FUNCTIONAL ANALYSIS, MEASURE THEORY AND REAL ANALYSIS

ZHIYAO XIONG

CONTENTS

1. Introduction . . . 1
2. Metric spaces . . . 2
   2.A. Basic concepts . . . 2
   2.B. Completeness, Baire category theorem . . . 3
   2.C. Compactness and boundedness, Arzelà–Ascoli theorem . . . 4
   2.D. Fixed point theorems . . . 6
3. Normed vector spaces and continuous operators . . . 9
   3.A. Basic concepts . . . 9
   3.B. Seminorm, balanced and absorbing convex set . . . 10
   3.C. Finite-dimensional NVS, Riesz lemma . . . 11
   3.D. Basic topology theorems — Banach spaces . . . 12
   3.E. Hahn–Banach theorem, division theorems . . . 14
   3.F. Deeper topology theorems — separability, Banach operator spaces, reflexivity . . . 17
   3.G. Weak(-*) topology, sequentially weak(-*) compactness, achieving distance . . . 19
   3.H. Complements, orthogonal relations . . . 22
   3.I. Compact operators, Fredholm property . . . 24
   3.J. Operator equations, spectrum on $\mathbb{C}$ . . . 28
4. Hilbert space . . . 31
   4.A. Basic knowledge . . . 31
   4.B. Orthogonal systems, orthogonal relations . . . 31
   4.C. Riesz representation theorem, Bilinear forms, Lax–Milgram theorem . . . 34
   4.D. Sovling elliptic PDE . . . 36
5. Spectrum theory . . . 39
   5.A. Fredholm alternative, solvability problems . . . 39
   5.B. Spectrum on $\mathbb{R}$, eigenvalues of epplitic PDE, compactness of spectrum . . . 40
   5.C. Spectrum of compact operator . . . 42
   5.D. Case of Hilbert spaces — bilinear forms, self-adjoint . . . 43
   5.E. Compact self-adjoint operators, eigenvector spaces, spectral decomposition . . . 46
   5.F. Epplitic PDE for $L = \Delta$ . . . 47

6. Measurable spaces and measurable maps . . . 48
6.A. Basic concepts, methods of judging measurability . . . 48
6.B. Pointwise limit of measurable maps . . . 49
6.C. Approximation — pointwise limit of simple functions . . . 50
6.D. Pullback — Borel σ-algebra, product measurable space . . . 51
6.E. Measurability of sections . . . 53
7. Measures . . . 54
7.A. Basic knowledge . . . 54
7.B. Approximations — Egoroff's theorem, convergence in measure . . . 55
7.C. Completion, completele spaces . . . 56
7.D. Outer measure, elementary sets . . . 57
7.E. Correspondence between outer measures and measures . . . 58
7.F. Premeasure on algebra, from premeasure to outer measure . . . 58
7.G. From premeasure to measure — extension comparison . . . 59
7.H. From elementary family to algebra, product measure . . . 60
7.I. From algebra to σ-algebra . . . 61
8. Lebesgue measure . . . 62
8.A. Construction of Lebesgue measure . . . 62
8.B. Elementary sets — boxes and cubes, Lebesgue (outer) measure . . . 62
8.C. Examples — generalized Cantor sets, Cantor-Lebesgue function . . . 64
8.D. Filling problems . . . 65
8.E. Regularity, Radon measures on $\mathbb{R}^n$ . . . 66
8.F. Regularity of Lebesgue measure, F_σ and G_δ sets . . . 68
8.G. Existence of non-measurable sets . . . 70
8.H. Uniqueness, product property . . . 71
9. Integration on a measure space . . . 73
9.A. Integration on L^+ . . . 73
9.B. (Improved) monotone convergence theorem, Fatou's lemma . . . 74
9.C. Integration of measurable functions, the dominated convergence theorem . . . 76
9.D. Product measure, Fubini-Tonelli theorem . . . 79
9.E. The transformation of integration, polar coordinates . . . 81
9.F. Riemann integration v.s. Lebesgue intergration . . . 83
10. Signed measures and differentiation . . . 86
10.A. Signed measure . . . 86
10.B. Decomposition theorems, standard integration representation . . . 87
10.C. Variation, integration, derivative, absolutely continuous . . . 88
10.D. Radon-Nikodym-Lebesgue theorem . . . 89

10.E. Lebesgue decomposition, Radon-Nikodym derivative, Chain rule 90
10.F. Complex measure, Radon-Nikodym-Lebesgue theorem, total variation 91
10.G. Pointwise Differentiation on $\mathbb{R}^n$, coincidence of two derivatives 92
10.H. Hardy-Littlewood maximal function, Lebesgue differentiation theorem 94
10.I. A New framework of derivatives 95
11. Radon measures 97
11.A. LCH space, Radon measure, positive linear functionals on $C_c(X, \mathbb{C})$ 97
11.B. Riesz Representation Theorem 97
11.C. Conclusions of regularity — σ-finite, F_σ and G_δ, σ-compact 98
11.D. LSC functions, integration approximation 99
11.E. $C_c(X, \mathbb{C})^* = C_0(X, \mathbb{C})^*$, Jordan decomposition of $C_0(X, \mathbb{R})^*$ 100
11.F. Signed (or complex) Radon measure, Riesz representation theorem 102
11.G. Filling problems and covering theorems for $\mathbb{R}^n$, Vitali, Besicovitch 103
11.H. Differentiating Radon measures with respect to Radon measures. 105
11.I. Weak convergence of Radon measures 107
12. L^p spaces 109
12.A. L^p space 109
12.B. Hölder inequality, Minkowski inequality 109
12.C. Banach, separability, dual, weak compact, uniformly convex 111
12.D. Convolution, regularization. 113
13. Appendix A — Semicontinuity 117
References 117

1. Introduction

TBC.

2. Metric spaces

Metric space is the basic models of spaces. We start with a brief review of it.

2.A. **Basic concepts.**

Definition 2.1 (Metric space). *Let X be a nonempty set. A map $d : X \times X \to \mathbb{R}$ is called a* ***metric*** *if it satisfies the following properties:*

(1) $d(x, y) \geq 0$ for all $x, y \in X$, and $d(x, y) = 0$ iff $x = y$;
(2) $d(x, y) = d(y, x)$ for all $x, y \in X$;
(3) $d(x, y) \leq d(x, z) + d(z, y)$, for all $x, y, z \in X$.

We shall call (X, d) a ***metric space****.*

The open balls form the base for a *topology* on the metric space (X, d), making it a topological space. Next, we will introduce some useful notions of (X, d), many of which are related to topology.

Definition 2.2 (Basic notations of metric spaces). *Let (X, d) be a metric space. There are some well-known notions:*

(1) X is called ***complete*** *if every Cauchy sequence of points in X has a limit that is also in X.*
(2) For two subsets A and B, we say that A is ***dense*** *in B if $B \subset \overline{A}$.*
(3) X is called ***separable*** *if there exists a countable dense subset.*
(4) A subset $A \subset X$ is called ***nowhere dense*** *if its closure has empty interior.*
(5) A subset $A \subset X$ is said to be a ***meagre subset of*** *X, or of* ***first category*** *in X if it is a countable union of nowhere dense subsets of X.*
(6) A subset $A \subset X$ is of ***second category*** *in X if it is not of first category in X.*
(7) A subset $A \subset X$ is called ***sequentially compact*** *if every sequence of points in A has a convergent subsequence convergeing to a point in X.*
(8) A subset $A \subset X$ is called ***compact*** *if every open cover of A has a finite subcover.*
(9) A subset $A \subset X$ is called ***bounded*** *if $A \subset B(x, r)$ for some $x \in X$ and for some $r > 0$.*
(10) A subset $A \subset X$ is called ***totally bounded*** *if for every $\varepsilon > 0$, there exists a finite collection of open balls in X of radius ε whose union contains A.*

A map $T : (X, d_1) \to (Y, d_2)$ is called an ***isometry*** *if $d_2(Tx, Ty) = d_1(x, y)$ for all $x, y \in X$, and is called an* ***isometric isomorphism*** *if it is a bijective isometry. Two metric spaces (X, d_1) and (Y, d_2) are called* ***isometric*** *if there is an isometric isomorphism from X to Y.*

Remark 2.3. There are some trivial conclusions.

(1) A closed subset A of X is nowhere dense iff its interior is empty.
(2) A subset is dense iff every nonempty open subset intersects it.
(3) A subset B is nowhere dense iff for each open subset U, $B \cap U$ is not dense in U.

Remark 2.4. Being separable makes it possible for us to do induction to a certain extent. See lemma 3.40.

Example 2.5. There are some basic examples.

(1) If (X, d) is a metric space, then (X, d_1) is also a metric space where $d_1(x, y) = \frac{d(x,y)}{1+d(x,y)}$.

(2) Let (s) be the space of all sequences of real numbers, and define

$$d(x,y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|x_j - y_j|}{1 + |x_j - y_j|}.$$

Then $((s), d)$ is separable and complete.
(3) $\mathbb{R}^n$ is separable and complete.
(4) $L^p(X, \mathcal{X}, \mu)$ $(1 \leq p \leq \infty)$ is a complete metric space.
(5) If Ω is Lebesgue measurable $(1 \leq p \leq \infty)$, then $L^\infty(\Omega, \mathcal{L}(\Omega), m)$ is separable.
(6) If Ω is Lebesgue measurable with $m(\Omega) > 0$, then $L^\infty(\Omega, \mathcal{L}(\Omega), m)$ is not separable.
(7) ℓ^p is separable $(1 \leq p \leq \infty)$, but ℓ^∞ is not separable.

Proposition 2.6. *The distance function is continuous with respect to each variable.*

Proof. Trivial. □

In the next we introduce some important properties of metric spaces.

2.B. **Completeness, Baire category theorem.** Although completeness is a property of the metric and not of the topology,[1] it will lead to some important conclusions related to topology. For instance, we have the Baire category theorem.

Theorem 2.7 (Completion)**.** *For any metric space (X, d), there is a complete metric space (Y, d_1) such that there exists a dense subset Y_1 satisfying that (Y_1, d_1) and (X, d) are isometric. Moreover, such (Y, d_1) is unique up to isometric isomorphism.*

Proof. Well-known. □

Theorem 2.8 (Cantor's intersection theorem)**.** *Let (X, d) be a complete metric space, and let (F_n) be a sequence of nonempty closed subsets satisfying:*
(1) $F_n \supset F_{n+1}$, $n = 1, 2, \cdots$.
(2) $d(F_n) = \sup_{x,y \in F_n} d(x,y) \to 0$, as $n \to \infty$.
Then there exists a unique $x \in \bigcap_{j=1}^{\infty} F_j$.

Proof. Since $F_n \neq \varnothing$ for each n, we choose a point $x_n \in F_n$ for each n. It follows that $(x_n)_{n=1}^{\infty}$ is a Cauchy sequence and hence converges to some $x \in X$, since (X, d) is complete. Note that for any fixed n, we have $x_m \in F_m \subset F_n$ for all $m \geq n$; it follows from the closedness of F_n that $x \in F_n$. Hence $x \in \bigcap_{n=1}^{\infty} F_n$. If there is another $y \in \bigcap_{n=1}^{\infty} F_n$, then $d(x,y) \leq d(F_n) \to 0$ and hence $d(x,y) = 0$; then $x = y$. □

Theorem 2.9 (Baire category theorem)**.** *There are several editions.*
(1) Let (X, d) be a complete metric space. Then for each countable collection of open dense subsets $(U_n)_{n=1}^{\infty}$, their intersection $\bigcap_{n=1}^{\infty} U_n$ is dense.
(2) Let (X, d) be a nonempty complete metric space. If X is the union of a countable family $(E_n)_{n=1}^{\infty}$ of closed subsets, then at least one of these closed subsets contains a nonempty open set.
(3) Let (X, d) be a complete metric space. Then X is of second category.

[1]It means that a complete metric space can be homeomorphic to a non-complete one. An example is $\mathbb{R}$, which is complete but homeomorphic to (0,1), which is not complete.

Proof. A subset is dense iff every nonempty open subset intersects it. Thus to show (1) it suffices to show that any nonempty open G in X has a point x in common with all of the U_n. Since U_1 is dense, G intersect U_1; thus there is a point x_1 and $0 < r_1 < 1$ such that $\overline{B(x_1, r_1)} \subset G \cap U_1$. Since each U_n is dense, we can continue recursively to find a pair of sequences x_n and $0 < r_n < n^{-1}$ such that $\overline{B(x_n, r_n)} \subset B(x_{n-1}, r_{n-1}) \cap U_n$. It follows that (x_n) is a Cauchy sequence and hence converges to some $x \in X$. By closedness, $x \in \overline{B(x_n, r_n)}$ for each n. Therefore, $x \in G$ and $x \in U_n$ for each n. We are done.

For (2), suppose for contradiction that each E_n is nowhere dense. Then E_n^c is dense for each n, and hence $\bigcap_{n=1}^{\infty} E_n^c = \left(\bigcup_{n=1}^{\infty} E_n\right)^c = \varnothing$ is dense by (1); a contradiction.

For (3), suppose for contradiction that X is of first category, and hence $X = \bigcup_{n=1}^{\infty} A_n$ in which A_n is nowhere dense. Fix $x_1 \in X$ and $r_1 < 1$. Since A_1 is nowhere dense, there is point x_2 and $0 < r_2 < 2^{-1}$ such that $\overline{B(x_2, r_2)} \subset B(x_1, r_1) \setminus A_1$. Since each A_n is nowhere dense, we can continue recursively to find a pair of sequences x_n and $0 < r_n < n^{-1}$ such that $\overline{B(x_n, r_n)} \subset B(x_{n-1}, r_{n-1}) \setminus A_{n-1}$. By theorem 2.8, there exists $x \in \bigcap_{n=1}^{\infty} \overline{B(x_n, r_n)}$. But $x \notin A_n$ for each n; a contradiction. We are done. □

Using Baire category theorem we can solve the following interesting problem.

Problem 2.10. *Suppose $f \in C^{\infty}(\mathbb{R})$ satisfying that $\forall\, x \in \mathbb{R}$, there exists $n_x \in \mathbb{N}$ such that $f^{(n_x)}(x) = 0$. Prove that f is a polynomial.*

Proof. See [Xio]. □

2.C. **Compactness and boundedness, Arzelà–Ascoli theorem.** Compactness is a vital topological properties. As in the case of $\mathbb{R}^n$, compactness is highly related to boundedness. Generally speaking, this relationship is related to completeness.

Theorem 2.11. *Let (X, d) be a metric space.*

(1) If a subset A is sequentially compact, then it is totally bounded.
(2) A subset A is compact iff A is sequentially compact and closed.
(3) If X is complete, *then a subset A is sequentially compact iff it is totally bounded.*

Proof. Well-known. □

Remark 2.12. By theorem 2.11 (2), when the background space is a metric space, sometimes we call A a **precompact** set if A is sequentially compact.

In the next we introduce a characterization of sequentially compact in $C([a, b], M)$.

Theorem 2.13 (Arzelà–Ascoli)**.** *Assume that M is a complete metric space. Donote all the continuous maps from $[a, b]$ to M by $C([a, b], M)$. Define*

$$d(x, y) = \sup_{t \in [a,b]} d_M(x(t), y(t)).$$

Then $(C([a, b], M), d)$ forms a complete metric space, and $A \subset C([a, b], M)$ is sequentially compact iff the following claims hold.

(a) A is bounded;
(b) For all $t \in [a, b]$ fixed. $A(t) = \{x(t) : t \in A\}$ is sequentially compact in M;

(c) *A is **uniformly equicontinuous**; i.e. for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_M(x(t), x(t')) < \varepsilon$ for all $x \in A$ and all $t, t' \in [a,b]$ with $|t' - t| < \delta$.*

Proof. Since d_M is continuous with respect to each variable, $f(t) := d_M(x(t), y(t))$ is continuous. Thus $f([a,b])$ is also compact, and hence $d(x,y) < +\infty$. In the next we prove that d is a metric.

(1) It's obvious that $d(x,y) = d(y,x) \geq 0$.
(2) If $d(x,y) = 0$, then $d_M(x(t), y(t)) = 0$, $\forall t \in [a,b]$, and hence $x = y$.
(3) Note that $\forall x, y, z \in C([a,b], M)$,

$$\begin{aligned} d(x,y) &= \sup_{t\in[a,b]} d_M(x(t), y(t)) \\ &\leq \sup_{t\in[a,b]} [d_M(x(t), z(t)) + d_M(z(t), y(t))] \\ &\leq \sup_{t\in[a,b]} d_M(x(t), z(t)) + \sup_{t\in[a,b]} d_M(z(t), y(t)) \\ &= d(x,z) + d(z,y). \end{aligned}$$

In the next we show that $(C([a,b], M), d)$ is complete. Given a Cauchy sequence $(x_n)_{n\geq 1}$ in $(C([a,b], M), d)$; then $(x_n(t))_{n\geq 1}$ is a Cauchy sequence in M, and hence has a limit point $x_0(t)$ since M is a complete metric space. It suffices to prove that $x_0 : [a,b] \to M$, $t \mapsto x_0(t)$, is continuous.

For all $\varepsilon > 0$, there exists $N > 0$ such that

$$d_M(x_m(t), x_n(t)) < \varepsilon \quad \forall t \in [a,b] \quad \forall m, n > N.$$

Fix some $n_0 > N$, since x_{n_0} is continuous on a compact set $[a,b]$, there exists $\delta > 0$ such that

$$d_M(x_{n_0}(t), x_{n_0}(t')) < \varepsilon \quad \forall |t - t'| < \delta,$$

and then for all $|t - t'| < \delta$,

$$\begin{aligned} d_M(x_0(t), x_0(t')) &\leq d_M(x_0(t), x_{n_0}(t)) + d_M(x_{n_0}(t), x_{n_0}(t')) + d_M(x_{n_0}(t'), x_0(t')) \\ &< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon. \end{aligned}$$

Thus x_0 is continuous. Now the first assertion follows. In the next we prove the second assertion.

Since $(C([a,b], M), d)$ is complete, we know that the following are equivalent.

(1) A is sequentially compact.
(2) A is totally bounded; i.e. for all $\varepsilon > 0$, A admits a finite ε-net.

"$\Longrightarrow$":

(a) Since A is totally bounded, then A is certainly bounded.
(b) Since any finite ε-net of A induces a finite ε-net of $A(t)$, $A(t)$ is also totally bounded. Note that M is also complete, and hence $A(t)$ is sequentially compact.
(c) For all $\varepsilon > 0$. A admits a finite ε-net $\{x_1, \cdots, x_n\}$. For each x_i there exists δ_i such that

$$d_M(x_i(t), x_i(t')) < \varepsilon \quad \forall |t - t'| < \delta_i$$

Put $\delta = \min_{1\le i\le n} \delta_i$. Then for all $x \in A$, there exists i such that $x \in B(x_i, \varepsilon)$, and then for $|t - t'| < \delta$

$$\begin{aligned} d_M(x(t), x(t')) &\le d_M(x(t), x_i(t)) + d_M(x_i(t), x_i(t')) + d_M(x_i(t'), x(t')) \\ &< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon \end{aligned}$$

Hence A is uniformly equicontinuous.

"$\Longleftarrow$": For all $\varepsilon > 0$, these exists $\delta > 0$ such that

$$d_M(x(t), x(t')) < \varepsilon \quad \forall x \in A \quad \forall |t - t'| < \delta$$

Find a δ-net of $[a, b]$, denoted by $\{t_1, \cdots, t_n\}$. Endow M^n with the natural metric $d = \sum_{i=1}^n d_M$, and then put

$$\Phi : C([a, b], M) \to M^n, \quad x \mapsto (x(t_1), \cdots, x(t_n))$$

Since $A(t_i)$ is sequentially compact for all i, $\prod_{i=1}^n A(t_i)$ is sequentially compact. Thus $\Phi(A) \subset \prod_{i=1}^n A(t_i)$ is sequentially compact, and hence has a finite ε-net, denoted by

$$\{\Phi(x_1), \cdots, \Phi(x_m)\}$$

Note that for all $x \in M$, we can find $i \in \{1, \cdots, m\}$, $j \in \{1, \cdots, n\}$ such that for all $t \in [a, b]$,

$$\begin{aligned} d_M(x(t), x_i(t)) &\le d_M(x(t), x(t_j)) + d_M(x(t_j), x_i(t_j)) + d_M(x_i(t_j), x_i(t)) \\ &< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon \end{aligned}$$

Hence A is totally bounded. □

2.D. **Fixed point theorems.** Finally, we introduce the Banach fixed point theorem and some of its applications.

