X} \mid \langle \bar{x}, y_i \rangle = 0, i = 1, \dots, n\}$. Fix any $y_0 \in V \neq \emptyset$, we have $V = y_0 + M^\perp$. By Theorem 13.36, $\exists! v_0 \in V$ such that $\|x - v_0\| = \min_{v \in V} \|x - v\|$ and a necessary and sufficient condition for any $v \in V$ to be v_0 is that $(x - v_0) \perp M^\perp$. Then, $(x - v_0) \in M^{\perp\perp} = M$, by Proposition 13.17. Then, $\exists(\beta_1, \dots, \beta_n) \in \mathbb{K}^n$ such that $x - v_0 = \sum_{i=1}^n \beta_i y_i$. Then, $v_0 = x - \sum_{i=1}^n \beta_i y_i$. Since $v_0 \in V$, we must have $\langle v_0, y_i \rangle = c_i, i = 1, \dots, n$. This leads to (13.2).

Any solution $(\beta_1, \dots, \beta_n) \in \mathbb{K}^n$ to (13.2), we have $v_0 \in V$ and $(x - v_0) \in M = M^{\perp\perp}$. By Theorem 13.36, v_0 is the unique minimizing vector for $\min_{v \in V} \|x - v\|$. This completes the proof of the theorem. $\square$

Much of the discussion in the above can be generalized from linear

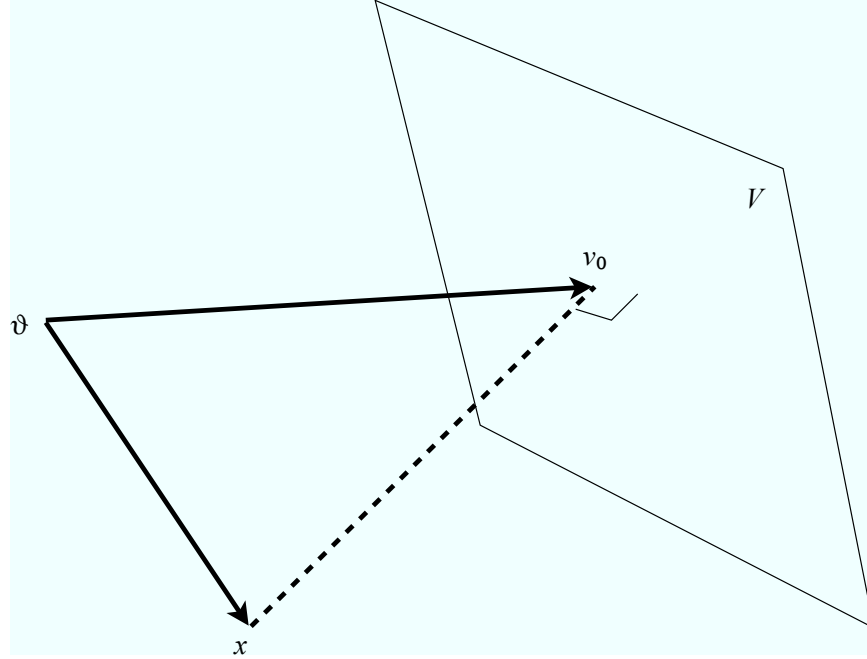


Figure 13.3: Reformulation of the Projection Theorem.

varieties to convex sets.

Theorem 13.38 *Let $\mathcal{X}$ be a real Hilbert space, $x \in \mathcal{X}$, $K \subseteq \mathcal{X}$ be a nonempty closed convex subset. Then, there is a unique vector $k_0 \in K$ such that $\|x - k_0\| = \min_{k \in K} \|x - k\|$. Furthermore, a necessary and sufficient condition for $k_0 \in K$ being the unique minimizing vector is that $\langle x - k_0, k - k_0 \rangle \leq 0, \forall k \in K$.*

The main idea of this theorem is illustrated in Figure 13.4, which shows that the angle between $x - k_0$ and $k - k_0$ is greater than or equal to 90° .

Proof First, we show the existence of k_0 . Let $\delta := \inf_{k \in K} \|x - k\| \in [0, \infty) \subset \mathbb{R}$ since $K \neq \emptyset$. Then, $\exists (k_i)_{i=1}^\infty \subseteq K$ such that $\lim_{i \in \mathbb{N}} \|x - k_i\| = \delta$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$, $\forall i \in \mathbb{N}$ with $n_0 \leq i$, we have $\delta \leq \|x - k_i\| < \sqrt{\delta^2 + \epsilon^2/4}$. $\forall i, j \in \mathbb{N}$ with $n_0 \leq i \leq j$, by Parallelogram Law 13.3, $\|(x - k_i) - (x - k_j)\|^2 + \|(x - k_i) + (x - k_j)\|^2 = 2\|x - k_i\|^2 + 2\|x - k_j\|^2$. This implies that $\|k_i - k_j\|^2 = 2\|x - k_i\|^2 + 2\|x - k_j\|^2 - 4\|x - (k_i + k_j)/2\|^2 < 4(\delta^2 + \epsilon^2/4) - 4\delta^2 = \epsilon^2$, where the inequality follows from the fact that K is convex and $(k_i + k_j)/2 \in K$. Hence, $(k_i)_{i=1}^\infty \subseteq K$ is Cauchy sequence. By Proposition 4.39, K is complete and $\lim_{i \in \mathbb{N}} k_i = k_0 \in K$. By Propositions 3.66 and 7.21, $\delta = \lim_{i \in \mathbb{N}} \|x - k_i\| = \|x - k_0\|$.

Next, we show the uniqueness of k_0 . Let $\bar{k} \in K$ be such that $\delta = \|x - \bar{k}\|$.

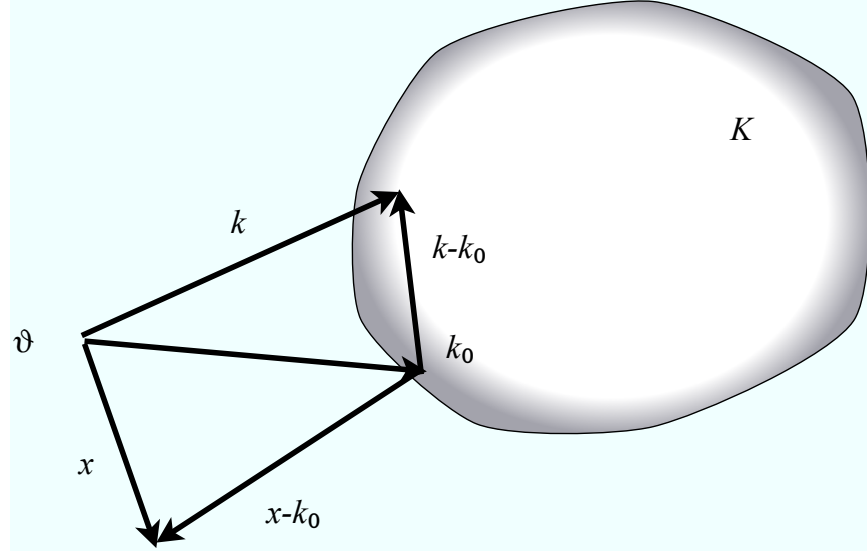


Figure 13.4: Projection to a convex set.