Theorem 2.14 (Banach fixed point theorem). *Let (X, d) be a complete metric space. If a map $T : X \to X$ satisfies*

$$1 > \alpha := \sup_{x\neq y} \frac{d(Tx, Ty)}{d(x, y)}. \tag{2.1}$$

Then T admits a unique fixed point x^*. Moreover, for any $x_0 \in X$, setting $x_{n+1} = Tx_n = T^{n+1}x_0$, $n = 0, 1, \cdots$, then $x_n \to x^*$ and

$$d(x_n, x^*) \le \frac{\alpha^n}{1-\alpha} d(Tx_0, x_0)$$

Proof. First we show the uniqueness of the fixed point. If x^* and y^* are two distinct fixed points of T, then we have

$$0 < d(x^*, y^*) = d(Tx^*, Ty^*) \le \alpha d(x^*, y^*),$$

a contradiction. In the next we show the existence of the fixed point. Fix $x_0 \in X$, setting $x_{n+1} = Tx_n = T^{n+1}x_0$ for each n, it follows that

$$d(x_{n+1}, x_n) \le \alpha^n d(Tx_0, x_0),$$

and hence

$$\begin{aligned} d(x_{n+k}, x_n) &\leq \sum_{j=1}^{k} d(x_{n+j}, x_{n+j-1}) \leq \sum_{j=1}^{k} \alpha^{n+j-1} d(Tx_0, x_0) \\ &\leq \left(\sum_{j=0}^{\infty} \alpha^j \right) \alpha^n d(Tx_0, x_0) = \frac{\alpha^n}{1-\alpha} d(Tx_0, x_0). \end{aligned}$$

Thus (x_n) is Cauchy; then (x_n) converges to some $x^* \in X$ with

$$x^* = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} Tx_n = Tx^*,$$

since T is continuous via (2.1).

Finally, letting $k \to \infty$ in the above formula, the last assertion follows. □

Corollary 2.15. *Suppose that X is a complete metric space, and $T : X \to X$ satisfies*

$$\inf_n \sup_{x\neq y} \frac{d(T^n x, T^n y)}{d(x,y)} < 1.$$

Then T admits a unique fixed point.

Proof. There exists n such that

$$\lambda := \sup_{x\neq y} \frac{d(T^n x, T^n y)}{d(x,y)} < 1.$$

Thus T^n admits a unique fixed point x_0 by theorem 2.14. Since $T(x_0)$ is also a fixed point of T^n, we get that $T(x_0) = x_0$. If $x_1 \neq x_0$ is another fixed point of T, then

$$1 = \sup_{x\neq y} \frac{d(T^n x, T^n y)}{d(x,y)} \leq \lambda < 1,$$

a contradiction. □

Remark 2.16. Sometimes T is not contractive but T^n is contractive. In this case T also admits a unique fixed point.

Proposition 2.17. *Suppose that X is a compact metric space, and $T : X \to X$ satisfies*

$$d(Tx, Ty) < d(x,y) \quad \forall x, y \in X \text{ with } x \neq y.$$

Then T admits a unique fixed point.

Proof. Put $f = d(f(x), x)$. Note that for all $x \neq y$

$$\begin{aligned} |f(x) - f(y)| &= |\rho(T(x), x) - \rho(T(y), y)| \\ &\leq \rho(T(x), T(y)) + \rho(x, y) \\ &\leq 2\rho(x, y). \end{aligned}$$

Thus f is continuous. Since X is compact, we can find $x_0 \in X$ with

$$f(x_0) = \inf_{x\in X} f(x).$$

Suppose for contradiction that $T(x_0) \neq x_0$; then

$$f(T(x_0)) = d(T^2(x_0), T(x_0)) < d(T(x_0), x_0) = f(x_0),$$

a contradiction. Thus $T(x_0) = x_0$. If $x_1 \neq x_0$ is another fixed point of T, then

$$d(x_0, x_1) = d(T(x_0), T(x_1)) < d(x_0, x_1),$$

a contradiction. □

Corollary 2.18. *Suppose that $f : \mathbb{R}^n \times [t_0 - \delta, t_0 + \delta] \to \mathbb{R}^n$ is continuous, and is Lipschitz with respect to the first variable $x \in \mathbb{R}^n$, i.e. there exists $L > 0$ such that for all $t \in (t_0 - \delta, t_0 + \delta)$, $x, y \in \mathbb{R}^n$ we have*

$$\|f(x,t) - f(y,t)\| \leq L\|x - y\|.$$

Then the following ODE

$$\begin{cases} \dfrac{dx}{dt} = f(x(t), t) \\ x(t_0) = x_0 \end{cases}$$

has a unique continuous solution on $[t_0 - \beta, t_0 + \beta]$ where $0 < \beta < \min\{\delta, 1/L\}$.

Proof. The above ODE is equivalent to the continuous solution of the following integration equation:

$$x(t) = x_0 + \int_{t_0}^{t} f(x(s), s)\, ds. \tag{2.2}$$

Setting $X = C\left([t_0 - \beta, t_0 + \beta], \mathbb{R}^n\right)$ and

$$T : X \to X, \quad x \mapsto (Tx)(t) = x_0 + \int_{t_0}^{t} f(x(s), s)\, ds.$$

then the continuous solution of (2.2) is equivalent to the fixed point of T. We have showed in theorem 2.13 that X is complete; hence it suffices to show that T is contractive. Note that

$$\begin{aligned} d(Tx, Ty) &= \max_{t\in[t_0-\beta,t_0+\beta]} \|(Tx)(t) - (Ty)(t)\| \\ &= \max_{t\in[t_0-\beta,t_0+\beta]} \left\| \int_{t_0}^{t} f(x(s), s)\, ds - \int_{t_0}^{t} f(y(s), s)\, ds \right\| \\ &\leq \max_{t\in[t_0-\beta,t_0+\beta]} \left| \int_{t_0}^{t} \|f(x(s), s) - f(y(s), s)\|\, ds \right| \\ &\leq L\beta \max_{t\in[t_0-\beta,t_0+\beta]} \|x(t) - y(t)\| \\ &= L\beta d(x, y). \end{aligned}$$

Then the conclusion follows from theorem 2.14. □

For more properties, one can refer to any nice related textbook.

3. Normed vector spaces and continuous operators

In this section we introduce the basic theory of normed vector spaces and continuous (i.e. bounded) operators.

3.A. Basic concepts.

Definition 3.1 (Normed vector space). *A **normed vector space** is a vector space E equipped with an $\mathbb{R}$-valued function $x \mapsto \|x\|$ satisfying:*

(1) $\|x\| \geq 0$ for all $x \in E$, and $\|x\| = 0$ iff $x = 0$;
(2) $\|\alpha x\| = |\alpha|\,\|x\|$, for all $x \in E$ and $\alpha \in \Lambda$;
(3) $\|x + y\| \leq \|x\| + \|y\|$, for all $x, y \in E$.

*We abbreviate "normed vector space" by **NVS**.*

Setting $d(x, y) = \|x - y\|$, an NVS $(E, \|\cdot\|)$ becomes a metric space, so the knowledge of metric space is applicable.

Definition 3.2 (Basic notions of NVS). *There are some basic notations:*

*(1) Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are called **equivalent** if there exist $C_1, C_2 > 0$ such that for all $x \in X$, we have $C_1\|x\|_1 \leq \|x\|_2 \leq C_2\|x\|_1$.*
*(2) An NVS E is called a **Banach space** if it is complete.*
*(3) Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be two NVSs, then their **direct sum** $E \oplus F$ denotes the NVS whose norm is given by $\|(x, y)\| = \|x\|_E + \|y\|_F$.*

Remark 3.3. Throughout this section, and unless otherwise specified, the vector spaces are over $\mathbb{C}$ or $\mathbb{R}$. In the next we denote $\mathbb{C}$ or $\mathbb{R}$ by Λ.

Definition 3.4 (Unbounded and bounded operators). *Let E and F be two NVSs over Λ. An **unbounded linear operator** from E into F is a linear map $A : D(A) \subset E \to F$ defined on a linear subspace $D(A) \subset E$. The set $D(A)$ is called the **domain** of A.*

*One says that an unbounded linear operator A is **bounded** if $D(A) = E$ and if there exists a constant $M > 0$ such that*

$$\|Ax\| \leq M\|x\|, \quad \forall x \in E.$$

*In particular, a **linear functional** on E is an unbounded linear operator $f : E \to \Lambda$, and a **bounded linear functional** on E is a bounded linear operator $f : E \to \Lambda$.*

Definition 3.5 (Kernel, range, graph and closedness). *For a unbounded linear operator $A : D(A) \to F$, the **kernel** of A is denoted by $N(A) := \{x \in D(A) : Ax = 0\}$, the **range** of A is denoted by $R(A) := \{Ax : x \in D(A)\}$, and the **graph** of A is denoted by $G(A) := \{(x, Ax) : x \in D(A)\}$. Moreover, A is called **closed** if $G(A)$ is closed in $E \oplus F$.*

Definition 3.6 (Operator spaces). ***The norm of a bounded operator** is defined by*

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

***The space of all bounded linear operators** from E to F is denoted by $\mathcal{L}(E, F)$, which is also an NVS.*

*In particular, the space of all bounded linear functionals on E is denoted by E^*, and is called the **dual space** of E.*

Definition 3.7 (Adjoint operators)**.** *Let $A : D(A) \subset E \to F$ be an unbounded linear operator that is* densely defined*. We shall introduce an unbounded operator $A^* : D(A^*) \subset F^* \to E^*$ as follows. First, one defines its domain:*

$$D(A^*) = \{v \in F^* : \exists c \geq 0 \text{ such that } |\langle v, Au\rangle| \leq c\,\|u\|\,, \forall u \in D(A)\}.$$

It's clear that $D(A^*)$ is a linear subspace of F^*. We shall now define A^*v. Given $v \in D(A^*)$, consider the map $g : D(A) \to \mathbb{R}$ defined by

$$g(u) = \langle v, Au\rangle \quad \forall u \in D(A)$$

We have

$$|g(u)| \leq c\,\|u\| \quad \forall u \in D(A)$$

By Hahn–Banach 3.24 there exists a linear functional $f : E \to \mathbb{R}$ that extends g such that

$$|f(u)| \leq c\,\|u\| \quad \forall u \in E$$

It follows that $f \in E^*$. Note that the extension of g is unique, since $D(A)$ is dense in E. Now set

$$A^*v = f$$

*The unbounded linear operator $A^* : D(A^*) \subset F^* \to E^*$ os called the **adjoint** of A.*

Remark 3.8. In addition to studying the operator itself, the operator is also used to reflect the properties of NVS. For instance, some information of the topology of an NVS is embodied by the properties of operators on it, especially the linear functionals.

Remark 3.9. We can study A via A^*. This is very useful in solving the operator equations, since we add topology into our consideration by this way, and the topology is the key point.

Now let's introduce some important properties of normed vector spaces and operators.

3.B. **Seminorm, balanced and absorbing convex set.** Seminorm, a generalization of norm, is also a means of constructing a norm. Seminorm is highly related to the balanced and absorbing convex sets. The relevant theory is as follows.

Definition 3.10 (Seminorm)**.** *Let X be a vector space. We call $p : X \to \mathbb{R}$ a **seminorm** if it satisfies the following two conditions:*

(1) (Subadditivity) *$p(x + y) \leq p(x) + p(y)$, $\forall x, y \in X$;*
(2) (Absolute homogeneity) *$p(\alpha x) = |\alpha| p(x)$ for all $x \in X$ and $\alpha \in \Lambda$;*

Proposition 3.11. *Let X be a vector space, and let $p : X \to \mathbb{R}$ be a seminorm. Then p is a norm iff $\{x \in X : p(x) = 0\} = \{0\}$.*

Proof. Trivial. □

Definition 3.12 (Balanced and absorbing). *Let X be a vector space. A subset M is call* ***balanced*** *if $\alpha M \subset M$ for any scalar α with $|\alpha| \leq 1$, and is call* ***absorbing*** *if for every $x \in X$, there exists $\varepsilon_x > 0$ such that $\alpha x \in M$ for any scalar α with $|\alpha| \leq \varepsilon_x$.*

Theorem 3.13 (Correspondence between seminorms and balanced absorbing convex sets). *Let X be a vector space. We denote the collection of all seminorms on X by $\mathcal{P}$, and denote the collection of balanced and absorbing convex subsets by $\mathcal{M}$. Let X be a vector space. If $p \in \mathcal{P}$, then $M_p := \{x \in M : p(x) \leq 1\} \in \mathcal{M}$ with*

$$p(x) = \inf\left\{\alpha : \alpha > 0, \alpha^{-1}x \in M\right\}.$$

If $M \in \mathcal{M}$, then $p_M(x) := \inf\left\{\alpha : \alpha > 0, \alpha^{-1}x \in M\right\} \in \mathcal{P}$.

Proof. Well-known. The details are in my handwritten notes. □

Moreover, sometimes we need to construct a *subadditive* and *positive homogeneous* function, such as in the *division problem*. The following is a basic construction method.

Proposition 3.14. *Let X be a vector space. If a convex subset M is absorbing, we define its* ***Minkowski functional*** *by*

$$p_M(x) := \inf\left\{\alpha : \alpha > 0, \alpha^{-1}x \in M\right\}.$$

Then p_M is subadditive and positive homogeneous; i.e. P_M satisfies:
(1) (Subadditivity) *$p_M(x+y) \leq p_M(x) + p_M(y)$, for all $x, y \in X$;*
(2) (Positive homogeneity) *$p_M(\lambda x) = \lambda p_M(x)$, for all $x \in X$ and for all $\lambda > 0$.*

Proof. Well-known. The details are in my handwritten notes. □

3.C. **Finite-dimensional NVS, Riesz lemma.** In the next we briefly introduce the properties of finite-dimensional normed vector spaces, and introduce Riesz lemma to show the difference when we deal with the topology of infinite-dimensional NVS.

Theorem 3.15. *Let $(X, \|\cdot\|)$ be an n-dimensional NVS, and let $e_1, e_2, \cdots, e_n$ be a basis of X. Then there exist $0 < C_1 \leq C_2$ such that for all $x = \sum_{j=1}^n x_j e_j \in X$, we have*

$$C_1\left(\sum_{j=1}^n |x_j|^2\right)^{1/2} \leq \|x\| \leq C_2\left(\sum_{j=1}^n |x_j|^2\right)^{1/2}.$$

Proof. Trivial. □

Corollary 3.16. *Any n-dimension NVS E over $\mathbb{R}$ (or $\mathbb{C}$) is homeomorphic to $\mathbb{R}^n$ (or $\mathbb{C}^n$), and hence is complete and separable. Moreover, any bounded subset of E is sequentially compact.*

Theorem 3.17 (Riesz lemma). *Let E be an NVS and let $M \subset E$ be a closed linear subspace with $M \neq E$. Then*

$$\forall \varepsilon > 0\ \exists u \in E \text{ such that } \|u\| = 1 \text{ and } dist(u, M) \geq 1 - \varepsilon$$

Proof. This conclusion is intuitive. Let $v \in E$ with $v \notin M$. Since M is closed, then

$$d = \operatorname{dist}(v, M) > 0,$$

and hence we can choose $m_0 \in M$ such that

$$d \leq \|v - m_0\| \leq \frac{d}{1-\varepsilon}.$$

Then

$$u := \frac{v - m_0}{\|v - m_0\|}$$

satisfies the required properties. Indeed, for every $m \in M$, we have

$$\|u - m\| = \left\| \frac{v - m_0}{\|v - m_0\|} - m \right\| \geq \frac{d}{\|v - m_0\|} \geq 1 - \varepsilon,$$

since $m_0 + \|v - m_0\|\, m \in M$. □

Remark 3.18. Riesz lemma 3.17 says that a strictly ascending chain of closed subspaces has the property of divergence to a certain degree.

Since compactness will lead to convergence, the following corollary is natural.

Corollary 3.19. *An NVS E is finite-dimensional iff the unit ball of E is sequentially compact.*

Proof. If E is infinite-dimensional, then there is a sequence (E_n) of finite-dimensional subspaces of E such that

$$E_{n-1} \subset E_n \quad E_{n-1} \neq E_n.$$

By Riesz lemma 3.17 there is a sequence (u_n) with $u_n \in E_n$ such that

$$\|u_n\| = 1 \quad \mathrm{dist}(u_n, E_{n-1}) \geq \frac{1}{2}.$$

In particular,

$$\|u_n - u_m\| \geq \frac{1}{2} \quad \forall m \neq n.$$

Thus the sequence (u_n) has no convergent subsequence, and hence the unit ball B_E is not sequentially compact. Then the conclusion follows from corollary 3.16. □

Hence we see that the topology of infinite-dimensional NVS is very different and much more complicated. In the next we will introduce some basic properties of it.

3.D. **Basic topology theorems — Banach spaces.** As we mentioned before, we will introduce the properties of continuous operators to reflect the properties of the topology of space.

Proposition 3.20. *A linear operator L between normed vector spaces X and Y is bounded iff it is a continuous linear operator.*

Proof. $\Longrightarrow$: Suppose that L is bounded. Then, for all vectors $x, h \in X$ with h nonzero we have

$$\|L(x+h) - L(x)\| = \|L(h)\| \leq M\|h\|.$$

It follows that L is continuous at x. Moreover, since the constant M does not depend on x, this shows that in fact L is uniformly continuous, and even Lipschitz continuous.

$\Longleftarrow$: Conversely, it follows from the continuity at 0 that there exists a $\delta > 0$ such that

$$\|L(h)\| = \|L(h) - L(0)\| \leq 1, \quad \forall h \in B_{2\delta}(0).$$

Thus, for all non-zero $x \in X$, one has

$$\|Lx\| = \left\| \frac{\|x\|}{\delta} L\left(\delta \frac{x}{\|x\|}\right) \right\| = \frac{\|x\|}{\delta} \left\| L\left(\delta \frac{x}{\|x\|}\right) \right\| \leq \frac{\|x\|}{\delta} \cdot 1 = \frac{1}{\delta} \|x\|.$$

This proves that L is bounded. □

Proposition 3.21. *Let $(X, \|\cdot\|)$ be an NVS, and let $f : X \to \Lambda$ be a linear functional. Then f is bounded iff $N := \{x \in X : f(x) = 0\}$ is a closed subspace of X. Moreover, if $f \in X^*$ and $f \not\equiv 0$, then for any $x_0 \in X$ with $f(x_0) \neq 0$, we have*

$$X = N + span\{x_0\}.$$

Proof. Trivial. □

Theorem 3.22 (Topology theorem). *There are some basic theorems related to topology.*

(1) If X and Y are Banach, $A \in \mathcal{L}(X, Y)$, and A is bijective, then $A^{-1} \in \mathcal{L}(Y, X)$.
(2) If X and Y are Banach, and $A : X \to Y$ is linear, then $A \in \mathcal{L}(X, Y)$ iff $G(A)$ is closed in $X \oplus Y$.
(3) If X is Banach and Y is an NVS, and $(A_\alpha)_{\alpha \in A}$ is a collection of operators in $\mathcal{L}(X, Y)$ that satisfies

$$\sup_{\alpha \in A} \|A_\alpha(x)\| < \infty, \quad \forall x \in X.$$

Then $\sup_{\alpha \in A} \|A_\alpha\| < \infty$.
(4) If X and Y are Banach, $A \in \mathcal{L}(X, Y)$, and A is surjective, then A is an open map.

Proof. Well-known. The details are in my handwritten notes. □

Corollary 3.23. *Let X be a Banach space, and let A and B be two closed linear subspace with $A \cap B = \{0\}$. Then*

$$\begin{aligned} A + B \text{ is closed} &\iff \exists C > 0 : \|a\| \leq C\|a + b\|, \quad \forall a \in A, b \in B \\ &\iff \pi_A : A + B \to A \text{ is continuous.} \end{aligned}$$

where $\pi_A : A + B \to A$ is the standard projection.

Proof. First we note that A and B are certainly complete, and that the second equivalence is obvious. In the next we prove the first equivalence.

If there exists $C > 0$ such that

$$\|a\| \leq C\|a + b\|, \quad \forall a \in A, b \in B. \tag{3.1}$$

Then we have

$$\|b\| = \|a + b - a\| \leq \|a + b\| + \|a\| \leq (C + 1)\|a + b\|, \quad \forall a \in A, b \in B. \tag{3.2}$$

Hence for a Cauchy sequence $(a_n + b_n)_{n=1}^\infty$ in $A + B$, where $a_n \in A$ and $b_n \in B$ for each n, it easily follows from (3.1) and (3.2) that (a_n) and (b_n) are Cauchy sequences in A and B respectively. Thus $a_n \to a$ for some $a \in A$ and $b_n \to b$ for some $b \in B$. It follows that $a_n + b_n \to a + b \in A + B$, and hence $A + B$ is closed.

If $A+B$ is closed, then $(A+B, \|\cdot\|_X)$ is Banach. On the other hand, note that $(A+B, \|\cdot\|_1)$ is Banach, where $\|a+b\|_1 = \|a\|_X + \|b\|_X$ for all $x \in X$ and $y \in Y$, since $(A, \|\cdot\|_X)$ and $(B, \|\cdot\|_X)$ are Banach. Note that

$$\mathrm{id} : (A+B, \|\cdot\|_1) \to (A+B, \|\cdot\|_X)$$

is bijective and bounded, since

$$\|a+b\|_X \leq \|a\|_X + \|b\|_X.$$

It follows from 3.22 (1) that the inverse operator

$$\mathrm{id} : (A+B, \|\cdot\|_X) \to (A+B, \|\cdot\|_1)$$

is also bounded, then there exists $C > 0$ such that

$$\|a\|_X + \|b\|_X \leq C\|a+b\|_X, \quad \forall a \in A, b \in B.$$

Then the conclusion follows. We are done. □

3.E. **Hahn–Banach theorem, division theorems.** As we have showed:

(A) The topology of an infinite-dimensional NVS is much more complicated.

If we can't find new tools, it's difficult to get more powerful results.

Another difficulty with infinite dimension is that we lose the concept of coordinates. Hence some standard conclusions in linear algebra don't work (sometimes if we add some conditions of topology, the conclusions in linear algebra may work). An intuitive result is as follows:

(B) An operator $A : E \to F$ cannot be analyzed componentwise *when F is infinite-dimensional. In other words, it's difficult for us to analyze* from local to global.

In fact, we have the Hahn–Banach theorem, which solves the extension problem, and hence help us to analyze from local to global. Also, just because of this, Hahn–Banach theorem becomes the desired new tool that can help us analyze topology.

In the next we will introduce Hahn–Banach theorem and show some of its direct applications. One will see how Hahn–Banach theorem helps us deal with difficulties (A) and (B) in following subsections.

Theorem 3.24 (Hahn–Banach-Bohnenblust). *There are several editions of extension theorems.*

(1) Let X be a vector space over $\mathbb{R}$, and let $p : X \to \mathbb{R}$ be a subadditive and positive homogeneous function; i.e. p satisfies:
   (1.1) (Subadditivity) *$p(x+y) \leq p(x) + p(y)$, for all $x, y \in X$;*
   (1.2) (Positive homogeneity) *$p(\lambda x) = \lambda p(x)$, for all $x \in X$ and for all $\lambda > 0$.*
   Let $Y \subset X$ be a linear subspace and let $f : Y \to \mathbb{R}$ be a linear functional such that

$$f(y) \leq p(y), \quad \forall y \in Y.$$

   Then there exists a linear functional $F : X \to \mathbb{R}$ satisfying:
   (1.1) $F(y) = f(y)$, for all $y \in Y$;
   (1.2) $F(x) \leq p(x)$, for all $x \in X$.

(2) Let X be a vector space over $\mathbb{C}$, let p be a seminorm on X, let $Y \subset X$ be a linear subspace, and let $f : Y \to \mathbb{C}$ be a linear functional such that

$$|f(y)| \leq p(y), \quad \forall y \in Y.$$

Then there exists a linear functional $F : X \to \mathbb{C}$ satisfying:
(2.1) $F(y) = f(y)$, for all $y \in Y$;
(2.2) $|F(x)| \leq p(x)$, for all $x \in X$.

(3) Let $(X, \|\cdot\|)$ be a normed vector space over Λ, let $Y \subset X$ be a linear subspace, and let $f : Y \to \Lambda$ be a bounded linear functional. Then there exists a bounded linear functional $F : X \to \Lambda$ satisfying:
(3.1) $F(y) = f(y)$, for all $y \in Y$;
(3.2) $\|F\| \leq \|f\|$.

Proof. Use the transfinite induction for (1), and then (2) and (3) follows. The details are in my handwritten notes. □

Before we deal with difficulties (A) and (B), we introduce some direct applications of Hahn–Banach-Bohnenblust theorem 3.24 first.

Corollary 3.25. *Let $(X, \|\cdot\|)$ be an NVS.*

(1) For all $x_0 \in X \setminus \{0\}$, there exists $f \in X^*$ satisfying
(1.1) $f(x_0) = \|x_0\|$;
(1.2) $\|f\| = 1$.

(2) Let $Y \subset X$ be its closed linear subspace, and let $x_0 \in X \setminus Y$, then exists $f \in X^*$ satisfying
(2.1) $f(y) = 0$, for all $y \in Y$;
(2.2) $f(x_0) = dist(x_0, Y)$;
(2.3) $\|f\| = 1$.

To a certain degree, the following division theorems are intuitive representation of Hahn–Banach-Bohnenblust theorem 3.24.

Theorem 3.26 (Mazur theorem). *Let $(X, \|\cdot\|)$ be an NVS, and let $K \subset X$ be closed and convex. If $x_0 \notin K$, then there exists r and a bounded linear functional $F : X \to \mathbb{R}$ satisfying*

$$F(x_0) > r, \quad \text{and} \quad F(x) \leq r, \quad \forall x \in K.$$

Proof. WLOG we can assume that $0 \in K$. Since $x_0 \notin K$, $\delta = \operatorname{dist}(x_0, K) > 0$. Setting

$$M = \overline{\left\{x \in X : \operatorname{dist}(x, K) < \frac{\delta}{3}\right\}},$$

then it's easy to see the following properties:

(1) M is a closed and convex subset;
(2) $B_{\delta/3}(0) \subset M$ (and hence M is absorbing);
(3) $x_0 \notin M$.

It follows from proposition 3.14 that the corresponding Minkowski functional

$$p_M(x) := \inf\left\{\alpha : \alpha > 0, \alpha^{-1}x \in M\right\}$$

is subadditive and positive homogeneous, and satisfies

$$p(x_0) > 1, \quad \text{and} \quad p(x) \leq 1, \quad \forall x \in M.$$

Setting

$$Y = \{cx_0 : c \in \mathbb{R}\}, \quad \text{and} \quad f : Y \to \mathbb{R}, \quad cx_0 \mapsto cp(x_0),$$

then

$$f(y) \leq p(y), \quad \forall y \in Y,$$

and hence it follows from Hahn–Banach-Bohnenblust theorem 3.24 (1) that there exists a linear functional $F : X \to \mathbb{R}$ satisfying:

(1) $F(y) = f(y)$, for all $y \in Y$;

(2) $F(x) \leq p(x)$, for all $x \in X$.

In particular, we have

$$|F(x)| \leq \max\{p(x), p(-x)\} \leq 1, \quad \forall x \in B_{\delta/3}(0) \subset M.$$

Thus F is bounded, and hence is exactly the desired bounded linear functional. □

Corollary 3.27. *Given a normed vector space $(X, \|\cdot\|)$ over $\mathbb{R}$, let $K_1, K_2 \subset X$ be two closed and convex subsets with*

$$d_0 = dist(K_1, K_2) > 0.$$

Then there exists $f \in X^*$ such that

$$\sup_{x \in K_1} f(x) < \inf_{y \in K_2} f(y).$$

Proof. Each point in $K_1 - K_2$ is of the form $a - b$, where $a \in K_1$ and $b \in K_2$. Letting $a - b$ and $a' - b'$ be two points in $K_1 - K_2$, note that for any $\theta \in [0, 1]$,

$$\theta\,(a - b) + (1 - \theta)\,(a' - b') = [\theta a + (1 - \theta)\,a'] - [\theta b + (1 - \theta)\,b'].$$

Thus $K_1 - K_2$ is convex. Since $d_0 = \text{dist}(K_1, K_2) > 0$, we have

$$B\left(0, \frac{d_0}{2}\right) \bigcap (K_1 - K_2) = \varnothing.$$

Setting $A = B(0, d_0/2) - (K_1 - K_2)$, then A is also convex for the same reason. Since

$$A = \bigcup_{y \in (K_1 - K_2)} \left(B\left(0, \frac{d_0}{2}\right) - y\right),$$

we know that A is open. Note that $0 \notin A$, by Mazur theorem 3.26 there is some $f \in X^*$ such that

$$f(z) < 0 \quad \forall z \in A,$$

that is,

$$f(x) < f(y) \quad \forall x \in B\,(0, d_0/2) \quad \forall y \in (K_1 - K_2).$$

Letting $\varepsilon = \frac{d_0}{4}\,\|f\|$, it induces that

$$\varepsilon < f(y) \quad \forall y \in (K_1 - K_2),$$

which means

$$f(a) - f(b) > \varepsilon \quad \forall a \in K_1 \quad \forall b \in K_2.$$

Thus $-f$ satisfies the requirement. □

3.F. **Deeper topology theorems — separability, Banach operator spaces, reflexivity.** As mentioned in subsection 3.E, Hahn–Banach-Bohnenblust theorem 3.24 is just the desired new tool that helps us to analyze topology. In the next we introduce some deeper topology theorems which are derived via it.

Proposition 3.28. *Let E be an NVS. If E^* is separable, then E is separable.*

Proof. Since E^* is separable, there exists a sequence $(f_n)_{n=1}^{\infty}$ in E^* such that $\|f_n\| = 1$ for each n and $\{f_n\}_{n=1}^{\infty}$ is dense in the unit sphere of E^*. Choose $x_n \in X$ for each n such that $\|x_n\| = 1$ and

$$|f_n(x_n)| > \frac{1}{2}, \quad n = 1, 2, \cdots.$$

Setting $Y = \overline{\operatorname{span}\{x_n\}_{n=1}^{\infty}}$, then Y is separable. It suffices to prove that $X = Y$. Suppose for contradiction that there exists $x_0 \in X \setminus Y$, then by corollary 3.25, there exists $f \in X^*$ satisfying

$$\|f\| = 1, \quad f(x_0) = d = \operatorname{dist}(x_0, Y), \quad \text{and} \quad f|_Y \equiv 0.$$

Then note that

$$\|f - f_n\| = \sup_{\substack{x \in X \\ \|x\| \le 1}} |f(x) - f_n(x)| \ge |f(x_n) - f_n(x_n)| > \frac{1}{2}, \quad n = 1, 2, \cdots.$$

A contradiction. □

Proposition 3.29. *Suppose that X and Y are two normed vector spaces over Λ. Then $\mathcal{L}(X, Y)$ is Banach iff Y is Banach.*

Proof. Suppose that Y is Banach. Given a Cauchy sequence (A_n) in $\mathcal{L}(X, Y)$, i.e. for all $\varepsilon > 0$, there exists $N > 0$ such that

$$\|A_n - A_m\| \le \varepsilon \quad \forall m, n \ge N.$$

Then for each $x \in X$,

$$\|A_n x - A_m x\| \le \|A_n - A_m\| \, \|x\| \le \varepsilon \|x\| \quad \forall m, n \ge N.$$

Thus $(A_n x)$ is Cauchy and hence converges. Put

$$A : X \to Y, \quad x \mapsto \lim_{n \to \infty} A_n x.$$

It follows that A is linear, and

$$\begin{aligned} \|Ax\| = \left\| \lim_{k \to \infty} A_k x \right\| &= \left\| A_N x + \lim_{k \to \infty} (A_k - A_N) x \right\| \\ &\le \|A_N x\| + \left\| \lim_{k \to \infty} (A_k - A_N) x \right\| \\ &\le \|A_N\| \, \|x\| + \varepsilon \|x\| = (\|A_N\| + \varepsilon) \|x\|. \end{aligned}$$

Hence $A \in \mathcal{L}(X, Y)$. Note that for each $x \in X$, we have

$$\|A_n x - A_m x\| \le \varepsilon \|x\| \quad \forall m, n \ge N \implies \|A_n x - Ax\| \le \varepsilon \|x\| \quad \forall n \ge N.$$

Thus

$$\|A_n - A\| = \sup_{\|x\| \le 1} \|A_n x - Ax\| \le \varepsilon,$$

and hence $\lim_{n\to\infty} A_n = A$.

Suppose that $\mathcal{L}(X, Y)$ is Banach. Given any Cauchy sequence (y_n) in Y, i.e. for all $\varepsilon > 0$, there exists $N > 0$ such that

$$\|y_n - y_m\| \le \varepsilon \quad \forall m, n \ge N.$$

Choose $0 \ne x_0 \in E$ and via Hahn-Banach theorem 3.24 choose $f_0 \in E^*$ such that $f_0(x_0) = \|x_0\|$ and $\|f_0\| = 1$. Put

$$A_n : X \to Y, \quad x \mapsto f_0(x) y_n.$$

It's clear that $A_n \in \mathcal{L}(X, Y)$. Note that $\forall m, n \ge N$, we have

$$\|A_n - A_m\| = \sup_{\|x\| \le 1} \|A_n x - A_m x\| = \sup_{\|x\| \le 1} |f_0(x)| \|y_n - y_m\| \le \|y_n - y_m\| \le \varepsilon.$$

Hence we know that (A_n) converges to some $A \in \mathcal{L}(X, Y)$ since $\mathcal{L}(X, Y)$ is Banach. Note that for each $x \in X$ we have

$$\|A_n x - Ax\| \le \|A_n - A\| \|x\|.$$

Hence $(A_n x)$ converges. Letting $x = x_0$, we get that $y_n \to \|x_0\|^{-1} A x_0$. □

An important difference of the topology of infinite-dimensional NVS is that E^{**} may not be canonically isomorphic to E. Now we introduce the concept of *reflexivity.*

Definition 3.30 (Reflexive)**.** *Let E be an NVS and let $J : E \to E^{**}$ be the* ***canonical injection*** *from E into E^{**}.*[2] *The space E is called* ***reflexive*** *if J is surjective, i.e., $J(E) = E^{**}$.*

Remark 3.31. Since J is an isometry and X^{**} is Banach, a necessary condition of being reflexive is that E is Banach.

Theorem 3.32. *Let X be an Banach space. Then X is reflexive iff every closed linear subspace Y of X is reflexive.*

Proof. It suffices to show that if X is reflexive then every closed linear subspace Y is reflexive. For any $F \in Y^{**}$, define

$$\widetilde{F}(f) = F\left(f|_Y\right), \quad \forall f \in X^*.$$

Then

$$\left|\widetilde{F}(f)\right| \le \|F\| \cdot \|f|_Y\| \le \|F\| \cdot \|f\|,$$

and hence $\widetilde{F} \in X^*$. Since X is reflexive, there exists $x_0 \in X$ with

$$J(x_0) = x_0^{**} = \widetilde{F}.$$

[2]For $v \in E$, we have $J(v) : E^* \to \Lambda$, $f \mapsto \langle f, v \rangle_{E^*, E}$.

Hence

$$x_0^{**}(f) = f(x_0) = \widetilde{F}(f) = F(f|_Y), \quad \forall f \in X^*. \tag{3.3}$$

It suffices to show that $x_0 \in Y$. Suppose for contraction that $x_0 \notin Y$, by corollary 3.25, there exists $f \in X^*$ satisfying

$$\|f\| = 1, \quad f(x_0) = d = \operatorname{dist}(x_0, Y), \quad \text{and} \quad f|_Y \equiv 0.$$

Then it follows from (3.3) that

$$d = f(x_0) = F(f|_Y) = 0,$$

a contraction. □

We will go deeper for studying the topology of NVS in the next subsection.

3.G. **Weak(-*) topology, sequentially weak(-*) compactness, achieving distance.** Now we introduce two aim-oriented concepts, the weak topology and the weak-* topology.

Definition 3.33 (Weak topology and weak-* topology)**.** *Let E be an NVS.*

*(1) The **weak topology** $\sigma(E, E^*)$ on E is the* coarsest *topology that makes all the maps $(f)_{f\in E^*}$ continuous.*
*(2) The **weak-* topology** $\sigma(E^*, E)$ on E^* is the* coarsest *topology that makes all the maps $(\langle \cdot, x\rangle)_{x\in E}$ continuous.*

Remark 3.34. The significance of definition 3.33 is just embodied in "coarsest". This is the weakest requirement to achieve the goal. Actually in lots of problems, it suffices to show the convergence in weak topology, which is the key point.

Remark 3.35. Let X be an infinite-dimensional NVS. Then X equipped with the weak topology is *not* metrizable,[3] and X equipped with the weak-* topology is *not* always metrizable.

By the above remark, we need to re-clarify some notions.

Definition 3.36 (Weak(-*) convergence)**.** *Let E be an NVS.*

*(1) We say that a sequence (x_n) in E **weakly converges** to x if (x_n) converges to x in the weak topology $\sigma(E, E^*)$, which is denoted by $x_n \rightharpoonup x$.*
*(2) We say that a sequence (x_n) in E **weakly-* converges** to x if (x_n) converges to x in the weak-* topology $\sigma(E^*, E)$, which is denoted by $x_n \overset{*}{\rightharpoonup} x$.*

Definition 3.37 (Sequentially weak(-*) compactness)**.** *Let E be an NVS.*

*(1) A subset $A \subset E$ is called **sequentially weak compact** if every sequence of points in A has a subsequence that weakly converges to a point in E.*
*(2) A subset $A \subset E$ is called **sequentially weak-* compact** if every sequence of points in A has a subsequence that weakly-* converges to a point in E.*

In the next we introduce some basic properties of weak(-*) convergence.

[3]One can refer to https://math.stackexchange.com/questions/1381759.

Proposition 3.38 (Weak(-*) convergence)**.** *Let (x_n) be a sequence in an NVS E. Then*

(1) $x_n \rightharpoonup x \Longleftrightarrow \langle f, x_n\rangle \to \langle f, x\rangle$, $\forall f \in E^*$.
(2) If $x_n \to x$, then $x_n \rightharpoonup x$.
(3) If $x_n \rightharpoonup x$, then $(\|x_n\|)$ is bounded and $\|x\| \le \liminf_{n\to\infty} \|x_n\|$.
(4) If $x_n \rightharpoonup x$ and if $f_n \to f$ in E^*, then $\langle f_n, x_n\rangle \to \langle f, x\rangle$.

Also, let (f_n) be a sequence in E^*. Then

(1) $f_n \overset{*}{\rightharpoonup} f \Longleftrightarrow \langle f_n, x\rangle \to \langle f, x\rangle$, $\forall x \in E$.
(2) If $f_n \to f$, then $f_n \rightharpoonup f$ in $\sigma(E^*, E^{**})$. If $f_n \rightharpoonup f$ in $\sigma(E^*, E^{**})$, then $f_n \overset{*}{\rightharpoonup} f$ in $\sigma(E^*, E)$.
(3) If $f_n \overset{*}{\rightharpoonup} f$, then $(\|f_n\|)$ is bounded and $\|f\| \le \liminf_{n\to\infty} \|f_n\|$.
(4) If $f_n \overset{*}{\rightharpoonup} f$ and if $x_n \to x$ in E, then $\langle f_n, x_n\rangle \to \langle f, x\rangle$.

Proof. One can refer to [Bre], in which we use the uniform boundedness principle, i.e. theorem 3.22 (3). □

Theorem 3.39. *Let X be an NVS, let $x_0 \in X$, and let (x_n) be a sequence of points in X. If $x_n \rightharpoonup x_0$, then $x_0 \in \overline{Co\{x_n\}}$, where $Co\{x_n\}$ is the* ***convex hull****.*

Proof. Set $K = \overline{\mathrm{Co}\{x_n\}}$. If $x_0 \notin K$, via Mazur theorem 3.26, there exists a bounded linear functional $f : X \to \mathbb{R}$ and some constant r such that

$$f(x_0) > r, \quad \text{and} \quad f(x) \le r, \quad \forall x \in K.$$

But it follows from the weak convergence that $f(x_n) \to f(x_0)$; a contradiction. □

The properties of weak(-*) convergence lead to the following properties of sequentially weak(-*) compactness if we add some conditions of topology.

Lemma 3.40. *Let X be a separable NVS. Then any bounded set in X^* is sequentially weakly-* compact.*

Proof. Suppose that $(f_n) \subset X$ satisfying $\|f_n\| \le M$, $n = 1, 2, \cdots$. It suffices to show that there is a weakly-* convergent subsequence.

Since X is separable, choose a sequence (x_n) of points such that $\{x_n\}$ is a dense subset. Since $(f_n(x_1))_{n=1}^\infty$ is bounded, there is a convergent subsequence $(f_{n_j,1}(x_1))_{j=1}^\infty$. Proceeding inductively, for each k we find a subsequence $((n_j, k))_{j=1}^\infty$ of $((n_j, k-1))_{j=1}^\infty$ such that $(f_{n_j,k}(x_k))_{j=1}^\infty$ converges. Hence $(f_{n_j,j})_{j=1}^\infty$ converges at each x_k.