The sequence $(k_i)_{i=1}^\infty := (k_0, \bar{k}, k_0, \bar{k}, \dots) \subseteq K$ satisfies $\lim_{i \in \mathbb{N}} \|x - k_i\| = \delta$. By the proof of the existence, $(k_i)_{i=1}^\infty$ is convergent. Then, we must have $\bar{k} = k_0$. Hence, k_0 is unique.

Next, we show that $\langle x - k_0, k - k_0 \rangle \leq 0, \forall k \in K$ by an argument of contradiction. Suppose this is not true. Then, $\exists k_1 \in K$ such that $\langle x - k_0, k_1 - k_0 \rangle =: \lambda > 0$. Consider the vector $k_\alpha := (1 - \alpha)k_0 + \alpha k_1$, where $\alpha \in (0, 1) \subset \mathbb{R}$ is to be determined. Since K is convex, then $k_\alpha \in K$. Note that $\|x - k_\alpha\|^2 = \|(x - k_0) - \alpha(k_1 - k_0)\|^2 = \|x - k_0\|^2 - 2\alpha \langle x - k_0, k_1 - k_0 \rangle + \alpha^2 \|k_1 - k_0\|^2 = \delta^2 - 2\alpha\lambda + \alpha^2 \|k_1 - k_0\|^2$. Then, for sufficiently small $\alpha \in (0, 1) \subset \mathbb{R}$, we have $\|x - k_\alpha\|^2 < \delta^2$, which contradicts with the definition of δ . Hence, $\langle x - k_0, k - k_0 \rangle \leq 0, \forall k \in K$.

Finally, we show that if $\bar{k} \in K$ satisfies $\langle x - \bar{k}, k - \bar{k} \rangle \leq 0, \forall k \in K$, then $\|x - \bar{k}\| = \min_{k \in K} \|x - k\|$. $\forall k \in K$, we have $\|x - k\|^2 = \|(x - \bar{k}) - (k - \bar{k})\|^2 = \|x - \bar{k}\|^2 - 2\langle x - \bar{k}, k - \bar{k} \rangle + \|k - \bar{k}\|^2 \geq \|x - \bar{k}\|^2$. Hence, $\|x - \bar{k}\| \leq \|x - k\|, \forall k \in K$.

This completes the proof of the theorem. $\square$

Chapter 14

Probability Theory

14.1 Fundamental Notions

Definition 14.1 Let $\Omega := (\Omega, \mathcal{B}, \mathbf{P})$ be a finite measure sapce with $\mathbf{P}(\Omega) = 1$. Then, it is said to be a probability measure space. Let $\mathcal{Y}$ be a separable reflexive Banach space over $\mathbb{K}$. A $\mathcal{Y}$ -valued random variable is a $\mathcal{B}$ -measurable function $x : \Omega \rightarrow \mathcal{Y}$. When $\mathcal{Y} = \mathbb{R}$, we will simply say x is a random variable. The integral $\int_{\Omega} x d\mathbf{P}$ is said to be the expectation of x , when it makes sense, which will be denoted by $\mathbf{E}(x)$. Let $\hat{\mathcal{B}} \subseteq \mathcal{B}$ be a σ -algebra on Ω and $\hat{\mathbf{P}} := \mathbf{P}|_{\hat{\mathcal{B}}}$. By Proposition 11.13, $\hat{\Omega} := (\Omega, \hat{\mathcal{B}}, \hat{\mathbf{P}})$ is a probability measure space. Let $\tilde{\mathcal{B}} \subseteq \mathcal{B}$ be another σ -algebra on Ω . We will say that $\hat{\mathcal{B}}$ and $\tilde{\mathcal{B}}$ are independent if, $\forall \hat{E} \in \hat{\mathcal{B}}, \forall \tilde{E} \in \tilde{\mathcal{B}}$, we have $\mathbf{P}(\hat{E} \cap \tilde{E}) = \mathbf{P}(\hat{E})\mathbf{P}(\tilde{E})$. Assume that $x \in \bar{\mathbf{L}}_1(\Omega, \mathcal{Y})$. Define $\nu : \mathcal{B} \rightarrow \mathcal{Y}$ by $\nu(E) = \int_E x d\mathbf{P}, \forall E \in \mathcal{B}$. By Proposition 11.116, ν is a finite $\mathcal{Y}$ -valued measure on $(\Omega, \mathcal{B})$ with $\mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ x d\mathbf{P}, \forall E \in \mathcal{B}$. By Proposition 11.107, $\hat{\nu} := \nu|_{\hat{\mathcal{B}}}$ is a $\mathcal{Y}$ -valued pre-measure on $(\Omega, \hat{\mathcal{B}})$. By Definition 11.99, we have $\mathcal{P} \circ \hat{\nu} \leq (\mathcal{P} \circ \nu)|_{\hat{\mathcal{B}}}$. Then, $(\Omega, \hat{\mathcal{B}}, \hat{\nu})$ is a finite $\mathcal{Y}$ -valued measure space. $\forall E \in \hat{\mathcal{B}}$ with $\hat{\mathbf{P}}(E) = 0$, we have $E \in \mathcal{B}, \mathbf{P}(E) = 0$, and $\mathcal{P} \circ \hat{\nu}(E) \leq \mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ x d\mathbf{P} = 0$. Hence, $\mathcal{P} \circ \hat{\nu} \ll \hat{\mathbf{P}}$. By Radon-Nikodym Theorem 11.171, there exists $f : \Omega \rightarrow \mathcal{Y}$ that is the Radon-Nikodym derivative of $\hat{\nu}$ with respect to $\hat{\mathbf{P}}$, where f is $\hat{\mathcal{B}}$ -measurable. The function f is unique as described in Proposition 11.167 and $f \in \bar{\mathbf{L}}_1(\hat{\Omega}, \mathcal{Y})$. Then, $[f] \in \mathbf{L}_1(\hat{\Omega}, \mathcal{Y})$ is said to be the conditional expectation of x given $\hat{\mathcal{B}}$ and is denoted by $\mathbf{E}(x|\hat{\mathcal{B}})$. $f \in \mathbf{E}(x|\hat{\mathcal{B}})$ is a $\mathcal{Y}$ -valued random variable and is said to be a version of the conditional expectation. $\forall E \in \hat{\mathcal{B}}$, we have