Then setting $f(x) = \lim_{j\to\infty} f_{n_j,j}(x)$, it easily follows that $f \in X^*$ and $f_{n_j,j} \overset{*}{\rightharpoonup} f$. □

Theorem 3.41. *Any bounded subset A of a reflexive space X is sequentially weakly compact.*

Proof. It suffices to show that given a bounded sequences $(x_n)_{n=1}^\infty$ in X, then there exist a subsequence $(x_{n_j})_{j=1}^\infty$ and a point $x \in X$ such that $x_{n_j} \rightharpoonup x$.

Set $Y = \overline{\mathrm{span}\{x_n\}_{n=1}^\infty}$, then Y is separable, and is reflexive via theorem 3.32. Then via proposition 3.28 we know that Y^* is separable. Hence, via lemma 3.40, $(x_n^{**}|_Y)$ has a subsequence $(x_{n_j}^{**}|_Y)$ that weakly-* converges to $x_0^{**}|_Y$ for some $x_0 \in Y$. It follows that

$$f(x_{n_j}) \to f(x_0), \quad \forall f \in Y^*,$$

and hence

$$f(x_{n_j}) = f|_Y(x_{n_j}) \to f|_Y(x_0) = f(x_0), \quad \forall f \in X^*.$$

We are done. □

After the above detailed analysis of topological properties, we can now prove the well-known property of achieving distance.

Corollary 3.42. *Let X be a reflexive space, and let K be a closed and convex subset. Then for every $x_0 \in K$, there exists $y_0 \in K$ such that*

$$\|x_0 - y_0\| = \inf_{y \in K} \|x_0 - y\|.$$

Proof. WLOG we assume that $x_0 \notin K$. Choose a sequence $(y_n)_{n=1}^{\infty}$ of points in K such that

$$\|x_0 - y_n\| \le d + \frac{1}{n}, \quad \text{where} \quad d = \inf_{y \in K} \|x_0 - y\|.$$

It's clear that (y_n) is bounded, and hence, by theorem 3.41, has a weakly convergent subsequence $(y_{n_j})_{j=1}^{\infty}$ that converges to some $z \in X$; i.e. we have

$$f(x - y_{n_j}) \to f(x - z), \quad \forall f \in X^*.$$

By theorem 3.39, we know $z \in K$. By corollary 3.25 there exists $f \in X^*$ such that

$$f(x - z) = \|x - z\| \quad \text{and} \quad \|f\| = 1$$

Hence

$$\|x - z\| = \lim_{j \to \infty} f(x - y_{n_j}) \le \|f\| \cdot \|x - y_{n_j}\| \le d + \frac{1}{n_j}, \quad \forall j.$$

It follows that z is the desired point. □

The following are some other well-known properties of achieving distance.

Problem 3.43. *The following facts are well-known:*

(1) If a Banach space E is uniformly convex, then E is reflexive;
(2) The minimizing point in corollary 3.42 is unique when E is uniformly convex;
(3) Every Hilbert space is uniformly convex;
(4) $L^p(X, \mathcal{X}, \mu)$ is uniformly convex if $1 < p < \infty$ and X is σ-finite.

Proof. See [Xio]. □

Problem 3.44. *Let (x_n) be a sequence in ℓ^1, and $x_0 \in \ell^1$. Prove that $x_n \rightharpoonup x_0 \iff x_n \to x_0$.*

Proof. See [Xio]. □

For more basic properties, one can refer to [Bre].

3.H. **Complements, orthogonal relations.** Next, we will show how Hahn-Banach-Bohnenblust theorem 3.24 helps us deal with difficulties (B) in subsection 3.E.

Hilbert spaces are typical cases that help us understand the framework. Roughly speaking, Hahn-Banach-Bohnenblust theorem 3.24 helps us find the orthogonal decomposition and helps us find a "basis" (i.e. a complete orthogonal system). See section 4 for the detailed introduction.

For general spaces, we still have the conclusions for *complements* and *orthogonal relations*. One can regard them as the generalization of the theory for Hilbert spaces.

Definition 3.45. *Let $G \subset E$ be a closed subspace of a Banach space. A subspace $L \subset E$ is said to be a* ***topological complement*** *or simply a* ***complement*** *of G if L satisfies:*

(1) L is closed;
(2) $G \cap L = \{0\}$ and $G + L = E$.

Lemma 3.46. *Let $(X, \|\cdot\|)$ be an NVS over Λ.*

(1) If $f_1, \cdots, f_n \in X^*$ are linear independent, then there exist $e_1, \cdots, e_n \in X$ satisfying:
  (1.1) $f_j(e_k) = \delta_{jk}$, $j, k = 1, 2, \cdots$;
  (1.2) We have

$$X = \left(\bigcap_{j=1}^{n} N(f_j)\right) \oplus \mathit{span}\left\{e_j\right\}_{j=1}^{n},$$

  where for every $x \in X$ we have

$$\left(x - \sum_{j=1}^{n} f_j(x)e_j\right) \in \bigcap_{j=1}^{n} N(f_j).$$

(2) If $Y \subset X$ is an n-dimensional linear subspace, then there exists a closed linear subspace Z of X such that $X = Y \oplus Z$.

Proof. For (1), we define

$$\varphi : X \to \Lambda^n, \quad x \mapsto (f_1(x), f_2(x), \cdots, f_n(x))$$

Then it follows easily from the linear independence that φ is surjective. Now choose e_j for each j such that

$$\varphi(e_j) = (\underbrace{0, \cdots, 0}_{i-1}, 1, 0 \cdots, 0)$$

then (1.1) follows. Moreover, note that

$$f_k\left(x - \sum_{j=1}^{n} f_j(x)e_j\right) = f_k(x) - \sum_{j=1}^{n} f_j(x)f_k(e_j) = f_k(x) - f_k(x) = 0, \quad \forall k.$$

It's clear that $\left(\bigcap_{j=1}^{n} N(f_j)\right) \cap \operatorname{span}\left\{e_j\right\}_{j=1}^{n} = \{0\}$, and then (1) follows.

For (2), setting $Y = \operatorname{span}\left\{e_j\right\}_{j=1}^{n}$, and then there exist $f_1, \cdots, f_n \in Y^*$ with

$$f_j(e_k) = \delta_{jk}, \quad j, k = 1, 2, \cdots.$$

By Hahn–Banach theorem 3.24, we extend f_j to a bounded linear functional $F_j : X \to \Lambda$ for each j, and then

$$X = \left(\bigcap_{j=1}^{n} N(F_j) \right) \oplus Y$$

by (1). Then (2) follows from proposition 3.21. □

Proposition 3.47. *Let $(X, \|\cdot\|)$ be an Banach space over Λ.*

(1) Every finite-dimensional subspace Y admits a complement.

(2) Every closed finite-codimensional subspace M admits a complement.[4]

(3) Let $N \subset E^*$ be a subspace of dimension p. Then

$$G = \{x \in E : \langle f, x \rangle = 0 \quad \forall f \in N\} = N^\perp$$

is closed and admits a complement of dimension p.

Proof. (1) and (3) follow immediately from lemma 3.46.

For (2), assume that $X = M + N$ for some finite-dimensional space $N \subset X$. We may always assume that $M \cap N = \{0\}$ (otherwise by (1) we can choose a complement N' of $M \cap N$ in N). Then the conclusion follows, since finite-dimensional subspace is always closed. □

Definition 3.48 (Orthogonal)**.** *Let E be an NVS. If $M \subset E$ is a linear subspace we set*

$$M^\perp = \{f \in E^* : \langle f, x \rangle = 0, \forall x \in M\}$$

If $N \subset E^*$ is a linear subspace we set

$$N^\perp = \{x \in E : \langle f, x \rangle = 0, \forall f \in N\}$$

*Note that, by definition, $N^\perp$ is a subset of E rather than E^{**}. It is clear that $M^\perp$ (resp. $N^\perp$) is a closed linear subspace of E^* (resp. E). We say that $M^\perp$ (resp. $N^\perp$) is the space **orthogonal** to M (resp. N).*

Proposition 3.49. *Let E be an NVS, and let $M \subset E$ be a linear subspace. Then*

$$(M^\perp)^\perp = \overline{M}.$$

Let $N \subset E^*$ be a linear subspace. Then

$$(N^\perp)^\perp \supset \overline{N}.$$

Proof. It is clear that $N \subset (N^\perp)^\perp$ and since $(N^\perp)^\perp$ is closed we have $\overline{N} \subset (N^\perp)^\perp$. It is also clear that $M \subset (M^\perp)^\perp$, and since $(M^\perp)^\perp$ is closed we have $\overline{M} \subset (M^\perp)^\perp$. Conversely, suppose for contradiction that

$$x_0 \in (M^\perp)^\perp \quad x_0 \notin \overline{M}.$$

Via Mazur theorem 3.26 we get

$$\langle f, x \rangle < \alpha < \langle f, x_0 \rangle \quad \forall x \in M,$$

[4] Let M be a subspace of a Banach space X. M has finite codimension if there exists a finite-dimensional space $N \subset X$ such that $M + N = X$.

for some bounded linear functional $f : E \to \mathbb{R}$ and some $\alpha \in \mathbb{R}$. Since M is linear space we get

$$\langle f, x \rangle = 0 \quad \forall x \in M,$$

and hence $\langle f, x_0 \rangle > 0$. Therefore $f \in M^\perp$ and consequently $\langle f, x_0 \rangle = 0$, a contradiction. $\square$

The following propositions are vital, in which topology also plays an important role.

Theorem 3.50 (Orthogonal relations). *Let E and F be two NVSs, and let $A : D(A) \subset E \to F$ be an unbounded linear operator that is densely defined and closed. Then,*

(1) $N(A) = R(A^*)^\perp$;
(2) $N(A^*) = R(A)^\perp$;
(3) $N(A)^\perp \supset \overline{R(A^*)}$;
(4) $N(A^*)^\perp = \overline{R(A)}$.

Moreover, if E^* is reflexive then $N(A)^\perp = \overline{R(A^*)}$.

Proof. One can refer to [Bre]. $\square$

Theorem 3.51 (Equivalence between closedness and orthogonal relations). *Let E and F be Banach, and let $A : D(A) \subset E \to F$ be an unbounded linear operator that is densely defined and closed. The following properties are equivalent:*

(1) $R(A)$ is closed.
(2) $R(A^*)$ is closed.
(3) $R(A) = N(A^*)^\perp$.
(4) $R(A^*) = N(A)^\perp$.

Proof. One can refer to [Bre]. $\square$

Remark 3.52. Coming back to Hilbert spaces again, the theory of complements and orthogonal relations will be more powerful when we deal with Hilbert spaces. One can see that theorem 4.5, proposition 4.11, proposition 4.12 and theorem 4.14 give us a powerful method to prove vital theorems like Riesz representation theorem 4.16.

3.I. **Compact operators, Fredholm property.** In the next we introduce the compact operators, which are typical examples of operators that make a good use of the preceding conclusions.

By adding the property of compactness, compact operators enjoy the Fredholm property, which is the core topic of this subsection.

Definition 3.53 (Compact operator). *Let E and F be two Banach spaces. A bounded operator $T \in \mathcal{L}(E, F)$ is said to be* ***compact*** *if $T(B_E)$ is sequentially compact. The set of all compact operators from E into F is denoted by $\mathcal{K}(E, F)$. For simplicity one writes $\mathcal{K}(E) = \mathcal{K}(E, E)$.*

Example 3.54. Any finite-rank operator is compact.

Proposition 3.55. *Let E, F and G be three Banach spaces.*

(1) If $T \in \mathcal{L}(E,F)$ and $S \in \mathcal{K}(E,F)$ (resp. $T \in \mathcal{K}(E,F)$ and $S \in \mathcal{L}(F,G)$), then $S \circ T \in \mathcal{K}(E,G)$.
(2) If $T \in \mathcal{K}(E,F)$, then $T^* \in \mathcal{K}(F^*,E^*)$. And conversely.
(3) $\mathcal{K}(E,F)$ is a closed linear subspace of $\mathcal{L}(E,F)$.
*(4) **(Completely continuous)** If $T \in \mathcal{K}(E,F)$, and if $u_n \rightharpoonup u$, then $Tu_n \to Tu$ in F.*

Proof. One can refer to [Bre]. □

Remark 3.56. Property (4) will produce non-trivial effects if we combine it with compact embedding theorems. For instance, if $\Omega \subset \mathbb{R}^n$ is bounded and smooth, then a bounded sequence of functions $(u_k)_{k\geq 1}$ in $H^1(\Omega)$ has a subsequence $(u_{k_j})_{j\geq 1}$ that converges strongly to some $u \in L^2(\Omega)$.

Theorem 3.57 (Fredholm)**.** *Suppose that E is a Banach space and $T \in \mathcal{K}(E)$. Then*

(1) $N(I-T)$ is finite-dimensional.
(2) $R(I-T)$ is closed.
(3) $N(I-T) = \{0\} \Longleftrightarrow R(I-T) = E$.
(4) $\dim N(I-T) = \dim N(I-T^*)$.

Proof. We will give the ideas of these conclusions and then solve them.

(1) *The conclusion is just a direct corollary of theorem 3.17.* Let $E_1 = N(I-T)$. Then $B_{E_1} \subset T(B_E)$ and thus B_{E_1} is compact. By Theorem 3.17, E_1 must be finite-dimensional.
(2) *The conclusion originate in the fact that compactness will lead to the property of convergence. Suppose that*

$$f_n = u_n - Tu_n \to f.$$

We aim to show that $f \in R(I-T)$. Our idea is to find a sequence (x_k) such that

$$f_{n_k} = x_k - Tx_k, \quad x_k \to x, \quad \text{and} \quad Tx_k \to Tx.$$

and then we will get $f = x - Tx \in R(I-T)$. Clearly, we only need to find a sequence (x_k) such that $f_{n_k} = x_k - Tx_k$ and

$$Tx_k \to Tx. \tag{3.4}$$

Note that we can't achieve this idea directly via (u_n), since (3.4) is certainly supposed to be derived by the compactness, which requires that the sequence is bounded. Therefore, in order to find appropriate (x_k), we must make the following transformation first. Note that if $v_n \in N(I-T)$, we have

$$f_n = \widetilde{u}_n - T\widetilde{u}_n \quad \text{where} \quad \widetilde{u}_n = u_n - v_n.$$

Naturally, we want to minimize $\|\widetilde{u}_n\|$. Set $d_n = \text{dist}\,(u_n, N(I-T))$. Since $N(I-T)$ is finite-dimensional and hence is homeomorphic to some $\mathbb{R}^n$, we can choose v_n such that $\|\widetilde{u}_n\| = d_n$. Now it suffices to prove that $(T\widetilde{u}_n)$ has a convergent subsequence, and hence it suffices to prove that $(\widetilde{u}_n)$ is bounded via the compactness.

Suppose for contradiction that $(\widetilde{u}_n)$ is unbounded. Note that[1]

$$\begin{aligned}
(\widetilde{u}_n) \text{ is unbounded} \quad &\iff \quad \exists\, (\widetilde{u}_{n_k}) \text{ such that } \left\|\widetilde{u}_{n_k}\right\| \to \infty \\
&\iff \quad \exists\, (\widetilde{u}_{n_k}) \text{ such that } \frac{\widetilde{u}_{n_k} - T\widetilde{u}_{n_k}}{\left\|\widetilde{u}_{n_k}\right\|} \to 0 \\
&\iff \quad \exists \left(\omega_k = \frac{\widetilde{u}_{n_k}}{\left\|\widetilde{u}_{n_k}\right\|}\right) \text{ such that } \omega_k - T\omega_k \to 0.
\end{aligned}$$

Since (ω_k) is bounded, there exists (ω_{n_k}) such that $T\omega_{n_k} \to z$, and hence

$$\omega_{n_k} \to z \quad \text{ where } z \in N(I-T).$$

Thus $\operatorname{dist}\left(\omega_{n_k}, N(I-T)\right) \to 0$. But note that

$$\operatorname{dist}\left(\omega_{n_k}, N(I-T)\right) = \operatorname{dist}\left(\frac{\widetilde{u}_{n_k}}{\left\|\widetilde{u}_{n_k}\right\|}, N(I-T)\right) = \frac{1}{d_{n_k}}\operatorname{dist}\left(\widetilde{u}_{n_k}, N(I-T)\right) = 1.$$

Contradiction. Thus we can find appropriate (x_k).

(3) *Roughly speaking, our idea is that the condition that*

$$N(I-T) = \{0\} \quad \textit{and} \quad R(I-T) \subsetneq E$$

will give us a strictly decresing chain of closed subspaces which has the property of divergence (to a certain degree), and this will contradict the fact compactness will lead to the property of convergence. Besides, the converse propblem is related to the dual propblem, and hence we can use proposition 3.55.

Suppose for contradiction that $E_1 = R(I-T) \neq E_1$ and $N(I-T) = \{0\}$. Then letting $E_n = (I-T)^n(E)$, we obtain a decresing chain of closed subspaces since

$$E_{n+1} = (I-T)^n(E_1) \subset (I-T)^n(E) = E_n.$$

Moreover, it's a strictly decresing chian, since if

$$E_{n+1} = (I-T)^n(E_1) = (I-T)^n(E) = E_n \quad \text{ for some } n \in \mathbb{N},$$

then $E_1 = E$ (since $(I-T)$ is injective), a contradiction. Now via Riesz's lemma, we may construct a sequence (u_n) such that

$$u_n \in E_n \quad \|u_n\| = 1 \quad \operatorname{dist}(u_n, E_{n+1}) \geq 1/2 \quad \forall n \in \mathbb{N}.$$

Then for $n > m$ we have

$$Tu_n - Tu_m = (-u_n + Tu_n + u_m - Tu_m + u_n) - u_m \geq \operatorname{dist}(u_m, E_{m+1}) \geq 1/2.$$

This is impossible, since T is a compact operator. Hence

$$N(I-T) = \{0\} \implies R(I-T) = E. \tag{3.5}$$

Conversely, assume that $R(I-T) = E$. Via theorem 3.50 we know

$$N(I-T^*) = R(I-T)^{\perp} = \{0\}.$$

[1] Normalization is the basic method and the following equivalence is natural. One can easily see that the unboundedness will lead to a contradiction.

Since $T^* \in K(E^*)$, property (3.5) yields $R(1-T^*) = E^*$. Then theorem 3.50 yields

$$N(I-T) = R(I-T^*)^\perp = \{0\}.$$

Now we get the conclusion.

(4) *Set* $d = \dim N(I-T)$ *and* $d^* = \dim N(1-T^*)$. *Naturally, we want to shift all the dimensions into E and then compare them.*

By (1) and proposition 3.55 (2) we know $d^* < +\infty$, and then via proposition 3.47, theorem 3.51 and (2) we know that

$$R(I-T) = N(I-T^*)^\perp$$

has a complement in E, denoted by F, of dimension d^*.

(4.1) We will first prove that $d^* \leq d$.

Suppose for contradiction that $d < d^*$. Then there is a linear map

$$\Lambda : N(I-T) \to F$$

that is injective and not surjective.[2] We will derive the contradiction similarly as in footnote 2 via (3).

First we define the projection P. Via (1), $N(I-T)$ admits a complement in E via proposition 3.47. Thus, by corollary 3.23, there exists a continuous projection

$$P : E \twoheadrightarrow N(I-T).$$

Then we put[3]

$$\widetilde{f} = -\Lambda \circ P + (I-T) : E \to E, \quad u \mapsto -\Lambda \circ Pu + (I-T)u.$$

Clearly, $N(\widetilde{f}) = \{0\}$. Note that $I - \widetilde{f} = \Lambda \circ P + T$ is compact since $\Lambda \circ P$ has finite rank. Then via (3) we know $R(\widetilde{f}) = E$, which certainly contradicts the fact that Λ is not surjective.

(4.2) Applying (4.1) to T^*, we obtain

$$\dim N(I-T^{**}) \leq \dim N(I-T^*) \leq \dim N(I-T).$$

Based on the canonical injection from E to E^{**}, we know via theorem 3.50

$$N(I-T^{**}) \supset N(I-T)$$

and therefore $d = d^*$.

□

Remark 3.58. If we change I into an isomorphism A, the conclusion is also true. Certainly one can prove this similarly, but this can also be proved by the properties of composition; that is, using the fact that $A \circ (I-T) = A - A \circ T$ where $A \circ T$ is also compact.

---

[2] *Now our aim is to find the contradiction. Take the simplest situation that E is finite-dimensional as an example. Then it's obvious that there doesn't exist such* Λ, *since*

$$f : E \to N \oplus R, \quad u \mapsto (Pu, (I-T)u)$$

is an isomorphism and hence together with $E \cong F \oplus R$ *implies* $\dim F = \dim N$, *where P is a natural projection in the natural isomorphism* $E \xrightarrow{\cong} N \oplus E/N$.

[3] *We only know that F and R are complements mutually. Hence we transform f naturally. In another word, we turn to emphasize the fact that* $E \cong F \oplus R$ *and find the contradiction based on this.*

3.J. **Operator equations, spectrum on $\mathbb{C}$.** At the end of this section, we make a brief introduction to operator equations and the spectrums of operators (on $\mathbb{C}$). One can refer to section 5 for a further study.

Operator equations are the equations about operators and elements, which also give us a new perspective to show the strength of preceding results. For instance, note that

$$\begin{aligned} y \in R(A) \quad &\Longleftrightarrow \quad Ax = y \text{ has a solution;} \\ N(A) = \{0\} \quad &\Longleftrightarrow \quad \text{if } Ax = y \text{ has a solution, then the solution is unique.} \end{aligned}$$

Then theorem 3.57 yields

$$\begin{aligned} T \in \mathcal{K}(E) : (I - T)x = y &\text{ has at most one solution } \forall y \\ \Longleftrightarrow (I - T)x = y &\text{ has at least one solution } \forall y. \end{aligned}$$

Clearly, the preceding theorems correspond to conclusions of solving operator equations.

As we do with linear maps, in the next we focus on the basic operator equation

$$(\lambda I - A)(x) = y.$$

This leads to the concept of *spectrum*.

Definition 3.59 (Spectrum on $\mathbb{C}$)**.** *Let X be a Banach space over $\mathbb{C}$, and let $A \in \mathcal{L}(X, X)$.*

*(1) The **resolvent set**, denoted by $\rho(A)$, is defined by*

$$\rho(A) = \{\lambda \in \mathbb{C} : (\lambda I - A) \textit{ is bijective}\}$$

*(2) The **spectrum**, denoted by $\sigma(T)$, is the complement of the resolvent set, i.e., $\sigma(T) = \mathbb{C} \setminus \rho(T)$.*

(3) Furthermore,

$$\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$$

where

$$\begin{aligned} \sigma_p(A) &= \{\lambda \in \mathbb{C} : N(\lambda I - A) \neq 0\}; \\ \sigma_c(A) &= \left\{\lambda \in \mathbb{C} : N(\lambda I - A) = 0, R(\lambda I - A) \neq X, \overline{R(\lambda I - A)} = X\right\}; \\ \sigma_r(A) &= \left\{\lambda \in \mathbb{C} : N(\lambda I - A) = 0, \overline{R(\lambda I - A)} \neq X\right\}. \end{aligned}$$

Remark 3.60. Clearly, $\mathbb{C} = \rho(A) \cup \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$.

In the next we introduce some basic properties of spectrums for general operators.

Theorem 3.61. *Let X be a Banach space over $\mathbb{C}$, and let $A \in \mathcal{L}(X, X)$. Then*

*(1) **($\sigma(A)$ is bounded)** If*

$$|\lambda| > r(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \inf_{n\geq 1} \|A^n\|^{1/n},$$

then $\lambda \in \rho(A)$, and

$$(\lambda I - A)^{-1} = \sum_{n=0}^{\infty} \frac{A^n}{\lambda^{n+1}},$$

where RHS is the limit of $S_m = \sum_{n=0}^{m} \frac{A^n}{\lambda^{n+1}}$ with respect to the normal operator norm.

(2) ***($\rho(A)$ is open)*** *If $\lambda_0 \in \rho(A)$, then*

$$r_{\lambda_0} := \lim_{n\to\infty} \|(\lambda_0 I - A)^{-n}\|^{1/n} < \infty,$$

we have $B(\lambda_0, r_{\lambda_0}^{-1}) \subset \rho(A)$, and

$$(\lambda I - A)^{-1} = \sum_{n=0}^{\infty} (\lambda_0 I - A)^{-(n+1)} (\lambda_0 - \lambda)^n, \quad \forall |\lambda - \lambda_0| < r_{\lambda_0}^{-1}.$$

(3) ***(The radius of spectrum)*** *we have*

$$\sigma(A) \neq \varnothing \quad \text{and} \quad \sup_{\lambda\in\sigma(A)} |\lambda| = r(A).$$

(4) ***(Adjoint operator)*** *We have*

$$\sigma(A) = \sigma(A^*) \quad \text{and} \quad (\lambda I - A^*)^{-1} = \left((\lambda I - A)^{-1}\right)^*, \quad \forall \lambda \in \rho(A).$$

Proof. Well-known. The details are in my handwritten notes. □

Remark 3.62. For every $A \in \mathcal{L}(X, X)$, $\sigma(A)$ is a compact set.

Example 3.63. Let $F \subset \mathbb{C}$ be an arbitrary compact subset, and let $\{\alpha_n\}_{n=1}^{\infty}$ be an arbitrary dense subset of F. Then for $1 \leq p < \infty$, the operator

$$A : \ell^p \to \ell^p, \quad (x_1, x_2, \cdots, x_n, \cdots) \mapsto (\alpha_1 x_1, \alpha_2 x_2, \cdots, \alpha_n x_n, \cdots)$$

satisfies that $F = \sigma(A)$.

Proposition 3.64. *Suppose that E is a Banach space and that $\{T_n\} \subset \mathcal{L}(E)$ $(n = 1, 2, 3, \cdots)$ converges to $T \in \mathcal{L}(E)$. Let λ_0 be a regular value of T. Then λ_0 is also a regular value of T_n when n is sufficiently large, and*

$$\lim_{n\to\infty} (\lambda_0 I - T_n)^{-1} = (\lambda_0 I - T)^{-1}.$$

Proof. Note that

$$\lambda_0 I - T_n = \lambda_0 I - T - (T_n - T) = (\lambda_0 I - T)\left(I - (\lambda_0 I - T)^{-1}(T_n - T)\right) \tag{3.6}$$

and note that

$$\|(\lambda_0 I - T)^{-1}(T_n - T)\| \leq \|(\lambda_0 I - T)^{-1}\| \cdot \|T_n - T\| \xrightarrow{n\to\infty} 0.$$

Hence there exists $N > 0$ such that

$$\|(\lambda_0 I - T)^{-1}(T_n - T)\| < 1 \quad \forall n > N,$$

and hence

$$I - (\lambda_0 I - T)^{-1}(T_n - T)$$

also has bounded inverse operator by theorem 3.61 (1). Hence for $n > N$, by (3.6) we know that λ_0 is also a regular value of T_n. Moreover,

$$\begin{aligned} (\lambda_0 I - T_n)^{-1} &= \left(I - (\lambda_0 I - T)^{-1}(T_n - T)\right)^{-1} (\lambda_0 I - T)^{-1} \\ &= \left(\sum_{k=0}^{\infty} \left[(\lambda_0 I - T)^{-1}(T_n - T)\right]^k\right) (\lambda_0 I - T)^{-1}, \end{aligned}$$

and hence

$$
\begin{aligned}
\|(\lambda_0 I - T_n)^{-1} - (\lambda_0 I - T)^{-1}\| &= \left\| \left(\sum_{k=1}^{\infty} [(\lambda_0 I - T)^{-1}(T_n - T)]^k \right) (\lambda_0 I - T)^{-1} \right\| \\
&\leq \sum_{k=1}^{\infty} \|(\lambda_0 I - T)^{-1}(T_n - T)\|^k \cdot \|(\lambda_0 I - T)^{-1}\| \\
&= \frac{\|(\lambda_0 I - T)^{-1}(T_n - T)\|}{1 - \|(\lambda_0 I - T)^{-1}(T_n - T)\|} \cdot \|(\lambda_0 I - T)^{-1}\| \\
&\leq \frac{\|T_n - T\|}{1 - \|(\lambda_0 I - T)^{-1}\| \cdot \|T_n - T\|} \cdot \|(\lambda_0 I - T)^{-1}\|^2 \xrightarrow{n \to \infty} 0
\end{aligned}
$$

We are done. □

We end with a typical example.

Proposition 3.65. *Suppose that $K(s, t)$ is a continuous function on $[a, b] \times [a, b]$. Define $A : C[a, b] \to C[a, b]$ by*

$$(Ax)(t) = \int_a^t K(s, t)x(s)ds$$

If $\lambda \neq 0$, then for all $y \in C[a, b]$, the equation

$$\lambda x(t) - (Ax)(t) = y(t)$$

has unique solution.

Proof. It's clear that A is (totally) continuous. It suffices to prove that

$$r(A) = \lim_{n \to \infty} \|A^n\|^{1/n} = 0$$

by theorem 3.61 (1). Put $M = \max_{[a,b] \times [a,b]} |K(s, t)|$. Note that

$$|(Ax)(t)| = \left| \int_a^t K(s, t)x(s)ds \right| \leq M(t - a)\|x\|$$

and by induction we easily know that

$$|(A^n x)(t)| \leq \frac{(t - a)^n M^n}{n!} \|x\|.$$

Hence

$$\|A^n\| \leq \frac{(b - a)^n M^n}{n!},$$

and then

$$r(A) = \lim_{n \to \infty} \|A^n\|^{1/n} \leq \lim_{n \to \infty} \frac{(b - a)M}{\sqrt[n]{n!}} = 0.$$

Done. □

We will go further with the spectrum theory in section 5 by using the conclusions in subsection 3.I for compact operators.

4. Hilbert space

Before we develop the operator theory, we might as well introduce the basic theory of Hilbert spaces and study some interesting examples.

Hilbert spaces have better properties than Banach spaces. Many important PDE problems can be solved in the framework of Hilbert spaces.

4.A. Basic knowledge.

Definition 4.1 (Inner product space)**.** *Let H be a vector space over Λ. If a map*

$$(\cdot,\cdot) : H \times H \to \Lambda$$

*is called an **inner product** on H if it satisfies the following properties:*

(1) (Positive definite) $(x,x) \geq 0$, *for all $x \in H$, and* $(x,x) = 0$ *iff* $x = 0$.
(2) (Conjugate symmetric) $(x,y) = \overline{(y,x)}$ *for all $x, y \in H$.*
(3) (Linear in its first argument) $(\alpha x + \beta y, z) = \alpha(x,y) + \beta(y,z)$ *for all $x, y, z \in H$ and $\alpha, \beta \in \Lambda$.*

*We shall call such a space $(H,(\cdot,\cdot))$ an inner product space. The induced norm is $\|x\| = (x,x)^{\frac{1}{2}}$, and we call $(H,(\cdot,\cdot))$ a **Hilbert space** if H is complete.*

Remark 4.2. $(z, \alpha x + \beta y) = \overline{\alpha}(z,x) + \overline{\beta}(z,y)$.

Proposition 4.3. *Let H be an inner product space over Λ. Then*

(1) ***(Schwarz inequality)*** $|(x,y)|^2 \leq (x,x)(y,y)$, *for all $x, y \in H$.*
(2) $(H, \|\cdot\|)$ *is an NVS where* $\|x\| = (x,x)^{\frac{1}{2}}$.
(3) $(\cdot,\cdot)$ *is continuous with respect to each variable.*
(4) ***(Polarization identity)*** *We have*

$$(x,y) = \begin{cases} \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right), & \Lambda = \mathbb{R} \\ \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) + \frac{i}{4}\left(\|x+iy\|^2 - \|x-iy\|^2\right), & \Lambda = \mathbb{C} \end{cases}$$

(5) ***(Parallelogram law)*** *We have*

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

Conversely, if a normed vector space satisfies the parallelogram law, it forms an inner product space via polarization identity.

Proof. Trivial. □

4.B. Orthogonal systems, orthogonal relations.

In the next we introduce two basic tools, the *orthogonal relations (of subspaces)* and the *orthogonal systems.*

Definition 4.4 (Orthogonal)**.** *Let H be an inner product space, and let M be a linear subspace. The **orthogonal complement** of M is defined as*

$$M^\perp = \{x \in H : \langle x,y \rangle = 0, \forall y \in M\}.$$

*For $x, y \in H$, we say that x and y are **orthogonal** if $\langle x,y \rangle = 0$, which is denoted $x \perp y$.*

Theorem 4.5. *Let H be a Hilbert space, and let L be a closed linear subspace. Then $H = L \oplus L^\perp$ and $(L^\perp)^\perp = L$.*

Proof. Note that $L \cap L^\perp = \{0\}$, since if $y \in (L \cap L^\perp)$ then $(y, y) = 0$. Note that

$$\|a + b\|^2 = \langle a + b, a + b\rangle = \langle a, a\rangle + \langle b, b\rangle = \|a\|^2 + \|b\|^2, \quad \forall a \in L, b \in L^\perp.$$

Note that $L^\perp$ is closed via the continuity of inner product. It follows form corollary 3.23 that $M := L + L^\perp$ is closed. Suppose for contradiction that there exists $0 \neq x_0 \in H \setminus M$; then by corollary 3.42 and problem 3.43, there exists $y_0 \in M$ with

$$\|x_0 - y_0\| = \inf_{y \in M} \|x_0 - y\|.$$

It easily follows that $0 \neq (x_0 - y_0) \perp M$. But it's obvious that $M^\perp = \{0\}$; a contraction.

Now it follows that $H = L \oplus L^\perp = L^\perp \oplus (L^\perp)^\perp$. Since $L \subset (L^\perp)^\perp$, it follows then $(L^\perp)^\perp = L$. □

Remark 4.6. *An alternative method to show the contradiction*: suppose for contradiction that there exists $0 \neq x \in H \setminus (L \oplus L^\perp)$. Via corollary 3.25 there exists $f \in H^*$ with $f(L \oplus L^\perp) = \{0\}$ and $f(x) \neq 0$. Via Riesz representation theorem 4.16, there exists $y \in H$ with $f = (y, \cdot)$. Thus $y \in (L \oplus L^\perp)^\perp$, and hence $(y, y) = 0$, i.e. $y = 0$, which contradicts that $(y, x) \neq 0$.

We can also use the conclusions of orthogonal system to prove this theorem.

Corollary 4.7. *Let H be a Hilbert space, and let L be a closed linear subspace. Via theorem 4.5, let $P_L : H \to L$ be the standard projection. Then for $x \in H$,*

$$\|x - P_M(x)\| = \inf_{y \in M} \|x - y\|.$$

Proof. Trivial. □

Corollary 4.8. *Let H be Hilbert and let $M \subset H$. Then $(M^\perp)^\perp$ is the smallest closed space that contains M.*

Proof. I.e. we need to prove that

$$(M^\perp)^\perp = \overline{\operatorname{span}(M)}.$$

It's clear that LHS is closed since the inner product is continuous. Hence it follows from $M \subset (M^\perp)^\perp$ that $\overline{\operatorname{span}(M)} \subset (M^\perp)^\perp$. Note that

$$M \subset \overline{\operatorname{span}(M)} \implies \overline{\operatorname{span}(M)}^\perp \subset M^\perp \implies (M^\perp)^\perp \subset \overline{\operatorname{span}(M)} = \left(\overline{\operatorname{span}(M)}^\perp\right)^\perp,$$

where the last equality comes from theorem 4.5. We are done. □

In the next we develop the theory of orthogonal system.

Definition 4.9 (Orthogonal system). *Let H be an inner product space, and let $\{e_j\}_{j \in J}$ be a family of points in H. $\{e_j\}_{j \in J}$ is called an* ***orthogonal system*** *if $(e_j, e_k) = \delta_{jk}$ for all $j, k \in J$. An orthogonal system is called* ***normal*** *if $\left(\operatorname{span}\{e_j\}_{j \in J}\right)^\perp = \{0\}$, and is called* ***complete*** *if* *we have*

$$x = \sum_{j\in J}(x, e_j)e_j, \quad \forall x \in H.$$

where RHS must be a finite sum or a countable sum that converges.

Remark 4.10. An orthogonal system is normal iff the representation of 0 is unique. Also note that being complete implies being normal.

Proposition 4.11. *For any nonempty inner product space H, there exists a normal orthogonal system on H.*

Proof. Trivial. Use Zorn's lemma. □

Proposition 4.12. *Let H be an inner product space, and let $\{e_j\}_{j\in J}$ be an orthogonal system. Then for each $x \in H$, the set $A := \{j \in J : (x, e_j) \neq 0\}$ is at most countable, and we have*

$$\sum_{j\in J} |(x, e_j)|^2 := \sum_{j\in A} |(x, e_j)|^2 \leq \|x\|^2. \tag{4.1}$$

Proof. If A is finite, then

$$x - y \perp y, \text{ where } y = \sum_{k=1}^{n} (x, e_{j_k})\, e_{j_k},$$

and hence

$$\|x\|^2 = \|x - y\|^2 + \|y\|^2 \geq \|y\|^2 = \sum_{k=1}^{n} |(x, e_{j_k})|^2.$$

Then (4.1) follows. It follows that

$$A_n := \left\{ j \in J : |(x, e_j)| \geq \frac{1}{n} \right\}$$

is a finite set for each n, and hence $A = \bigcup_{n=1}^{\infty} A_n$ is at most countable. Finally, jsut take limit to get (4.1) based on the fintie conclusion. □

Example 4.13. Set $x_1 = \sum_{n=1}^{\infty} 2^{-n} e_n$, where $e_n = (\underbrace{0, \cdots, 0}_{n-1}, 1, 0, \cdots)$, $n = 1, 2, \cdots$, and set

$$X = \left\{ \sum_{k=2}^{m} \alpha_k e_k + \alpha_1 x_1 : \alpha_k \in \mathbb{R}, k = 1, 2, \cdots. \right\}$$

Then X is a linear subspace of ℓ^2. Note that $\{e_k\}_{k=2}^{\infty}$ is normal but not complete.

Theorem 4.14. *Let H be a Hilbert space, and let $\{e_j\}_{j\in J}$ be an orthogonal system. Then $\{e_j\}_{j\in J}$ is normal iff $\{e_j\}_{j\in J}$ is complete, and we have*

$$\|x\|^2 = \sum_{j\in J} |(x, e_j)|^2.$$

Proof. "$\Longleftarrow$" is trivial.

"$\Longrightarrow$": write $A = \{j \in J : (x, e_j) \neq 0\} = \{e_1, e_2, \cdots\}$ and set $x_n = \sum_{k=1}^{n} (x, e_k)e_k$. It follows from (4.1) that (x_n) is Cauchy and hence converges to some $x' \in X$. Since $\{e_j\}_{j\in J}$ is normal, it's easy to see that $x = x'$. Then the result easily follows. □

Moreover, we can give a new method to prove theorem 4.5.

A new proof of theorem 4.5: Since $L \cap L^{\perp} = \{0\}$, it suffices to show that $L + L^{\perp} = H$.

Given $x \in H$. Since L is a closed linear subspace, L itself is also a Hilbert space. Based on proposition 4.11 and theorem 4.14, there exists a complete orthogonal system $\{e_j\}_{j \in J}$ of L. By proposition 4.12 we know that

$$A := \{j \in J : (x, e_j) \neq 0\} \text{ is at most countable, and } \sum_{j \in A} (x, e_j)e_j \text{ converges in } M.$$

Then we set

$$y = \sum_{j \in A} (x, e_j)e_j.$$

It follows that $y \in M$ and $x - y \perp e_j$ for each $j \in J$, and hence $x - y \perp L$. We are done. $\square$

Corollary 4.15. *If H is a separable Hilbert space over Λ, and let $\{e_j\}_{j \in J}$ be a complete orthogonal system. Then J is at most countable, and H is isometrically isometric to Λ^n or $\ell^2(\Lambda)$.*

Proof. Trivial. $\square$

4.C. **Riesz representation theorem, Bilinear forms, Lax–Milgram theorem.** Using the preceding basic tools, in the next we establish the theory of bilinear forms.

Theorem 4.16 (Riesz representation theorem). *Let H be a Hilbert space. Then for any $f \in H^*$, there exists a unique $y \in H$ such that*

$$f(x) = (x, y), \quad \forall x \in H,$$

and $\|f\| = \|y\|$.

WLOG, in the following proofs we assume that $\Lambda = \mathbb{R}$.

Method 1: orthogonal relations. Note that $N(f)$ is closed and is of codimension 1 (for the case that $f \not\equiv 0$). By theorem 4.5, we just choose an appropriate element in $N(f)^{\perp}$. $\square$

Method 2: orthogonal system. Let $\{e_j\}_{j \in J}$ be a complete orthogonal system of H via proposition 4.11 and theorem 4.14. Setting $f_j = f(e_j)$, we calim that

(1) $A(f) := \left\{j \in J : f_j \neq 0\right\}$ is at most countable;
(2) $\sum_{j \in A(f)} |f_j|^2 < \infty$.

One can easily show that

$$A_n(f) := \left\{j \in J : |f_j| \geq n^{-1}\right\}$$

is finite by using the boundedness of f, and hence $A(f) = \bigcup_{n=1}^{\infty} A_n(f)$ is at most countable. (2) also easily follows from the boundedness of f. Then setting $y = \sum_{j \in J} f_j e_j$ we obtain the result. $\square$

In the next we study the general bilinear forms.