$$\begin{aligned} \int_E x d\mathbf{P} &= \nu(E) = \hat{\nu}(E) = \int_E f d\hat{\mathbf{P}} = \int_E f d\mathbf{P} \\ \int_E \mathcal{P} \circ f d\hat{\mathbf{P}} &= \mathcal{P} \circ \hat{\nu}(E) = \int_E \mathcal{P} \circ f d\mathbf{P} \leq \mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ x d\mathbf{P} < \infty \end{aligned}$$

Definition 14.2 Let $(\Omega, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a topological space, $\mathcal{F} \subseteq \mathcal{B}$, and $x : \Omega \rightarrow \mathcal{Y}$ be $\mathcal{B}$ -measurable. We will denote the σ -algebra generated by $\mathcal{F}$ by $\sigma(\mathcal{F}) \subseteq \mathcal{B}$; and denote the σ -algebra generated by $\{E \in \mathcal{B} \mid E = x_{\text{inv}}(O), O \in \mathcal{O}_{\mathcal{Y}}\}$ by $\sigma(x) \subseteq \mathcal{B}$. Clearly, $\sigma(x)$ is the smallest σ -algebra on which x is measurable.

Proposition 14.3 Let $\Omega := (\Omega, \mathcal{B}, \mathbf{P})$ be a probability measure space, $\mathcal{Y}$ be a separable reflexive Banach space over $\mathbb{K}$, $x \in \bar{L}_1(\Omega, \mathcal{Y})$, $(x_n)_{n=1}^{\infty} \subseteq \bar{L}_1(\Omega, \mathcal{Y})$, $\hat{\mathcal{B}} \subseteq \mathcal{B}$ be a σ -algebra on Ω , $\hat{\Omega} := (\Omega, \hat{\mathcal{B}}, \hat{\mathbf{P}} := \mathbf{P}|_{\hat{\mathcal{B}}})$, $f \in \mathbf{E}(x|\hat{\mathcal{B}})$, and $f_n \in \mathbf{E}(x_n|\hat{\mathcal{B}})$, $\forall n \in \mathbb{N}$. Then, the following statements hold.

- (i) $\mathbf{E}(f) = \mathbf{E}(x)$.
- (ii) Let x be $\hat{\mathcal{B}}$ -measurable, then $[x] = \mathbf{E}(x|\hat{\mathcal{B}}) \in L_1(\hat{\Omega}, \mathcal{Y})$.
- (iii) Let $\alpha, \beta \in \mathbb{K}$, then $\mathbf{E}(\alpha x_1 + \beta x_2|\hat{\mathcal{B}}) = \alpha \mathbf{E}(x_1|\hat{\mathcal{B}}) + \beta \mathbf{E}(x_2|\hat{\mathcal{B}})$.
- (iv) Let $F \subseteq \mathcal{Y}$ be a nonempty closed convex set with $\vartheta_{\mathcal{Y}} \in F$, and $x : \Omega \rightarrow F$, then $f(\omega) \in F$ a.e. $\omega \in \hat{\Omega}$.
- (v) Let $\mathcal{Y} = \mathbb{R}$, $0 \leq x_n(\omega) \leq x_{n+1}(\omega) \leq x(\omega)$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} x_n(\omega) = x(\omega)$, $\forall \omega \in \Omega$, then $f_n \leq f_{n+1}$ a.e. in $\hat{\Omega}$ and $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\hat{\Omega}$.
- (vi) Let $\mathcal{Y} = \mathbb{R}$, $0 \leq x_n(\omega) < \infty$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$, and $\bar{f} \in \mathbf{E}(\liminf_{n \in \mathbb{N}} x_n|\hat{\mathcal{B}})$, then $\bar{f} \leq \liminf_{n \in \mathbb{N}} f_n$ a.e. in $\hat{\Omega}$.
- (vii) Let $g : \Omega \rightarrow [0, \infty) \subset \mathbb{R}$ be integrable over Ω , $\mathcal{P} \circ x_n(\omega) \leq g(\omega)$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} x_n = x$ a.e. in Ω , then $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\hat{\Omega}$.
- (viii) Let $F \subseteq \mathcal{Y}$ be a nonempty closed convex set, $x : \Omega \rightarrow F$, $c : F \rightarrow \mathbb{R}$ be a convex functional with $[c, F]$ being closed, $c \circ x : \Omega \rightarrow \mathbb{R}$ being absolutely integrable over Ω , and $\bar{f} \in \mathbf{E}(c \circ x|\hat{\mathcal{B}})$, then $c \circ f \leq \bar{f}$ a.e. in $\hat{\Omega}$.
- (ix) Let $\tilde{\mathcal{B}} \subseteq \hat{\mathcal{B}}$ be a σ -algebra on Ω , then $\mathbf{E}(f|\tilde{\mathcal{B}}) = \mathbf{E}(x|\tilde{\mathcal{B}})$.
- (x) Let $\mathcal{Z}$ be a separable reflexive Banach space over $\mathbb{K}$, $A : \Omega \rightarrow B(\mathcal{Y}, \mathcal{Z})$ be $\hat{\mathcal{B}}$ -measurable, and Ax be absolutely integrable over Ω , and $\bar{f} \in \mathbf{E}(Ax|\hat{\mathcal{B}})$, then $\bar{f} = Af$ a.e. in $\hat{\Omega}$.
- (xi) Let $\tilde{\mathcal{B}} \subseteq \mathcal{B}$ be a σ -algebra on Ω , $\hat{\mathcal{B}}$ and $\sigma(\tilde{\mathcal{B}} \cup \sigma(x))$ are independent, then $f \in \mathbf{E}(x|\sigma(\tilde{\mathcal{B}} \cup \hat{\mathcal{B}}))$.