Definition 4.17 (Bilinear form)**.** *Let H be a Hilbert space over Λ. A map*

$$\varphi : H \times H \to \Lambda$$

is called a ***bilinear form*** *if it satisfies:*
(1) $\varphi(\alpha x + \beta y, z) = \alpha\varphi(x, z) + \beta\varphi(y, z)$, for all $x, y, z \in H$ and $\alpha, \beta \in \Lambda$;
(2) $\varphi(z, \alpha x + \beta y) = \overline{\alpha}\varphi(z, x) + \overline{\beta}\varphi(z, y)$, for all $x, y, z \in H$ and $\alpha, \beta \in \Lambda$.
Moreover, a bilinear form is said to be
(1) ***bounded*** *if there exists a constant $M > 0$ such that*

$$|\varphi(x, y)| \leq M\|x\|\|y\|, \quad \forall x, y \in H.$$

(2) ***coercive*** *if there exists a constant $\alpha > 0$ such that*

$$\varphi(x, x) \geq \alpha\|x\|^2, \quad \forall x \in H.$$

Proposition 4.18. *Let H be a Hilbert space over Λ, and let φ be a bilinear form on H. Then*

$$\begin{aligned} \varphi \text{ is bounded} \quad &\iff \quad \exists! A \in \mathcal{L}(H) : \varphi(x, y) = (Ax, y) \\ &\iff \quad \exists! B \in \mathcal{L}(H) : \varphi(x, y) = (x, By) \end{aligned}$$

Proof. Trivial. □

Proposition 4.19. *Let H be a Hilbert space over Λ, and let $A \in \mathcal{L}(H)$. If there exists $\alpha > 0$ with*

$$(Ax, x) \geq \alpha\|x\|^2,$$

then A is bijective.

Proof. If $Ax = 0$, then $\|x\| = 0$ and hence $x = 0$. Thus A is injective.

If $x \perp R(A)$, then $\|x\| = 0$ and hence $x = 0$. Thus $R(A)^\perp = \{0\}$, and it suffices to show that $R(A)$ is closed by theorem 4.5. Suppose that $y_n = Ax_n \to y_0 \in H$. Then

$$\begin{aligned} \|x_n - x_m\|^2 &\leq \alpha^{-1}(A(x_n - x_m), x_n - x_m) \\ &\leq \alpha^{-1}\|A(x_n - x_m)\| \cdot \|x_n - x_m\| \\ &= \alpha^{-1}\|y_n - y_m\| \cdot \|x_n - x_m\| \end{aligned}$$

It follows that (x_n) is Cauchy and hence converges to some $x_0 \in H$. Then by the continuity of A we have

$$Ax_0 = \lim_{n\to\infty} Ax_n = \lim_{n\to\infty} y_n = y_0,$$

which implies that $y_0 \in R(A)$. □

To sum up, we obtain the Lax–Milgram theorem:

Corollary 4.20 (Lax-Milgram theorem)**.** *Let H be a Hilbert space, and let φ be a bounded and coercive bilinear form. Then for all $f \in H^*$, there exists a unique $y \in H$ such that*

$$f(x) = \varphi(x, y), \quad \forall x \in H.$$

Proof. It follows directly from proposition 4.18, proposition 4.19 and the Riesz representation theorem 4.16. □

4.D. **Sovling elliptic PDE.** An important application of Lax–Milgram theorem 4.20 is sovling elliptic PDE.

Theorem 4.21. *Let $\Omega \subset \mathbb{R}^N$ be open and bounded. Given the elliptic operator*

$$L : H^1(\Omega) \to H^{-1}(\Omega), \quad u \mapsto -\partial_j \left(a^{ij}\partial_i u + d^j u\right) + b^i \partial_i u + cu,$$

where $a^{ji} = a^{ij}$, $a^{ij} \in L^\infty(\Omega)$ and there exist constants $0 < \lambda < \Lambda$ such that

$$\lambda|\xi|^2 \le a^{ij}(x)\xi_i\xi_j \le \Lambda|\xi|^2 \quad \forall \xi \in \mathbb{R}^n \quad \forall x \in \Omega; \tag{4.2}$$

$$\sum_{i=1}^n \left\|b^i\right\|_{L^n(\Omega)} + \sum_{i=1}^n \left\|d^i\right\|_{L^n(\Omega)} + \|c\|_{L^{n/2}(\Omega)} \le \Lambda. \tag{4.3}$$

Suppose that $v \in H^{-1}(\Omega)$, $g \in H^1(\Omega)$. Then there exist $\overline{\mu} > 0$, such that for $\mu \ge \overline{\mu}$, the (Dirichlet) elliptic equation

$$\begin{cases} Lu + \mu jiu = v \\ u - g \in H_0^1(\Omega) \end{cases}$$

has a unique solution $u \in H^1(\Omega)$, where

$$i : H_0^1(\Omega) \hookrightarrow L^2(\Omega) \text{ is the compact imbedding.}$$
$$j : L^2(\Omega) \to H^{-1}(\Omega), \quad u \mapsto (u, \cdot)_{L^2(\Omega)}.$$

Proof. Note that our (Dirichlet) elliptic equation can be transformed into

- Finding $u \in H_0^1(\Omega)$ such that $Lu + \mu jiu = w$, where $w \in H^{-1}(\Omega)$.

Note that

$$\begin{aligned} \langle Lu + \mu jiu, v\rangle &= \left\langle -\partial_j \left(a^{ij}\partial_i u + d^j u\right) + b^i\partial_i u + (c+\mu)u, v\right\rangle \\ &= \left\langle a^{ij}\partial_i u + d^j u, \partial_j v\right\rangle + \left\langle b^i\partial_i u + (c+\mu)u, v\right\rangle \end{aligned}$$

Thus the equation is equivalent to

- Finding $u \in H_0^1(\Omega)$ such that $a(u,\cdot) = w$, where
$$\begin{aligned} a : H_0^1(\Omega) \times H_0^1(\Omega) &\to \mathbb{R} \\ (u,v) &\mapsto \int_\Omega \left(a^{ij}\partial_i u \partial_j v + d^j u \partial_j v + b^i(\partial_i u)v + (c+\mu)uv\right) dx \end{aligned}$$
is a continuous bilinear form.[4]

Now we claim that

(i) There exists $\overline{\mu} > 0$ such that a is coercive for $\mu \ge \overline{\mu}$.

Note that by Lax–Milgram theorem (corollary 4.20), the conclusion will follow from (i). Thus it suffices to prove (i). Note that

- Claim (i) is easy if the coefficients are in $L^\infty(\Omega)$.

[4] Continuity (i.e. boundedness) follows from (4.2), (4.3), Hölder inequality, Sobolev inequality and

$$|xAy^\top| \le \sqrt{xAy^\top(xAy^\top)^\top} = \sqrt{xA(y^\top y)Ax^\top} = |y|\sqrt{xAAx^\top} = \frac{|y|}{|x|}\sqrt{(xAx^\top)(xAx^\top)} \le \Lambda|x||y|$$

But we only have (4.3). Our idea is to show that the gap between (4.3) and L^∞ can be controlled. Note that via Poincaré inequality there exists $c_0 > 0$ such that

$$\|\nabla u\|_{L^2(\Omega)} \geq \frac{2c_0}{\lambda} \|u\|_{H_0^1(\Omega)} \quad \forall u \in H_0^1(\Omega).$$

Then we choose $0 < \varepsilon < c_0$ and then find b_1^i, b_2^i, d_1^i, d_2^i, c_1, c_2 and k via lemma 4.22 such that

$$\begin{aligned}
\sum_{i=1}^n \left\|b_1^i\right\|_{L^\infty(\Omega)} + \sum_{i=1}^n \left\|d_1^i\right\|_{L^\infty(\Omega)} + \|c_1\|_{L^\infty(\Omega)} &\leq k \\
\sum_{i=1}^n \left\|b_2^i\right\|_{L^n(\Omega)} + \sum_{i=1}^n \left\|d_2^i\right\|_{L^n(\Omega)} + \|c_2\|_{L^{n/2}(\Omega)} &\leq \varepsilon
\end{aligned}$$

Put

$$\begin{aligned}
a_1(u,v) &= \int_\Omega \left(a^{ij}\partial_i u \partial_j v + d_1^j u \partial_j v + b_1^i(\partial_i u)v + c_1 uv\right) dx \\
a_2(u,v) &= \int_\Omega \left(d_2^j u \partial_j v + b_2^i(\partial_i u)v + c_2 uv\right) dx \\
a_3(u,v) &= \left(k + \frac{2k^2}{\lambda}\right) \int_\Omega uv dx
\end{aligned}$$

Then[5]

$$\begin{aligned}
a_1(u,u) &\geq \lambda \|\nabla u\|_{L^2(\Omega)}^2 - k\left(\|u\|_{L^2(\Omega)} \|\nabla u\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)} \|\nabla u\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)}\right) \\
&= \lambda \|\nabla u\|_{L^2(\Omega)}^2 - k\|u\|_{L^2(\Omega)}^2 - 2k\|\nabla u\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \\
&= \frac{\lambda}{2} \|\nabla u\|_{L^2(\Omega)}^2 - \left(k + \frac{2k^2}{\lambda}\right) \|u\|_{L^2(\Omega)}^2 + \frac{\lambda}{2} \|\nabla u\|_{L^2(\Omega)}^2 + \frac{2k^2}{\lambda} \|u\|_{L^2(\Omega)}^2 - 2k\|\nabla u\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \\
&\geq \frac{\lambda}{2} \|\nabla u\|_{L^2(\Omega)}^2 - \left(k + \frac{2k^2}{\lambda}\right) \|u\|_{L^2(\Omega)}^2 \\
&\geq c_0 \|u\|_{H_0^1(\Omega)}^2 - \left(k + \frac{2k^2}{\lambda}\right) \|u\|_{L^2(\Omega)}^2
\end{aligned}$$

and just like in footnote 4, via Hölder inequality and Sobolev inequality, we have

$$|a_2(u,u)| \leq \varepsilon \|u\|_{H_0^1(\Omega)}^2$$

Thus

$$a_1(u,u) + a_2(u,u) + a_3(u,u) \geq (c_0 - \varepsilon) \|u\|_{H_0^1(\Omega)}^2$$

which proves claim (i). Hence the conclusion follows. □

Lemma 4.22. *Given $f \in L^p(\Omega)$ and $\varepsilon > 0$. Then we can find $f = f_1 + f_2$ such that*

$$\sup_{x\in\Omega} |f_1(x)| \leq k(\varepsilon) \qquad \|f_2\|_{L^p(\Omega)} \leq \varepsilon$$

[5] Note that if we use $c_1 \|\nabla u\|_{L^2(\Omega)} + c_2 \|u\|_{L^2(\Omega)}$ to controll $-\|\nabla u\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)}$, we can use a small c_1. But if we use $c_1 \|\nabla u\|_{L^2(\Omega)}$ and Poincaré inequality to controll $-\|\nabla u\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)}$, we can't use a arbitrarily small c_1.

Proof. Put $A_k = \{x \in \Omega : |f| < k\}$, $B_k = \Omega \setminus A_k$ and

$$f_{1k} = f\chi_{A_k} \qquad f_{2k} = f\chi_{B_k}$$

Then we know

$$f_{1k} + f_{2k} = f \qquad \sup_{x\in\Omega} |f_1(x)| \leq k$$

for all k. Note that $f \in L^p(\Omega)$ implies that $m(B_k) \to 0$ as $k \to \infty$, and then via the Lebesgue dominated convergence theorem 9.25 we know

$$\lim_{k\to+\infty} \|f_{2k}\|_{L^p(\Omega)} = 0$$

Thus $\forall \varepsilon > 0$, we can find an appropriate k such that f_{1k} and f_{2k} satisfy the requirements. □

Corollary 4.23. *In theorem 4.21, if $\mu \geq \overline{\mu}$, the operator $L + \mu ji : H_0^1(\Omega) \to H^{-1}(\Omega)$ is actually an isomorphism.*

Proof. Note that $L + \mu ji : H_0^1(\Omega) \to H^{-1}(\Omega)$ is continuous since the corresponding bilinear form a is continuous. Thus the conclusion follows from theorem 4.21 and theorem 3.22. □

Example 4.24. The operator $-\Delta + cji : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isomorphism, where $c \geq 0$ a.e. and $c \in L^{n/2}(\Omega)$.

Proof. The corresponding bilinear form is

$$a(u,v) = \int_\Omega (\partial_i u \cdot \partial_i v + cuv)\, dx$$

Then via Hölder inequality, Sobolev inequality and Poincaré inequality we have

$$|a(u,v)| \leq \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} + c_1 \|u\|_{H_0^1(\Omega)}^2 \|v\|_{H_0^1(\Omega)}^2 \leq c_2 \|u\|_{H_0^1(\Omega)}^2 \|v\|_{H_0^1(\Omega)}^2$$

and via Poincaré inequality we have

$$a(u,u) \geq \|\nabla u\|_{L^2(\Omega)}^2 \geq c_3 \|u\|_{H_0^1(\Omega)}^2$$

Hence the conclusion follows. □

5. Spectrum theory

In this section we revisit the concept of spectrum. For special operators we will go further by using the preceding conclusions.

5.A. Fredholm alternative, solvability problems.

Basic on the preceding theorems, we make some preparation.

Theorem 5.1 (Fredholm alternative). *Let E be a Banach space and $T \in \mathcal{K}(E)$. Then*

(1) $N(I-T)$ is finite-dimensional.
(2) $R(I-T)$ is closed, and more precisely $R(I-T) = N(I-T^*)^{\perp}$.
(3) $N(I-T) = \{0\} \Longleftrightarrow R(I-T) = E$.
(4) $\dim N(I-T) = \dim N(I-T^*)$.

Proof. Recall that we have proved (1)(3)(4) and that $R(I-T)$ is closed in theorem 3.57. Note that $(I-T)^* = I - T^*$. Hence (2) follows from via theorem 3.51. □

Remark 5.2. Similarly to remark 3.58, if we change I into an isomorphism A, the conclusion is also true.

An important application of Fredholm is to deal with the *solvability problem*.

Remark 5.3. The Fredholm alternative deals with the **solvability** of the equation

$$u - Tu = f$$

It says that

(1) *either* for every $f \in E$ the equation $u - Tu = f$ has a unique solution,
(2) *or* the homogeneous equation $u - Tu = 0$ admits n linearly independent solutions, and in this case, the inhomogeneous equation $u - Tu = f$ is solvable iff f satisfies n orthogonal conditions, i.e.,

$$f \in N(1-T^*)^{\perp}.$$

Again, we take epplitic PDE as a typical example.

Theorem 5.4 (Solvability of elliptic PDE). *Let $\Omega \subset \mathbb{R}^N$ be open and bounded. Consider the elliptic operator*

$$L : H^1(\Omega) \to H^{-1}(\Omega), u \mapsto -\partial_j \left(a^{ij}\partial_i u + d^j u\right) + b^i \partial_i u + cu,$$

where $a^{ji} = a^{ij}$, $a^{ij} \in L^{\infty}(\Omega)$ and there exist constants $0 < \lambda < \Lambda$ such that

$$\lambda|\xi|^2 \le a^{ij}(x)\xi_i\xi_j \le \Lambda|\xi|^2 \quad \forall \xi \in \mathbb{R}^n \quad \forall x \in \Omega;$$

$$\sum_{i=1}^{n} \left\|b^i\right\|_{L^n(\Omega)} + \sum_{i=1}^{n} \left\|d^i\right\|_{L^n(\Omega)} + \|c\|_{L^{n/2}(\Omega)} \le \Lambda.$$

Suppose that $v \in H^{-1}(\Omega)$. Then for the (Dirichlet) elliptic equation

$$Lu = v, \quad u \in H_0^1(\Omega),$$

we have

(1) either *for every $v \in H^{-1}(\Omega)$ the equation has a unique solution,*

(2) or *the homogeneous equation* $Lu = 0$ *admits* n *linearly independent solutions, and in this case, the inhomogeneous equation* $Lu = f$ *is solvable iff* f *satisfies* n *orthogonal conditions.*

Remark 5.5. Note that any (Dirichlet) elliptic equation can be transformed into finding $u \in H_0^1(\Omega)$ such that $Lu = v$, where $v \in H^{-1}(\Omega)$. Theorem 5.4 is just a direct corollary of theorem 4.21 and theorem 5.1.

Proof. Note that the equation is equivalent to:

$$\text{Finding } u \in H_0^1(\Omega) \text{ such that } (L + \mu ji)u - \mu jiu = v, \text{ where } \mu \in \mathbb{R} \text{ and}$$
$$i : H_0^1(\Omega) \hookrightarrow L^2(\Omega) \text{ is the compact imbedding;}$$
$$j : L^2(\Omega) \to H^{-1}(\Omega), \quad u \mapsto (u, \cdot)_{L^2(\Omega)}.$$

We can find an appropriate μ such that $L+\mu ji$ is isomorphism (corollary 4.23). Therefore, the elliptic equation is equivalent to:

$$\text{Finding } u \in H_0^1(\Omega) \text{ such that } u - Tu = w, \text{ where}$$
$$T = \mu(L + \mu ji)^{-1} ji \text{ is compact via proposition 3.55;}$$
$$w = (L + \mu ji)^{-1}(v) \in H_0^1(\Omega).$$

Thus as remark 5.3 claims,

(1) *either* for every $f \in H_0^1(\Omega)$ the equation $u - Tu = f$ has a unique solution,
(2) *or* the homogeneous equation $u - Tu = 0$ admits n linearly independent solutions, and in this case, the inhomogeneous equation $u - Tu = f$ is solvable iff f satisfies n orthogonal conditions, i.e.,

$$f \in N(1 - T^*)^\perp$$

Then the conclusion follows. □

Remark 5.6. In theorem 5.4, Fredholm alternative 5.1 also implies that any eigenspace of a elliptic operator is finite-dimensional.

5.B. **Spectrum on $\mathbb{R}$, eigenvalues of epplitic PDE, compactness of spectrum.** In the next we revisit the concept of spectrum, and show how it helps us analyze the eigenvalues of elliptic PDE. In this section, our spectrum theory is on $\mathbb{R}$ (not on $\mathbb{C}$ as before).

Definition 5.7 (Spectrum on $\mathbb{R}$). *Let E be Banach, and let $T \in \mathcal{L}(E)$.*

*(1) The **resolvent set**, denoted by $\rho(T)$, is defined by*

$$\rho(T) = \{\lambda \in \mathbb{R} : (T - \lambda I) \text{ is bijective from } E \text{ onto } E\}.$$

*(2) The **spectrum**, denoted by $\sigma(T)$, is the complement of the resolvent set, i.e., $\sigma(T) = \mathbb{R} \setminus \rho(T)$.*

*(3) A real number λ is said to be an **eigenvalue** of T if*

$$N(T - \lambda I) \neq \{0\},$$

*and $N(T - \lambda I)$ is the corresponding **eigenspace**.*

(4) The set of all eigenvalues is denoted by $EV(T)$.

Remark 5.8. It's clear that $EV(T) \subset \sigma(T)$. In general, this inclusion can be strict. One can refer to [Bre] section 6.3 for the example of strict inclusion.

In addition to helping us analyze the operator itself, the spectrum has many uses. When we deal with the eigenvalues problem of epplitic PDE, the spectrum theory will play a vital role.

Proposition 5.9. *Using the notations in theorem 4.21, then there exists $\mu \in \mathbb{R}$ such that*

$$S = (L + \mu ji)^{-1} ji$$

is a compact operator, and then

$$\lambda \in \mathbb{R} : N(L + \lambda ji) \neq \{0\} \Longleftrightarrow \lambda \neq \mu \quad \text{and} \quad \frac{1}{\mu - \lambda} \in EV(S). \tag{5.1}$$

Proof. Theorem 4.21 yields $\overline{\mu}$. By corollary 4.23 and proposition 3.55 (1), for $\mu \geq \overline{\mu}$, S is a compact operator. Note that for $\mu \neq \lambda$ and $\mu \geq \overline{\mu}$ we have

$$N(L + \lambda ji) = N\left(L + \mu ji + (\lambda - \mu)ji\right) = N\left(I + (\lambda - \mu)S\right) = N\left(S - \frac{1}{\mu - \lambda} I\right). \tag{5.2}$$

Then property 5.1 follows. If $\lambda = \mu$, it's clear that $N(L + \lambda ji) = \{0\}$. □

Remark 5.10. We will show in the next subsection that $\sigma(S) \setminus \{0\} = EV(S) \setminus \{0\}$.