Chapter 15

Numerical Methods

15.1 Newton's Method

Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces and $F : \mathcal{X} \rightarrow \mathcal{Y}$. To solve the equation $F(x) = \vartheta_{\mathcal{Y}}$, we may use Newton's method. Assume that F is twice Fréchet differentiable, we set $x_{n+1} = x_n - (F^{(1)}(x_n))^{-1}F(x_n)$, $\forall n \in \mathbb{N}$. Then, when x_1 is sufficiently close to a solution, the sequence $(x_n)_{n=1}^{\infty}$ converges to a solution x_{opt} with $F(x_{\text{opt}}) = \vartheta_{\mathcal{Y}}$. This result is formalized in the following proposition.

Proposition 15.1 *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D_1 \subseteq \mathcal{X}$, $F : D_1 \rightarrow \mathcal{Y}$, and $x_1 \in \mathcal{X}$. Assume that $\exists \beta_1, \eta_1, h_1, K \in [0, \infty) \subset \mathbb{R}$ such that*

- (i) $h_1 := \beta_1 \eta_1 K \leq \frac{1}{2}$;
- (ii) $\text{dom}(F^{(2)}) \supseteq \overline{\mathcal{B}}_{\mathcal{X}}\left(x_1, \frac{2-2h_1}{2-3h_1}\eta_1\right) =: D \subseteq \overline{\mathcal{B}}_{\mathcal{X}}(x_1, 2\eta_1)$ and, $\forall x \in D$, $\|F^{(2)}(x)\| \leq K$;
- (iii) $F^{(1)}(x_1)$ is bijective, $\|(F^{(1)}(x_1))^{-1}\| \leq \beta_1$, and $\|(F^{(1)}(x_1))^{-1} \cdot F(x_1)\| \leq \eta_1$;

Then, $\exists (x_n)_{n=1}^{\infty} \subseteq D$, defined by $x_{n+1} = x_n - (F^{(1)}(x_n))^{-1}F(x_n)$, $\forall n \in \mathbb{N}$, that converges to $x_{\text{opt}} \in D$ with $F(x_{\text{opt}}) = \vartheta_{\mathcal{Y}}$.

Proof We will use mathematical induction to prove the following claim.

Claim 15.1.1 $\forall n \in \mathbb{N}$, we have

- (a) $F^{(1)}(x_n)$ is bijective and $\|(F^{(1)}(x_n))^{-1}\| \leq \beta_n := \frac{\beta_{n-1}}{1-h_{n-1}} \in [0, \infty) \subset \mathbb{R}$;
- (b) $\|(F^{(1)}(x_n))^{-1}F(x_n)\| \leq \eta_n := \frac{h_{n-1}\eta_{n-1}}{2(1-h_{n-1})} \in [0, \infty) \subset \mathbb{R}$;

$$(c) \ x_{n+1} \in \overline{B}_y \left(x_1, \frac{2-2h_1}{2-3h_1} \left(1 - \left(\frac{h_1}{2(1-h_1)} \right)^n \right) \eta_1 \right);$$

$$(d) \ 0 \leq h_n := \beta_n \eta_n K \leq h_{n-1} \leq \frac{1}{2}.$$

Proof of claim: 1° Clearly, (a) – (d) are satisfied for $n = 1$.

2° Assume that (a) – (d) are satisfied for $n \leq k \in \mathbb{N}$.

3° Consider the case of $n = k + 1$. By Mean Value Theorem 9.23, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ such that $\|F^{(1)}(x_{k+1}) - F^{(1)}(x_k)\| \leq \|F^{(2)}(t_0 x_{k+1} + (1 - t_0)x_k)(x_{k+1} - x_k)\| \leq \|F^{(2)}(t_0 x_{k+1} + (1 - t_0)x_k)\| \|x_{k+1} - x_k\| \leq K\eta_k$. Then, $\|(F^{(1)}(x_k))^{-1}\| \|F^{(1)}(x_{k+1}) - F^{(1)}(x_k)\| \leq \beta_k \eta_k K = h_k < 1$. By Proposition 9.55, $F^{(1)}(x_{k+1})$ is bijective and

$$\begin{aligned} & \| (F^{(1)}(x_{k+1}))^{-1} - (F^{(1)}(x_k))^{-1} \| \\ & \leq \frac{\| (F^{(1)}(x_k))^{-1} \|^2 \|F^{(1)}(x_{k+1}) - F^{(1)}(x_k)\|}{1 - \| (F^{(1)}(x_k))^{-1} \| \|F^{(1)}(x_{k+1}) - F^{(1)}(x_k)\|} \\ & \leq \frac{\beta_k^2 K \eta_k}{1 - \beta_k \eta_k K} = \frac{h_k \beta_k}{1 - h_k} \end{aligned}$$

This leads to $\| (F^{(1)}(x_{k+1}))^{-1} \| \leq \| (F^{(1)}(x_k))^{-1} \| + \| (F^{(1)}(x_{k+1}))^{-1} - (F^{(1)}(x_k))^{-1} \| \leq \beta_k + h_k \beta_k / (1 - h_k) = \beta_k / (1 - h_k) = \beta_{k+1}$. Hence, (a) holds.

Note that $F(x_{k+1}) = F(x_{k+1}) - F(x_k) - F^{(1)}(x_k)(x_{k+1} - x_k)$. By Taylor's Theorem 9.48, $\exists t_1 \in (0, 1) \subset \mathbb{R}$ such that $\|F(x_{k+1})\| \leq \frac{1}{2} \|F^{(2)}(t_1 x_{k+1} + (1 - t_1)x_k)\| \|x_{k+1} - x_k\|^2 \leq K\eta_k^2/2$. Then, $\| (F^{(1)}(x_{k+1}))^{-1} F(x_{k+1}) \| \leq \beta_k K \eta_k^2 / (2 - 2h_k) = \frac{h_k \eta_k}{2(1-h_k)} = \eta_{k+1}$. Hence, (b) holds.

Note that $\|x_{k+2} - x_{k+1}\| = \| (F^{(1)}(x_{k+1}))^{-1} F(x_{k+1}) \| \leq \eta_{k+1} = \frac{h_k}{2(1-h_k)} \eta_k \leq \frac{h_1}{2(1-h_1)} \eta_k \leq \left(\frac{h_1}{2(1-h_1)} \right)^k \eta_1$. Then, $\|x_{k+2} - x_1\| \leq \|x_{k+2} - x_{k+1}\| + \|x_{k+1} - x_1\| \leq \left(\frac{h_1}{2(1-h_1)} \right)^k \eta_1 + \frac{2-2h_1}{2-3h_1} \left(1 - \left(\frac{h_1}{2(1-h_1)} \right)^k \right) \eta_1 = \frac{2-2h_1}{2-3h_1} \left(1 - \left(\frac{h_1}{2(1-h_1)} \right)^{k+1} \right) \eta_1$. Hence, (c) holds.

Note that $h_{k+1} = \beta_{k+1} \eta_{k+1} K = \frac{h_k^2}{2(1-h_k)^2}$. Since $0 \leq h_k \leq 1/2$, then $0 \leq \frac{h_k}{2(1-h_k)^2} \leq 1$. This implies that $0 \leq h_{k+1} \leq h_k \leq 1/2$. Hence, (d) holds.

This completes the induction process and the proof of the claim. $\square$

Then, the sequence $(x_n)_{n=1}^\infty \subseteq D$ is well-defined. $\forall n \in \mathbb{N}$, $\|x_{n+2} - x_{n+1}\| = \| (F^{(1)}(x_{n+1}))^{-1} F(x_{n+1}) \| \leq \eta_{n+1} \leq \left(\frac{h_1}{2(1-h_1)} \right)^n \eta_1$. Then, $(x_n)_{n=1}^\infty$ is a Cauchy sequence since $0 \leq \frac{h_1}{2(1-h_1)} \leq 1/2 < 1$. It then converges to $x_{\text{opt}} \in D$ since D is closed and $\mathcal{X}$ is complete. Note that $F(x_n) = F^{(1)}(x_n)(x_n - x_{n+1})$, $\forall n \in \mathbb{N}$. By Proposition 9.7, we have

F and $F^{(1)}$ are continuous at x_{opt} . Then, $F(x_{\text{opt}}) = \lim_{n \in \mathbb{N}} F(x_n) = \lim_{n \in \mathbb{N}} F^{(1)}(x_n)(x_n - x_{n+1}) = F^{(1)}(x_{\text{opt}})(x_{\text{opt}} - x_{\text{opt}}) = \vartheta_{\mathcal{Y}}$, where the first equality follows from Proposition 3.66; and the third equality follows from Propositions 3.66, 3.67, 7.23, and 7.65. This completes the proof of the proposition. $\square$

It is numerically expensive to use Newton's Method due to the necessity of inverting $F^{(1)}(x_n)$ at each step of the iteration. But the benefit of this method is the quadratic convergence of $(x_n)_{n=1}^{\infty}$ to x_{opt} , which is summarized in the following proposition.

Proposition 15.2 *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D_1 \subseteq \mathcal{X}$, $F : D_1 \rightarrow \mathcal{Y}$, and $x_1 \in \mathcal{X}$. Assume that $\exists \beta_1, \eta_1, h_1, K, M_1, M_2, M_4 \in [0, \infty) \subset \mathbb{R}$ such that*

$$(i) \quad h_1 := \beta_1 \eta_1 K \leq \frac{1}{2};$$

$$(ii) \quad \text{dom}(F^{(3)}) \supseteq D_2 \supseteq \overline{\mathcal{B}}_{\mathcal{X}}\left(x_1, \frac{2-2h_1}{2-3h_1}\eta_1\right) =: D, \text{ where } D_2 \subseteq D_1 \text{ is an open set in } \mathcal{X}, \forall x \in D_2, F^{(1)}(x) \text{ is bijective, and, } \forall x \in D, \\ \|F^{(2)}(x)\| \leq K, \| (F^{(1)}(x))^{-1} \| \leq M_1, \|F^{(3)}(x)\| \leq M_2, \text{ and } \|F(x)\| \leq M_4;$$

$$(iii) \quad \| (F^{(1)}(x_1))^{-1} \| \leq \beta_1 \text{ and } \| (F^{(1)}(x_1))^{-1} F(x_1) \| \leq \eta_1;$$

Then, $\exists (x_n)_{n=1}^{\infty} \subseteq D$, defined by $x_{n+1} = x_n - (F^{(1)}(x_n))^{-1} F(x_n)$, $\forall n \in \mathbb{N}$, that converges to $x_{\text{opt}} \in D$ with $F(x_{\text{opt}}) = \vartheta_{\mathcal{Y}}$. Furthermore, $\forall n \in \mathbb{N}$, $\|x_{n+1} - x_{\text{opt}}\| \leq c \|x_n - x_{\text{opt}}\|^2$, where $c := (M_1 K + 2M_1^3 K^2 M_4 + M_1^2 M_2 M_4)/2$.

Proof By Proposition 15.1, $(x_n)_{n=1}^{\infty} \subseteq D$ is well defined and converges to $x_{\text{opt}} \in D$ with $F(x_{\text{opt}}) = \vartheta_{\mathcal{Y}}$. Define $T : D_2 \rightarrow \mathcal{X}$ by $T(x) = x - (F^{(1)}(x))^{-1} F(x)$, $\forall x \in D_2$. Then, by Propositions 9.34, 9.44, 9.45, 9.41, 9.40, and 9.55, T is twice Fréchet differentiable. This leads to $T^{(1)}(x)(h_1) = h_1 - (F^{(1)}(x))^{-1} F^{(1)}(x)(h_1) + (F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F(x)) = (F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F(x))$, $\forall x \in D_2$, $\forall h_1 \in \mathcal{X}$. Then, $T^{(1)}(x_{\text{opt}})(h_1) = \vartheta_{\mathcal{X}}$, $\forall h_1 \in \mathcal{X}$, since $F(x_{\text{opt}}) = \vartheta_{\mathcal{Y}}$. This implies that $T^{(1)}(x_{\text{opt}}) = \vartheta_{\mathcal{B}(\mathcal{X}, \mathcal{X})}$. The second order derivative of T is given by, $\forall x \in D_2$, $\forall h_1, h_2 \in \mathcal{X}$,

$$\begin{aligned} T^{(2)}(x)(h_2)(h_1) &= -(F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} \\ &\quad \cdot F(x))) + (F^{(1)}(x))^{-1} F^{(3)}(x)(h_2)(h_1)((F^{(1)}(x))^{-1} F(x)) \\ &\quad + (F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1} \\ &\quad \cdot F(x)) + (F^{(1)}(x))^{-1} F^{(1)}(x)(h_2)) \\ &= -(F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} \\ &\quad \cdot F(x))) + (F^{(1)}(x))^{-1} F^{(3)}(x)(h_2)(h_1)((F^{(1)}(x))^{-1} F(x)) \end{aligned}$$

$$\begin{aligned}
& -(F^{(1)}(x))^{-1}F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1}F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1}F(x))) \\
& + (F^{(1)}(x))^{-1}F^{(2)}(x)(h_1)(h_2)
\end{aligned}$$