In the next we introduce the compactness of the spectrum, which is derived by the contraction mapping principle.

Theorem 5.11. *Let E be Banach, and let $T \in \mathcal{L}(E)$. The spectrum $\sigma(T)$ is compact and*

$$\sigma(T) \subset \left[-\|T\|, +\|T\|\right].$$

Proof. We prove it by two steps.

(1) Claim: $\sigma(T) \subset \left[-\|T\|, +\|T\|\right]$.

Our idea is: the existence and uniqueness of solution is equivalent to the existence and uniqueness of fixed point, and the fixed point is related to boundedness, since we have the contraction mapping principle.

Let $\lambda \in \mathbb{R}$ be such that $|\lambda| > \|T\|$. It suffices to show that $T - \lambda I$ is bijective. Now just note that given $f \in E$, the equation $Tu - \lambda u = f$ has a unique solution, since it may be written as $u = \lambda^{-1}(Tu - f)$ and the contraction mapping principle applies.

(2) Claim: $\rho(T)$ is open.

We have the similar idea: the existence and uniqueness of solution is equivalent to the existence and uniqueness of fixed point, and then we try to use the contraction mapping principle.

Let $\lambda_0 \in \rho(T)$, $\lambda \in \mathbb{R}$ (close to λ_0) and $f \in E$. We try to solve

$$Tu - \lambda u = f.$$

This is equivalent to

$$Tu - \lambda_0 u = f + (\lambda - \lambda_0)u,$$

i.e.,

$$u = (T - \lambda_0 I)^{-1} [f + (\lambda - \lambda_0)u].$$

Via contraction mapping principle again, there exists a unique solution if

$$|\lambda - \lambda_0| < \frac{1}{\|(T - \lambda_0 I)^{-1}\|}.$$

Hence $\rho(T)$ is open.

□

5.C. **Spectrum of compact operator.** In the next we introduce the properties of the spectrum of a compact operator.

In addition to using the preceding theorems for compact operators, *another basic idea is to transform information into properties of convergence (or divergence).*

Theorem 5.12. *Let E be Banach with $\dim E = \infty$, and let $T \in \mathcal{K}(E)$. Then we have*

(a) $0 \in \sigma(T)$,
(b) $\sigma(T) \setminus \{0\} = EV(T) \setminus \{0\}$,
(c) one of the following cases holds:
  (i) $\sigma(T) = \{0\}$,
  (ii) $\sigma(T) \setminus \{0\}$ *is a finite set,*
  (iii) $\sigma(T) \setminus \{0\}$ *is a sequence converging to* 0.

Proof. The first two conclusions are just direct corollaries of the preceding theorems.

(a) Suppose for contraction that $0 \notin \sigma(T)$. Then T is bijective and hence $I = T \circ T^{-1}$ is compact via theorem 3.22 and proposition 3.55. Thus B_E is compact and hence $\dim E < \infty$ via corollary 3.19; a contradiction.

(b) Let $\lambda \in \sigma(T)$, $\lambda \neq 0$. We shall prove that λ is an eigenvalue. Suppose not, that $N(T - \lambda I) = \{0\}$. Then by theorem 5.1 (3), we know that $R(T - \lambda I) = E$ and therefore $\lambda \in \rho(T)$; a contradiction.

For (c), it suffices to show that:

For any $n \in \mathbb{Z}_+$, the set $\sigma(T) \cap \{\lambda \in \mathbb{R} : |\lambda| \geq 1/n\}$ is either empty or finite.

Via theorem 5.11, it suffices to prove that:

All the points of $\sigma(T) \setminus \{0\}$ are isolated points.

Let $(\lambda_n)_{n \geq 1}$ be a sequence of distinct real numbers such that:

$$\lambda_n \in \sigma(T) \setminus \{0\} \quad \forall n \in \mathbb{N}, \quad \text{and} \quad \lambda_n \to \lambda.$$

It suffices to prove:

$$\lambda = 0. \tag{5.3}$$

Recall that our conditions are $T \in \mathcal{K}(E)$ and $\lambda_n \to \lambda$ where $\lambda_n \in EV(T) \setminus \{0\}$ for all n. A natural idea is that we transform both conditions into properties of convergence (or divergence) and then compare them to get (5.3).

We first transform the latter condition into some property of convergence by the following three steps, where the first two steps are a typical method, via using the tool Riesz lemma 3.17, to build a framework of solving problems.

(1) *For each $\lambda_n \in EV(T) \setminus \{0\}$ we find $e_n \neq 0$ with $e_n \in N(T - \lambda_n I)$. Let E_n be the space spanned by $\{e_1, \cdots, e_n\}$. We claim that $E_n \subsetneq E_{n+1}$, for all n.*

It suffices to check that for all n, the vectors $e_1, \cdots, e_n$ are linearly independent. The proof is by induction on n. Assume that this holds up to n and suppose that

$$e_{n+1} = \sum_{i=1}^{n} \alpha_i e_i.$$

Then by definitions we have

$$Te_{n+1} = \sum_{i=1}^{n} \alpha_i \lambda_i e_i = \lambda_{n+1} \sum_{i=1}^{n} \alpha_i e_i.$$

It follows that $\alpha_i(\lambda_i - \lambda_{n+1}) = 0$ for $i = 1, \cdots, n$ and hence $\alpha_i = 0$ for $i = 1, \cdots, n$; a contradiction. Hence we have proved that $E_n \subsetneq E_{n+1}$, for all n.

(2) *Now via Riesz lemma 3.17 we can construct a sequence $(u_n)_{n\geq 1}$ such that $u_n \in E_n$, $\|u_n\| = 1$ and $\mathrm{dist}(u_n, E_{n-1}) \geq 1/2$ for all $n \geq 2$.*

(3) *(Now we use the condition $\lambda_n \to \lambda$.)* It's clear that $(T - \lambda_n I)E_n \subset E_{n-1}$. Thus for $2 \leq m < n$ we have

$$\begin{aligned} \left\| \frac{Tu_n}{\lambda_n} - \frac{Tu_m}{\lambda_m} \right\| &= \left\| \frac{Tu_n - \lambda_n u_n}{\lambda_n} - \frac{Tu_m - \lambda_m u_m}{\lambda_m} + u_n - u_m \right\| \\ &\geq \mathrm{dist}(u_n, E_{n-1}) \geq \frac{1}{2}. \end{aligned}$$

Suppose for contradiction that $\lambda \neq 0$. Then via $\lambda_n \to \lambda$ we get a sequence $(Tu_n)_{n\geq 1}$ which has no convergent subsequence.

Now note that a compact operator sends a bounded sequence into a sequence that has a convergent subsequence; a contradiction. Hence we prove (5.3), and (c) follows. □

5.D. **Case of Hilbert spaces — bilinear forms, self-adjoint.** Now let us consider the better situation that $E = H$ is a Hilbert space and $T \in \mathcal{L}(H)$. We have more precise characterizations and have more tools in Hilbert spaces, for which one can refer to section 4.

For instance, by proposition 4.18, any operator $T \in \mathcal{L}(H)$ can be represented by the corresponding continuous bilinear form[6]

$$a : H \times H \to \mathbb{R}, \quad (u, v) \mapsto (Tu, v),$$

and we have Lax–Milgram theorem 4.20 to characterize H^*. This gives us a good breakthrough point for a new and better theory.

First of all, let's make a more precise conclusion than theorem 5.11.

[6]In Hilbert space H, the continuity of T is equivalent to the continuity of a via Cauchy-Schwarz inequality.

Theorem 5.13. *Let $T \in \mathcal{L}(H)$ and let a be the corresponding bilinear form. Set*

$$m = \inf_{\substack{u \in H \\ |u|=1}} a(u,u) \quad \text{and} \quad M = \sup_{\substack{u \in H \\ |u|=1}} a(u,u)$$

Then $\max\{|m|, |M|\} \leq \|T\|$ *and* $\sigma(T) \subset [m, M]$.

Proof. Note that

$$|a(u,v)| = |(Tu,v)| \leq \|Tu\| \cdot \|v\| \leq \|T\| \cdot \|u\| \cdot \|v\|.$$

It follows that $\max\{|m|, |M|\} \leq \|T\|$. Let $\lambda > M$; we will prove that $\lambda \in \rho(T)$. Note that

$$(\lambda u - Tu, u) \geq (\lambda - M)\|u\|^2 \quad \forall u \in H.$$

Applying Lax–Milgram theorem 4.20, we deduce that $\lambda I - T$ is bijective and thus $\lambda \in \rho(T)$. Similarly, any $\lambda < m$ belongs to $\rho(T)$ and therefore $\sigma(T) \subset [m, M]$. □

We have the following natural question:

Question 5.14. *Do m and M belong to $\sigma(T)$?*

Consider the problem of M first. By definition we can find a sequence (u_n) such that $\|u_n\| = 1$ for all n and

$$a(u_n, u_n) \to M \quad \text{as } n \to \infty.$$

Now we claim the following lemma, which is a natural sufficient condition.

Lemma 5.15. *If a is* symmetric, *then $(T - MI)(u_n) \to 0$, and hence $M \in \sigma(T)$.*

Proof. We first prove that $(T - MI)(u_n) \to 0$. Note that

$$b(u,v) = M(u,v)_{H,H} - a(u,v)$$

is symmetric and satisfies

$$b(v,v) \geq 0 \quad \forall v \in H.$$

Hence b satisfies the Cauchy-Schwarz inequality

$$|b(u,v)| \leq b(u,u)^{\frac{1}{2}} b(v,v)^{\frac{1}{2}}.$$

Now put

$$w_n = \frac{Mu_n - Tu_n}{\|Mu_n - Tu_n\|}.$$

It follows that

$$\|(T - MI)(u_n)\| = b\,(u_n, w_n) \leq b(u_n, u_n)^{\frac{1}{2}} b(w_n, w_n)^{\frac{1}{2}}.$$

Note that

$$b(u_n, u_n) = M - a(u_n, u_n) \to 0,$$

and

$$b(w_n, w_n) = M - a(w_n, w_n) \leq M - m.$$

Thus we get $\|(T - MI)(u_n)\| \to 0$.

Now suppose for contradiction that $M \in \rho(T)$. Then

$$u_n = (T - MI)^{-1}((T - MI)(u_n)) \to 0.$$

Contradiction. Hence the conclusion follows. □

Here we use the fact that **any symmetric bilinear form has Cauchy-Schwarz inequality**. We add the property of symmetry and go on.

Definition 5.16 (Self-adjoint)**.** *An operator $T \in \mathcal{L}(H)$ is said to be* ***self-adjoint*** *if $T^* = T$, which is equivalent to that the corresponding bilinear form a is* ***symmetric****, i.e.*

$$(Tu, v) = (u, Tv) \quad \forall u, v \in H.$$

Now we get a deeper conclusion.

Theorem 5.17. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator. Set*

$$m = \inf_{\substack{u \in H \\ |u|=1}} a(u,u) \quad \text{and} \quad M = \sup_{\substack{u \in H \\ |u|=1}} a(u,u).$$

Then $\sigma(T) \subset [m, M]$. Moreover, $m \in \sigma(T)$, $M \in \sigma(T)$ and $||T|| = \max\{|m|, |M|\}$.

Proof. We have proved in theorem 5.13 that

$$\max\{|m|, |M|\} \leq ||T|| \quad \text{and} \quad \sigma(T) \subset [m, M].$$

It follows from lemma 5.15 that $M \in \sigma(T)$. Similarly $m \in \sigma(T)$. Now it suffices to show that

$$||T|| \leq \max\{|m|, |M|\}.$$

Put $\mu = \max\{|m|, |M|\}$. Note that

$$\begin{aligned} 4|a(u,v)| &= |a(u+v, u+v) - a(u-v, u-v)| \\ &\leq \mu\left(||u+v||^2 + ||u-v||^2\right) \\ &= 2\mu\left(||u||^2 + ||v||^2\right). \end{aligned}$$

This is not homogeneous. Setting $v = tw$ we get

$$t|a(u,w)| \leq \frac{1}{2}\mu\left(||u||^2 + t^2||w||^2\right) \quad \forall t \in \mathbb{R}.$$

Hence

$$|a(u,w)| \leq \mu||u|| \cdot ||w|| \quad \forall v, w \in H.$$

For $u \neq 0$, putting $w = (Tu)/||Tu||$ we get

$$||Tu|| \leq \mu||u|| \quad \forall u \in H.$$

Hence $||T|| \leq \mu$, which completes the proof. □

We have the following important corollary.

Corollary 5.18. *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator. Then*

- $T = 0 \iff \sigma(T) = \{0\}$.

Let $A \subset H$ be a linear subspace. Then the following properties are equivalent

(1) $T(A) \subset A$ and $\sigma(T|_A) = \{0\}$.
(2) $A \subset N(T)$.

Proof. Obviously it follows from theorem 5.17. □

Remark 5.19. For a self-adjoint operator, the equivalence serves us new methods to prove that $T = 0$ or that $A \subset N(T)$ for a linear subspace A.

5.E. **Compact self-adjoint operators, eigenvector spaces, spectral decomposition.** Now let's put the preceding good properties together. In the next we will analyze the properties of a *compact self-adjoint operator*.

Definition 5.20 (Eigenvector space)**.** *Based on theorem 5.12, we can let $(\lambda_n)_{n\geq 0}$ be the sequence of all (distinct)* ***eigenvalues*** *of T with $\lambda_0 = 0$ for a compact operator. Then the* ***eigenvector spaces*** *are given by*

$$E_n = N(T - \lambda_n I) \quad \forall n \geq 0.$$

Remark 5.21. Note that via definition and Fredholm alternative (theorem 5.1) we know

$$0 < \dim E_n < \infty \quad \forall n \geq 1.$$

Now we give the first natural conclusion.

Theorem 5.22. *The closed spaces $(E_n)_{n\geq 0}$ are mutually orthogonal.*

Proof. Since the kernel of a bounded operator is closed, E_n is closed for each n. Note that for $u \in E_m$ and $v \in E_n$ with $m \neq n$ we have

$$a(u, v) = \lambda_m (u, v)_{H,H} = \lambda_n (u, v)_{H,H}.$$

Therefore $(u, v)_{H,H} = 0$. □

The second natural conclusion is a property of decomposition, which is derived based on corollary 5.18. This is a more precise characterization of the spectrum that makes full use of the typical properties of Hilbert space.

Theorem 5.23. *Let F be the vector space spanned by the spaces $(E_n)_{n\geq 0}$. Then F is dense in H.*

Proof. Via theorem 4.5, it suffices to prove that $F^\perp = \{0\}$. Basically, we have

- $T(F) \subset F$; obviously this is a direct corollary of definition.
- $T(F^\perp) \subset F^\perp$; indeed, given $u \in F^\perp$ we have

$$(Tu, v) = (u, Tv) = 0 \quad \forall v \in F.$$

- $T_0 = T|_{F^\perp} : F^\perp \to F^\perp$ is a self-adjoint compact operator.

It's very natural to analyze the spectrum of T_0 notting the definition of F. Actually it's obvious that

- $\sigma(T_0) = \{0\}$.

Suppose not; suppose that some $\lambda \neq 0$ belongs to $\sigma(T_0)$. Then via theorem 5.12 this implies that there is some $u \in F^\perp$, $u \neq 0$, such that

$$Tu = \lambda u.$$

Therefore $\lambda = \lambda_n$ for some n via definition. Thus

$$u \in E_n \subset F,$$

which implies $u \in F \cap F^\perp$. However $F \cap F^\perp = \{0\}$; a contradiction.

Now recall corollary 5.18; $T(F^\perp) \subset F^\perp$ and $\sigma(T_0) = \{0\}$ imply that $F^\perp \subset N(T) \subset F$. Thus $F = \{0\}$ since $F \cap F^\perp = \{0\}$. Hence the conclusion follows. □

To sum up, we have the following fundamental theorem.

Theorem 5.24 (Spectral Decomposition)**.** *Let H be a Hilbert space and let T be a compact self-adjoint operator. Then there exists a complete orthogonal system composed of eigenvectors of T.*

Proof. Via theorem 5.22 and theorem 5.23, we just choose in each subspace $(E_n)_{n\geq 0}$ a complete orthogonal system (the existence of such orthogonal systems follows form proposition 4.11 and theorem 4.14), and then union of these orthogonal systems is clearly a complete orthogonal system of H, composed of eigenvectors of T. □

5.F. **Epplitic PDE for $L = \Delta$.** Now we introduce an application in PDE. Recall proposition 5.9 and example 4.24.

Theorem 5.25. *Suppose that the (Dirichlet) epplitic PDE*

$$-\Delta u = \lambda u, \quad u \in H_0^1(\Omega)$$

has a nonzero solution. Then

(1) All possible λ's form a sequence

$$0 < \lambda_1 < \lambda_2 < \cdots$$

where

$$\lim_{n\to\infty} \lambda_n = +\infty$$

(2) The corresponding eigenspaces E_{λ_n}'s satisfy

$$0 < \dim E_{\lambda_n} < \infty \quad \forall n \geq 1$$

(3) There exists a Hilbert basis of $H_0^1(\Omega)$ composed of eigenfunctions.

Proof. Note that $\lambda > 0$ follows from example 4.24. Besides, we can put

$$S = (-\Delta)^{-1} ji$$

and then via (5.1) and theorem 5.12 we get (1). Moreover, via Fredholm alternative 5.1 and (5.2) we get (2), and via theorem 5.24 and (5.2) we get (3). □

Remark 5.26. Note that $N(S) = \{0\}$ here. Generally speaking, we have $0 \leq \dim E_0 \leq +\infty$. Theorem 4.21, example 4.24, proposition 5.9, theorem 5.4 and theorem 5.25 are the main results about PDE in this note.

6. Measurable spaces and measurable maps

In this section we introduce the basic theory of measurable spaces and measurable maps

6.A. **Basic concepts, methods of judging measurability.**

Definition 6.1 (Measurable spaces, measurable maps)**.** *There are some basic concepts.*

*(1) A **measurable spaces** $(X, \mathcal{X})$ is a set X, together with a collection $\mathcal{X}$ of subsets of X which form a σ-algebra.*
*(2) A subset of X is said to be **measurable** with respect to the measurable space if $A \in \mathcal{X}$.*
*(3) We say that one σ-algebra $\mathcal{X}$ on a set X is **coarser** than another $\mathcal{X}'$ if $\mathcal{X} \subset \mathcal{X}'$.*
*(4) A map $f : X \to Y$ from one measurable space $(X, \mathcal{X})$ to another $(Y, \mathcal{Y})$ is said to be **measurable** if $f^{-1}(E) \in \mathcal{X}$ for all $E \in \mathcal{Y}$.*

Proposition 6.2. *Let $(\mathcal{X}_\alpha)_{\alpha\in A}$ be an arbitrary family of σ-algebras on X.*

*(1) The **intersection** $\bigwedge_{\alpha\in A} \mathcal{X}_\alpha := \bigcap_{\alpha\in A} \mathcal{X}_\alpha$ of $(\mathcal{X}_\alpha)_{\alpha\in A}$ is another σ-algebra on X.*
*(2) Given any collection $\mathcal{F}$ of sets on X, the σ-algebra **generated by** $\mathcal{F}$ is defined as the intersection of all the σ-algebras containing $\mathcal{F}$, which is denoted by $\mathcal{B}[\mathcal{F}]$.*
*(3) The **join** $\bigvee_{\alpha\in A} \mathcal{X}_\alpha$ of $(\mathcal{X}_\alpha)_{\alpha\in A}$ is defined as $\mathcal{B}[\bigcup_{\alpha\in A} \mathcal{X}_\alpha]$.*

Remark 6.3. The σ-algebra generated by $\mathcal{F}$ is also the *coarsest* algebra for which all sets in $\mathcal{F}$ are measurable.

Example 6.4. The open sets $\mathcal{F}$ of a topological space $(X, \mathcal{F})$ generate a σ-algebra, known as the **Borel σ-algebra** $\mathcal{B}_X$ of that space.

Example 6.5. The **Lebesgue σ-algebra** $\mathcal{L}$ of Lebesgue measurable sets on a Euclidean space $\mathbb{R}^n$ is the join of the Borel σ-algebra $\mathcal{B}_{\mathbb{R}^n}$ and of the algebras of null sets and their complements (also called co-null sets). See theorem 8.19.

Example 6.6. Let $\overline{\mathbb{R}} = [-\infty, +\infty]$ be the extended real number system. We define Borel sets in $\overline{\mathbb{R}}$ by $\mathcal{B}_{\overline{\mathbb{R}}} = \left\{E \subset \overline{\mathbb{R}} : E \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\right\}$.