Then, $\forall x \in D$, $\|T^{(2)}(x)\| = \sup_{h_1, h_2 \in X, \|h_1\| \leq 1, \|h_2\| \leq 1} \|T^{(2)}(x)(h_2)(h_1)\| \leq 2M_1^3K^2M_4 + M_1^2M_2M_4 + M_1K = 2c$. $\forall n \in \mathbb{N}$, by Taylor's Theorem 9.48, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ such that $\|x_{n+1} - x_{\text{opt}}\| = \|T(x_n) - T(x_{\text{opt}})\| = \|T(x_n) - T(x_{\text{opt}}) - T^{(1)}(x_{\text{opt}})(x_n - x_{\text{opt}})\| \leq \frac{1}{2} \|T^{(2)}(t_0x_n + (1 - t_0)x_{\text{opt}})\| \|x_n - x_{\text{opt}}\|^2 \leq c \|x_n - x_{\text{opt}}\|^2$. This completes the proof of the proposition. $\square$

Appendix A

Elements in Calculus

A.1 Some Formulas

Proposition A.1 $\sum_{i=0}^n \prod_{j=1}^m (i+j) = \frac{1}{m+1} \frac{(n+m+1)!}{n!}, \forall n, m \in \mathbb{Z}_+.$

Proof We will use mathematical induction on $n \in \mathbb{Z}_+$ to prove the result. 1° $n = 0$. $\forall m \in \mathbb{Z}_+$, we have LHS = $m!$ = RHS. This case is proved.

2° Assume the result holds for $n = k \in \mathbb{Z}_+.$

3° Consider the case when $n = k + 1 \in \mathbb{N}$. $\forall m \in \mathbb{Z}_+,$

$$\begin{aligned} \text{LHS} &= \sum_{i=0}^{k+1} \prod_{j=1}^m (i+j) = \sum_{i=0}^k \prod_{j=1}^m (i+j) + \prod_{j=1}^m (k+1+j) \\ &= \frac{1}{m+1} \frac{(k+m+1)!}{k!} + \frac{(k+m+1)!}{(k+1)!} = \frac{1}{m+1} \frac{(k+m+2)!}{(k+1)!} \end{aligned}$$

where the third equality follow from the inductive assumption. This completes the induction process.

Hence, the result holds. $\square$

A.2 Convergence of Infinite Sequences

Proposition A.2 (Raabe's Test) Let $(x_n)_{n=1}^\infty \subseteq (0, \infty) \subset \mathbb{R}$. Then, the following statements hold.

- (i) If there exists $a \in (1, \infty) \subset \mathbb{R}$ and $n_0 \in \mathbb{N}$ such that $\frac{x_{n+1}}{x_n} \leq 1 - \frac{a}{n}, \forall n \geq n_0$, then $\sum_{n=1}^\infty x_n \in (0, \infty) \subset \mathbb{R}$.
- (ii) If there exists $a \in (-\infty, 1] \subset \mathbb{R}$ and $n_0 \in \mathbb{N}$ such that $\frac{x_{n+1}}{x_n} \geq 1 - \frac{a}{n}, \forall n \geq n_0$, then $\sum_{n=1}^\infty x_n = \infty$.

Proof (Bartle) (i) Under the assumption, $\forall n \geq n_0$,

$$nx_{n+1} \leq (n-1)x_n - (a-1)x_n$$

This implies that

$$(n-1)x_n - nx_{n+1} \geq (a-1)x_n > 0, \quad \forall n \geq n_0$$

Summing the above for $n = n_0, \dots, m$, $\forall m \geq n_0$, we have

$$0 < (a-1) \sum_{n=n_0}^m x_n \leq (n_0-1)x_{n_0} - mx_{m+1} < (n_0-1)x_{n_0} \in \mathbb{R}$$

Hence, we have $0 < \sum_{n=n_0}^m x_n < \frac{(n_0-1)x_{n_0}}{a-1} \in (0, \infty) \subset \mathbb{R}$, $\forall m \geq n_0$.

Hence, $\sum_{n=1}^{\infty} x_n \in (0, \sum_{n=1}^{n_0-1} x_n + \frac{(n_0-1)x_{n_0}}{a-1}] \subset \mathbb{R}$.

(ii) Under the assumption, we have

$$nx_{n+1} \geq (n-a)x_n \geq (n-1)x_n, \forall n \geq n_0$$

Then, $nx_{n+1} \geq n_0x_{n_0+1} =: c \in (0, \infty) \subset \mathbb{R}$, $\forall n \geq n_0$. This implies that $x_{n+1} \geq c/n$, $\forall n \geq n_0$, and hence $\sum_{n=1}^{\infty} x_n = \infty$.

This completes the proof of the proposition. $\square$

Bibliography

- Bartle, R. G. (1976). *The Elements Of Real Analysis*. 2nd ed. John Wiley & Sons, New York, NY.
- Luenberger, D. G. (1969). *Optimization by Vector Space Methods*. Wiley, New York.
- Mathematics Handbook Editors Group (ed.) (1979). *Mathematics Handbook*. Higher Education Publishing House, Beijing, China.
- Maunder, C. R. F. (1996). *Algebraic Topology*. Dover Publications, Mineola, N.Y.
URL: <http://www.loc.gov/catdir/description/dover031/95051359.html>
- Royden, H. L. (1988). *Real Analysis*. 3rd ed. Prentice Hall, Englewood Cliffs, NJ.
- Spivak, M. (1965). *Calculus on Manifolds*. Addison Wesley.
- Williams, D. (1991). *Probability with Martingales*. Cambridge University Press, New York, NY.

Index

- Absolute continuity of measure, 473
- Absolute integrability, 405
 - Banach space valued measure space, 438
- Adjoint operator, 203
- Admissible deviation, 253
- Alaoglu Theorem, 212
- Alexandroff One-Point Compactification, 120
- Algebra, 168
 - Boolean, 23
 - Generated by, 23
- Alignment, 194
- Almost everywhere, 370
 - Banach space valued measure, 419
 - Banach space valued pre-measure, 415
- Analytic radius, 313
- Antisymmetric, 20
- Ascoli-Arzelá Theorem, 111
- Baire, 86
 - Category Theorem, 87
- Ball
 - Closed, 73
 - Open, 73
- Banach measure space, 365
 - Banach space valued, 514
- Banach space, 158
- Banach space valued pre-measure, 414
 - On algebra, 538
- Basis
 - Topological space, 41
 - Vector space, 145
- Bernsteĭn Approximation Theorem, 528
- Bernsteĭn function, 527
- Bessel's Inequality, 651
- Bijjective, 21
- Bolzano-Weierstrass property, 104
- Borel measure, 594
- Borel sets, 361
- Borel-Cantelli Lemma, 351
- Borel-Lebesgue Theorem, 108
- Bound
 - Greatest lower, 25
 - Least upper, 25
 - Lower, 25
 - Upper, 25
- Bounded
 - Function, 153
 - Linear operator, 175
 - Set, 153
- Bounded Convergence Theorem, 391
- Bourbaki, 112
- Carathéodory Extension Theorem, 356, 534
- Cartesian metric, 80
- Cartesian product, 33
 - Normed linear space, 156
 - Vector space, 136
- Category
 - First, 48
 - Second, 48
 - Everywhere, 48
- Cauchy
 - Net, 87

- Sequence, 76
 - In measure, 377
- Cauchy-Schwarz Inequality, 150, 499, 637
- Chain Rule, 261
- Change of Variable, 543
- Choice, 26
- Closed Graph Theorem, 201, 202
- Commutative
 - Ring, 133
- Compact, 97, 163
 - Countably, 103
 - Locally, 114
 - Sequentially, 104
 - σ -, 123
- Complete, 76
 - Banach space valued measure space, 419
 - Banach space valued pre-measure space, 415
 - Measure space, 351
 - Metric measure space, 365
 - Banach space valued, 513
 - Sequence, 651
- Completely regular, 58
- Completion of a Banach space valued measure space, 421
- Completion of a Banach space valued pre-measure space, 416
- Completion of a measure space, 352
- Component, 51
- Composition, 21
- Concave functional, 225
- Conditional expectation, 659
- Cone, 143
 - Conjugate, 243
 - Negative, 243
 - Positive, 243, 244
- Conic segment, 143
- Conjugate
 - Group, 132
- Connected, 50
 - Arcwise, 54
 - Locally, 52
 - Simply, 54
- Continuity, 38
 - Absolute, 595
 - One-sided, 571
 - Semi-, 40
- Continuum, 102
- Contraction index, 288
- Contraction mapping, 288
- Contraction Mapping Theorem, 289
- Convergence in measure in $\mathcal{X}$, 376
 - Banach space valued measure, 419
- Convex
 - Functional, 225
 - Local, 214
 - Mapping, 244
- Convex combination, 142
- Convex hull, 141
- Coset, 165
- Countability
 - First axiom of, 41
 - Second axiom of, 41
- Countable, 23
- Covering, 42
- Cube, 93
 - Hilbert, 93
- Cumulative distribution function, 576
 - With origin, 587
- Curve, 54
 - Closed, 54
- Dense, 48
 - Nowhere, 48
- Derivative
 - Directional, 254
 - Fréchet, 254
 - Higher-order, 267
 - Partial, 255
 - Higher-order, 273
- Diameter, 612
- Dimension, 145
 - Co-, 655
- Dini's Lemma, 106

- Direct sum, 645
- Directed system, 59
- Distance, 75
- Domain, 21
- Dual functional, 247
- Dual space, 179
 - Second, 192
- Egoroff's Theorem, 375
- Eidelheit Separation Theorem, 221
- Empty set, 19
- Epigraph, 225
- Equicontinuity, 80
- Equivalence, 20
- Equivalent
 - Metric spaces, 75
 - Normed linear spaces, 163
- Essential supremum, 495
- Exhaustion, 123
- Expectation, 659
- Extensionality, 19
- Extreme point, 326
- Extreme subset, 326
- F_σ , 518
- Fatou's Lemma, 394, **467**
- Fenchel Duality Theorem, 236
- Field, 134
 - Sub-, 134
- Finite, 23
 - Banach space valued measure, 419
 - Banach space valued pre-measure, 415
 - Locally, 124
 - Banach space valued measure, 576
 - Topological measure space, 365
 - Measure, 351
 - Star, 126
- Finite intersection property, 97
- Fubini Theorem, 560, 562, 570
- Function, 21
 - Analytic, 313
 - Linear, 138
 - Measurable, 368
 - Of bounded variation, 575
- Functional, 138
 - Sublinear, 184
- Fundamental Theorem of Calculus, 623, 627
- G_δ , 518
- Generation process, 427
- Gram determinant, 648
- Gram matrix, 648
- Gram-Schmidt Procedure, 647
- Graph, 21
- Greatest element, 25
- Group, 131
 - Abelian, 132
 - Semi, 132
 - Abelian, 132
- Hahn Decomposition Theorem, 478
- Hahn-Banach Theorem
 - Extension form, 185
 - Complex, 187
 - Geometric form, 220
 - Simple version, 188
- Hausdorff
 - Maximal Principle, 26
 - Space, 47
- Heine-Borel, 99
- Hermitian adjoint, 645
- Hermitian operator, 645
- Hilbert space, 639
- Hölder's Inequality, 149, 497
- Homeomorphical isomeasure, 539
- Homeomorphism, 40
- Homomorphism, 132
 - Ring, 133
 - Vector space, 138
- Homotopic, 54
- Hyperplane, 217
 - Supporting, 221
- Ideal, 133
- Identity element, 133

- Image, 21
- Implicit Function Theorem, 295, 297
 - Analytic, 323
- Increment on semi-open rectangle, 575
- Indicator function, 379
- Induced measure, 540
- Infinity, 19
- Injective, 21
- Injective Mapping Theorem, 290
- Inner product, 637
- Integral
 - Finite Banach space valued measure space, 429
 - Finite measure space, 384
 - Infinite Banach space valued measure space, 429
 - Infinite measure space, 385
 - Nonnegative extended real-valued functions, 394
- Integration by Parts, 633
- Integration system, 383
- Interval, 359
- Inverse Function Theorem, 293
 - Analytic, 320
 - Global, 304, 305
- Inverse image, 21
- Invert, 300
- Invertible, 21
- Isomeasure, 539
- Isometry, 75
- Isomorphism, 132
 - Isometrical, 157
 - Ring, 133
 - Vector space, 138
- Iterated Limit Theorem, 94
- Jensen's Inequality, 413
- Joint Limit Theorem, 93
- Jordan Decomposition Theorem, 479, **480**
- Kernel
 - For measure, 424
- Krein-Milman, 327
- Kuhn-Tucker Theorem, 339
- $L_p(\mathcal{X}, \mathcal{Y})$, 494, 497
- Lagrange multiplier, 249, 334, 335
- Lattice, 168
- $\bar{L}_p(\mathcal{X}, \mathcal{Y})$, 494, 497
- Least element, 24
- Lebesgue Decomposition Theorem, 491
- Lebesgue Dominated Convergence Theorem, 402, **405**, 437, **440**, 467, 468
- Lebesgue measure, 360, 594
- Lebesgue number, 107
- Lebesgue outer measure, 361, 594
- Lebesgue point, 614
- Lebesgue-Stieltjes integral, 595
- Legendre polynomials, 652
- Limit, 59, 63
 - Inferior, 67, 68
 - Joint, 71
 - Superior, 67, 68
- Linear combination, 140
- Linear variety, 140
 - Generated by, 140
 - Closed, 156
- Linearly independent, 144
 - Vectors, 144
- Lipschitz continuity, 602
- m -dimensional Lebesgue measure space, 594
- Maximal, 24
- Mazur's Theorem, 220
- Meager, 48
 - Co-, 48
 - Non-, 48
- Mean Value Theorem, 51, 263, 264
- Measurable
 - Set, 349
 - Space, 349
 - With respect to an outer measure, 355

- Measure, 349
 - Banach space valued, 419
 - Induced outer, 356
 - Outer, 355
 - Product, 547
 - Space, 349
 - Vector, 463
- Measure on algebra, 356
 - Banach space valued, 533
- Metric, 73
 - Pseudo-, 89
- Metric measure space, 365
 - Banach space valued, 513
- Metric space, 73
 - Product, 80
- Metrizable, 92
- Minimal, 24
- Minkowski functional, 218
- Minkowski's Inequality, 150, 494
- Monotone class, 544
- Monotone Class Lemma, 544
- Monotone Convergence Theorem, 395
- Mutually singular measures, 473

- Natural mapping, 192
- Natural metric, 154
- Negative definite, 326
- Negative semi-definite, 326
- Net, 59
 - Joint, 71
 - Sub-, 62
- Newton's Method, 661
- Norm
 - Pseudo, 167
- Normal
 - Subgroup, 132
- Normal equation, 649
- Normal space, 47
- Normed linear measure space, 365
 - Banach space valued, 514
- Null space, 138

- Open mapping, 199
- Open Mapping Theorem, 200, 291

- Operator
 - Affine, 138
 - Linear, 138
- Order
 - Group, 132
- Ordering
 - Partial, 24
 - Total, 24
- Orthogonal, 194, 639
- Orthogonal complement, 194
 - Pre-, 194
- Orthonormal, 647
- Outer product of tensors, 312

- Pairing, 19
- Paracompact, 125
- Parallelogram Law, 638
- Partition of Unity, 118
- π -system, 544
- Point
 - Accumulation, 35
 - At infinity, 120
 - Beginning, 54
 - Boundary, 35
 - Cluster, 59
 - End, 54
 - Exterior, 35
 - Interior, 35
 - Relative, 156
 - Of closure, 35
- Point of minimum, 325
- Point of relative extremum, 325
- Point of relative minimum, 325
- Positive definite, 325
- Positive linear functional, 520
- Positive semi-definite, 326
- Power set, 19
- Pre-Hilbert space, 637
- Primal functional, 245
- Principal of a region, 573
- Probability measure space, 659
- Projection, 33, **43**
- Projection Theorem, 642, 655
- Proper, 121
 - Countably, 121

- Pythagorean Theorem, 639
- Quotient, 20
- Quotient space, 89, 165–167
- Raabe's Test, 665
- Radon-Nikodym derivative, 482
- Radon-Nikodym Theorem, 486, **488**
- Random variable, 659
- Range, 21
- Range space, 138
- Rectangle in $\mathbb{R}^m$, 572, 573
- Refinement, 100
- Reflexive, 20
- Region, 573
- Regular point, 334
- Regular space, 47
- Regularity, 19
- Relation, 20
- Replacement, 19
- Residual, 48
- Restriction, 21
- Riesz Representation Theorem, 507, 521, 531, 532, 634–636
- Riesz-Fréchet Theorem, 643
- Ring, 133
 - Sub-, 133
- Semialgebra, 366
- Separable, 48
- Separation, 50
- Set
 - Boundary, 35
 - Closed, 35
 - Closure, 35
 - Relative, 37
 - Complement, 35
 - Convex, 141
 - Exterior, 35
 - Interior, 35
 - Negative, 477
 - Open, 35
 - Relative, 156
 - Positive, 477
- σ -algebra, 23
 - Generated by, 23
- σ -finite
 - Banach space valued measure, 419
 - Measure, 351
- Simple function, 379
 - Banach space valued measure space, 431
- Stationary point, 335
- Stone-Čech compactification, 128
- Subordinate, 117
- Subset, 20
- Subspace, 137
 - Banach space valued measure space, 423
 - Banach space valued pre-measure space, 418
 - Generated by, 140
 - Measure space, 354
 - Proper, 137
 - Topological measure space, 366
- Support
 - Of a convex set, 222
 - Of a function, 117, **520**
- Support functional, 223
- Surjective, 21
- Surjective Mapping Theorem, 290
- Symmetric, 20
- Taylor Series, 313
- Taylor's Theorem, 285
- Tensor, 311
 - Transpose, 311
- Tietze's Extension Theorem, 56
- Tonelli Theorem, 559
- Topological measure space, 365
 - Banach space valued, 513
 - Product, 564
- Topological space, 35
- Topology, 35
 - Generated by, 38
 - Natural, 73
 - Of pointwise convergence, 57
 - Product, 42
 - Stronger, 38
 - Subset, 37

- Weak, 57, 206
- Weak*, 209
- Weaker, 38
- Total variation, 414, 419, 533, 538
 - Of function, 575
- Totally bounded, 79
- Transitive, 20
- Tychonoff
 - Space, 47
 - Theorem, 112
- Uniform
 - Boundedness Principle, 50
 - Cauchy sequence, 80
 - Continuity, 78
 - Convergence, 80
 - Equivalent
 - Metric spaces, 78
 - Homeomorphism, 78
- Union, 19
- Urysohn Metrization Theorem, 93
- Urysohn's Lemma, 55
- Vector space, 134
- Vitali, 613
- Vitali's Lemma, 613
- Well-ordering, 27
 - Principle, 27
- Zero-element, 132
- Zorn's Lemma, 27