Remark 6.7. It's clear that $\mathcal{B}_{\overline{\mathbb{R}}}$ can be generated by $\mathcal{A}_1 = \{[-\infty, a) : a \in \mathbb{R}\}$, $\mathcal{A}_2 = \{[-\infty, a] : a \in \mathbb{R}\}$, $\mathcal{A}_3 = \{(a, \infty] : a \in \mathbb{R}\}$, or $\mathcal{A}_4 = \{[a, \infty] : a \in \mathbb{R}\}$.

Remark 6.8. In the next we abbreviate the measurable space $(\overline{\mathbb{R}}, \mathcal{B}_{\overline{\mathbb{R}}})$ as $\overline{\mathbb{R}}$ or $[-\infty, +\infty]$, and we employ the notation "**real-valued**" to mean "$\overline{\mathbb{R}}$-valued".

Example 6.9 (Borel measurable). A map $f : X \to Y$ from one topological space to another is said to be **Borel measurable** if it is measurable once X and Y are equipped with their respective Borel σ-algebras.

In particular, **continuous maps are Borel measurable**, since the collection $\{E \in \mathcal{B}_Y : f^{-1}(E) \in \mathcal{B}_X\}$ is a σ-algebra that contains all open subsets of Y and hence is $\mathcal{B}_Y$. But the converse statement is false, and the counterexample can be easily derived via proposition 6.14.

Example 6.10 (Lebesgue measurable)**.** A map $f : \mathbb{R}^n \to Y$ is said to be **Lebesgue measurable** if it is measurable from $\mathbb{R}^n$ (with the Lebesgue σ-algebra) to Y (with the Borel σ-algebra).

In the next, we introduce some basic methods of judging measurability.

Proposition 6.11. *Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be measurable spaces. If $\mathcal{Y}$ is* generated *by $\mathcal{E}$, then $f : X \to Y$ is measurable iff $f^{-1}(E)$ is measurable for all $E \in \mathcal{E}$.*

Proof. The "only if" implication is trivial. For the converse statement, note that $\{E \subset Y : f^{-1}(E) \in \mathcal{X}\}$ is a σ-algebra that contains $\mathcal{E}$ and hence contains $\mathcal{Y}$. □

Corollary 6.12. *Let $f : (X, \mathcal{X}) \to \overline{\mathbb{R}}$ and $Y = f^{-1}(\mathbb{R})$. Then f is measurable iff $f^{-1}(\{-\infty\}) \in \mathcal{X}$, $f^{-1}(\{\infty\}) \in \mathcal{X}$, and f is measurable on Y.*

Proposition 6.13. *Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be measurable spaces. If $X = A \cup B$ where $A, B \in \mathcal{X}$, then a map $f : X \to Y$ is measurable iff f is measurable on A and on B.*

Proof. For any $E \in \mathcal{Y}$, $(f|_A)^{-1}(E) = f^{-1}(E) \cap A$, $(f|_B)^{-1}(E) = f^{-1}(B) \cap A$, and $f^{-1}(E) = (f|_A)^{-1}(E) \cup (f|_B)^{-1}(E)$. The result follows. □

6.B. **Pointwise limit of measurable maps.** In the next we introduce some basic properties related to pointwise limit.

Proposition 6.14. *Let $(X, \mathcal{X})$ be a measurable space, let $(Y, \mathcal{B}_Y)$ be a metric space equipped with its Borel σ-algebra, and let $f_n : X \to Y$ be measurable maps for each $n \in \mathbb{N}$ such that the pointwise limit of $\{f_n\}$ exists. Then $f(x) = \lim_{n\to\infty} f_n(x)$ is measurable.*

Proof. Since Y is a metric space, $\mathcal{B}_Y$ is generated by the open balls

$$\mathcal{G}_Y := \{B_Y(y, r) : y \in Y, r > 0\}.$$

Consequently, it suffices to show that $f^{-1}(G) \in \mathcal{X}$ for any $G \in \mathcal{G}_Y$. For $G := B_Y(y, r)$, we have

$$f(x) \in B_Y(y, r) \iff \exists k = k(x), N = N(x) \in \mathbb{N}, \forall n \geq N : f_n(x) \in B_Y\left(y, r - \frac{1}{k}\right).$$

Then,

$$f^{-1}(G) = \bigcup_{\substack{k\in\mathbb{N} \\ \frac{1}{k}<r}} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} f_n^{-1}\left(B_Y\left(y, r - \frac{1}{k}\right)\right)$$

and hence $f^{-1}(G) \in \mathcal{X}$. □

Proposition 6.15. *Let $(X, \mathcal{X})$ be a measurable space, let $(Y, \mathcal{B}_Y)$ be a complete metric space equipped with its Borel σ-algebra, and let $f_n : X \to Y$ be measurable maps for each $n \in \mathbb{N}$. Then $E := \{x : \lim_{n\to\infty} f_n(x)$ exists$\}$ is a measurable set.*

Proof. Define

$$A_{m,n,j} := \left\{x \in X : |f_n(x) - f_m(x)| < \frac{1}{j}\right\}.$$

Note that

$$x \in E \iff \forall j \in \mathbb{N},\ \exists M = M(j), N = N(j) \in \mathbb{N},\ \forall m \geq M, n \geq N : x \in A_{m,n,j}.$$

Hence

$$E = \bigcap_{j=1}^{\infty} \bigcup_{M \in \mathbb{N}} \bigcup_{N \in \mathbb{N}} \bigcap_{m \geq M} \bigcap_{n \geq N} A_{m,n,j},$$

which is measurable. □

Remark 6.16. After giving the concept of measure, we can show that if f_n converges to f a.e., then f_n converges to f almost uniformly. See Egoroff's theorem 7.7. We will derive more properties via measures, which will be introduced later.

Remark 6.17. Moreover, a measurable map can be approximated by a continuous map. In fact, under good conditions ($X = \mathbb{R}^n$, μ^* is a Borel regular outer measure, and $\mu^*(A) < \infty$), any measurable map f on A is almost continuous on A. This is *Lusin's theorem*. One can refer to [Evaa] section 1.2.

6.C. **Approximation — pointwise limit of simple functions.** The above properties inspire us that we may approximate a measurable map via simple measurable maps. We will apply this idea to measurable functions.

First we introduce some basic properties.

Proposition 6.18 (Properties of measurable functions)**.** *Let $(X, \mathcal{X})$ be a measurable space.*

(1) If $f, g : X \to [-\infty, +\infty]$ are measurable, then so are

$$f + g, \quad fg, \quad |f|, \quad \min\{f, g\}, \quad \text{and} \quad \max\{f, g\}.$$

The function $\frac{f}{g}$ is also measurable, provided $g \neq 0$ on X.

(2) If the functions $f_k : X \to [-\infty, +\infty]$ are measurable ($k = 1, 2, \cdots$), then

$$\inf_{k \geq 1} f_k, \quad \sup_{k \geq 1} f_k, \quad \liminf_{k \to \infty} f_k, \quad \text{and} \quad \limsup_{k \to \infty} f_k$$

are also measurable.

Proof. Trivial. Note that

$$\liminf_{k \to \infty} f_k = \sup_{m \geq 1} \inf_{k \geq m} f_k, \quad \limsup_{k \to \infty} f_k = \inf_{m \geq 1} \sup_{k \geq m} f_k$$

for the last assertion. □

Remark 6.19. If f is measurable, then the positive part $f^+(x) := \max\{f(x), 0\}$ and the negative part $f^-(x) := \max\{-f(x), 0\}$ are measurable. If f^+ and f^- are measurable, then $f = f^+ - f^-$ is measurable.

Now we show that **any measurable functions is a pointwise limit of some sequence simple functions**. This approximation is like doing a finer and finer division of the space, and the process of refinement needs to use the property of f, not just the property of the space.

Definition 6.20 (Simple function). *Let $(X, \mathcal{X})$ be a measurable space. A **simple function** on X is a finite linear combination of characteristic functions of sets in $\mathcal{X}$.*

Remark 6.21. A function $f : X \to \mathbb{R}$ (or $\mathbb{C}$) is simple iff f is measurable and the range of f is a finite subset of $\mathbb{R}$ (or $\mathbb{C}$).

Theorem 6.22 (From simple functions to measurable functions). *Let $(X, \mathcal{X})$ be a measurable space.*

(1) If $f : X \to [0, \infty]$ is measurable, there is a sequence (φ_n) of simple functions such that $0 \le \varphi_1 \le \varphi_2 \le \cdots \le f$, $\varphi_n \to f$ pointwise, and $\varphi_n \to f$ uniformly on any set on which f is bounded.

(2) If $f : X \to \mathbb{C}$ is measurable, there is a sequence (φ_n) of simple functions such that $0 \le |\varphi_1| \le |\varphi_2| \le \cdots \le |f|$, $\varphi_n \to f$ pointwise, and $\varphi_n \to f$ uniformly on any set on which f is bounded.

Proof. Suppose that f is real-valued. For $n = 0, 1, 2 \cdots$ and $0 \le k \le 2^{2n} - 1$, let

$$E_n^k = f^{-1}((k2^{-n}, (k+1)2^{-n}]) \quad \text{and} \quad F_n = f^{-1}((2^n, +\infty])$$

and define

$$\varphi_n = \sum_{k=0}^{2^{2n}-1} k2^{-n} \chi_{E_n^k} + 2^n \chi_{F_n}$$

Then $0 \le \varphi_1 \le \varphi_2 \le \cdots \le f$ and

$$\text{if } f(x) \le 2^n, \text{ then } 0 \le f(x) - \varphi_n(x) \le 2^{-n}$$

Then (1) follows.

Suppose that f is complex-valued. If $f = g + ih$, we can apply part (1) to the positive and negative parts of g and h, obtaining sequences $\psi_n^+, \psi_n^-, \zeta_n^+, \zeta_n^-$ of non-negative simple functions that increse to g^+, g^-, h^+, h^-. Let $\varphi_n = \psi_n^+ - \psi_n^- + i(\zeta_n^+ - \zeta_n^-)$; then (2) follows. □

Remark 6.23. Let's make a simple summary of approximations of measurable functions.

(1) In the pointwise sense, we have the above theorem and Egoroff's theorem 7.7. On the other hand, if X is LCH and f is LSC and non-negative, we have proposition 11.14 (5).
(2) We can approximate a measurable map by a continuous map. One can refer to [Evaa] section 1.2 for Lusin's theorem.
(3) In the sense of convergence in measure, we have theorem 7.10.
(4) In the sense of convergence in L^p, any $f \in L^p$ can be approximated by continuous functions, which will be showed in section 11. See corollary 11.10.

6.D. **Pullback — Borel σ-algebra, product measurable space.** In the next we introduce *pullback* and use it to study more spaces, such as the product spaces. Pullback is a typical example that embodies the interaction between spaces and maps on these spaces.

Definition 6.24 (Pullback). *Given a map $f : X \to Y$ from a set X to a measurable space $(Y, \mathcal{Y})$, the **pullback** $f^{-1}(\mathcal{Y})$ of $\mathcal{Y}$ is defined as the σ-algebra $f^{-1}(\mathcal{Y}) := \{f^{-1}(E) : E \in \mathcal{Y}\}$.*

More generally, given a family $(f_\alpha : X \to Y_\alpha)_{\alpha\in A}$ of maps into measurable spaces $(Y_\alpha, \mathcal{Y}_\alpha)$, we can from ***the σ-algebra*** $\bigvee_{\alpha\in A} f_\alpha^{-1}(\mathcal{Y}_\alpha)$ ***generated by the f_α's.***

Remark 6.25. Pullback is a basic tool to describe the relation between different σ-algebras and to construct new σ-algebra. From another perspective, $f^{-1}(\mathcal{Y})$ is the coarsest σ-algebra on X that makes f measurable, and $\bigvee_{\alpha\in A} f_\alpha^{-1}(\mathcal{Y}_\alpha)$ is the coarsest σ-algebra on X that makes f_α simultaneously measurable.

Example 6.26 (Restriction)**.** If E is a subset of a measurable space $(Y, \mathcal{Y})$, the pullback of $\mathcal{Y}$ under the inclusion map $\iota : E \to Y$ is called the **restriction** of $\mathcal{Y}$ to E, and is denoted by $\mathcal{Y}|_E$.

Proposition 6.27. *Let M be a topological space and let N be its open subset. Then $\mathcal{B}_N \subset \mathcal{B}_M$ and $\mathcal{B}_N = \mathcal{B}_M|_N$.*

Proof. If $E \subset N$ is open in N, then E is open in M and $\iota_N^{-1}(E) = E$. Hence $\mathcal{B}_N \subset \mathcal{B}_M$ and $\mathcal{B}_N \subset \mathcal{B}_M|_N$. On the other hand, the collection $\left\{E \in \mathcal{B}_M : \iota_N^{-1}(E) \in \mathcal{B}_N\right\}$ is easily seen to be a σ-algebra on M that contains the open subsets of M and hence $\mathcal{B}_M$. In other words, $\iota_N^{-1}(E) \in \mathcal{B}_N$ for all $E \in \mathcal{B}_M$, and hence $\mathcal{B}_M|_N \subset \mathcal{B}_N$. Thus $\mathcal{B}_N = \mathcal{B}_M|_N$. □

Example 6.28 (Cartesian product)**.** Let $(X_\alpha, \mathcal{X}_\alpha)_{\alpha\in A}$ be a family of measurable spaces, then the **Cartesian product** $\prod_{\alpha\in A} X_\alpha$ has canonical projection maps $\pi_\beta : \prod_{\alpha\in A} X_\alpha \to X_\beta$ for each $\beta \in A$. The product σ-algebra $\prod_{\alpha\in A} \mathcal{X}_\alpha$ is defined as the σ-algebra on $\prod_{\alpha\in A} \mathcal{X}_\alpha$ generated by the π_β's as in definition 6.24.

Proposition 6.29. *Let $(X_\alpha, \mathcal{X}_\alpha)_{\alpha\in A}$ be a family of measurable spaces, and let $\mathcal{X}_\alpha$ be generated by $\mathcal{F}_\alpha$, $\alpha \in A$. Then $\prod_{\alpha\in A} \mathcal{X}_\alpha$ is generated by $\mathcal{X}_1 = \left\{\pi_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathcal{F}_\alpha, \alpha \in A\right\}$. If A is at most countable and $X_\alpha \in \mathcal{F}_\alpha$ for all α, $\prod_{\alpha\in A} \mathcal{X}_\alpha$ is generated by $\mathcal{X}_2 = \left\{\prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{F}_\alpha\right\}$.*

Proof. Obviously $\mathcal{B}[\mathcal{X}_1] \subset \prod_{\alpha\in A} \mathcal{X}_\alpha$. On the other hand, for each α, the collection $\{E \subset X_\alpha : \pi_\alpha^{-1}(E) \subset \mathcal{B}[\mathcal{X}_1]\}$ is easily seen to be a σ-algebra on X_α that contains $\mathcal{F}_\alpha$ and hence $\mathcal{X}_\alpha$. In other words, $\pi_\alpha^{-1}(E) \in \mathcal{B}[\mathcal{X}_1]$ for all $E \in \mathcal{X}_\alpha$, $\alpha \in A$, and hence $\prod_{\alpha\in A} \mathcal{X}_\alpha \subset \mathcal{B}[\mathcal{X}_1]$.

Now suppose that A is at most countable. Note that if $E_\alpha \in \mathcal{X}_\alpha$, then $\pi_\alpha^{-1}(E_\alpha) = \prod_{\beta\in A} E_\beta$ where $E_\beta = X$ for $\beta \neq \alpha$. Hence $\prod_{\alpha\in A} \mathcal{X}_\alpha \subset \mathcal{B}[\mathcal{X}_2]$. Also note that $\prod_{\alpha\in A} E_\alpha = \bigcap_{\alpha\in A} \pi_\alpha^{-1}(E_\alpha) \in \prod_{\alpha\in A} \mathcal{X}_\alpha$, so $\mathcal{B}[\mathcal{X}_2] \subset \prod_{\alpha\in A} \mathcal{X}_\alpha$. □

Proposition 6.30. *Let $X_1, \cdots, X_n$ be metric spaces and let $X = \prod_{j=1}^n X_j$, equipped with the product metric. Then $\prod_{j=1}^n \mathcal{B}_{X_j} \subset \mathcal{B}_X$. If the X_j's are separable, then $\prod_{j=1}^n \mathcal{B}_{X_j} = \mathcal{B}_X$.*

Proof. By proposition 6.29, $\prod_{j=1}^n \mathcal{B}_{X_j}$ is generated by the sets $\pi_j^{-1}(U_j)$, $1 \le j \le n$, where U_j is open in X_j. Since these sets are open in X, we get $\prod_{j=1}^n \mathcal{B}_{X_j} \subset \mathcal{B}_X$.

Suppose now that C_j is a countable dense set in X_j. Let $\mathcal{F}_j$ be the collection of balls in X_j with rational radius and center in C_j. Then it's clear that $\mathcal{B}_{X_j}$ is generated by $\mathcal{F}_j$ and $\mathcal{B}_X$ is generated by $\left\{\prod_{j=1}^n E_j : E_j \in \mathcal{F}_j\right\}$. Hence $\prod_{j=1}^n \mathcal{B}_{X_j} = \mathcal{B}_X$ by proposition 6.29. □

Corollary 6.31. $\mathcal{B}_{\mathbb{R}^n} = \prod_{j=1}^n \mathcal{B}_{\mathbb{R}}$.

Remark 6.32. More generally, let $(X_\alpha)_{\alpha\in A}$ be an at most countable family of second countable topological spaces. Then the Borel σ-algebra of the product space (with the product topology) is equal to the product of the Borel σ-algebras of the factor spaces.

Remark 6.33. The Lebesgue σ-algebra on $\mathbb{R}^n$ is not the product of n copies of the one-dimensional Lebesgue σ-algebra, as it contains some additional null sets; however, it is the completion of that product. See theorem 8.19.

Proposition 6.34. *Let $(X, \mathcal{X})$ and $(Y_\alpha, \mathcal{Y}_\alpha)$ ($\alpha \in A$) be measurable spaces. Then $f : (X, \mathcal{X}) \to (\prod_{\alpha\in A} Y_\alpha, \prod_{\alpha\in A} \mathcal{Y}_\alpha)$ if measurable iff $f_\alpha = \pi_\alpha \circ f$ is measurable for all α.*

Proof. If f is measurable, so is each f_α since the composition of measurable maps is measurable. Conversely, if each f_α is measurable, then for all $E_\alpha \in \mathcal{Y}_\alpha$, $f^{-1}(\pi_\alpha^{-1}(E_\alpha)) = f_\alpha^{-1}(E_\alpha) \in \mathcal{X}$. Hence f is measurable via proposition 6.11 and proposition 6.29. $\square$

Corollary 6.35. *Let $(X, \mathcal{X})$ be a measurable space. A function $f : X \to \mathbb{C}$ is measurable iff $\Re f$ and $\Im f$ are measurable.*

Proof. This follows since $\mathcal{B}_\mathbb{C} = \mathcal{B}_\mathbb{R} \times \mathcal{B}_\mathbb{R}$ by proposition 6.34. $\square$

6.E. **Measurability of sections.** In the next we introduce some basic properties of the sections.

Definition 6.36 (Sections)**.** *Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be measurable spaces. If $E \subset X \times Y$, for $x \in X$ and $y \in Y$, we define the x-**section** E_x and the y-**section** E^y of E by*

$$E_x = \{y \in Y : (x, y) \in E\}, \quad \textit{and} \quad E^y = \{x \in X : (x, y) \in E\}.$$

*Also, if f is a map on $X \times Y$ we define the x-**section** f_x and the y-**section** f^y of f by*

$$f_x(y) = f^y(x) = f(x, y).$$

Proposition 6.37. *Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be measurable spaces. If E is measurable with respect to $\mathcal{X} \times \mathcal{Y}$, then the section E_x is measurable in $\mathcal{Y}$ for every $x \in X$, and similarly the section E^y is measurable in $\mathcal{X}$ for every $y \in Y$.*

Proof. We only prove that E_x is measurable in $\mathcal{Y}$ for every $x \in X$, and the other is similar. Given $x \in X$, we define

$$\lambda_x : Y \to X \times Y, \quad y \mapsto (x, y).$$

Via proposition 6.29, $\mathcal{X} \times \mathcal{Y}$ is generated by $\{A \times B : A \in \mathcal{X}, B \in \mathcal{Y}\}$. Note that

$$\forall A \in \mathcal{X}, \forall B \in \mathcal{Y} : \lambda_x^{-1}(A \times B) = \begin{cases} B, & x \in A; \\ \varnothing, & x \notin A. \end{cases}$$

It follows that λ_x is measurable, and hence $E_x = \lambda_x^{-1}(E)$ is measurable. $\square$

Corollary 6.38. *Sections of Borel-measurable sets are again Borel-measurable.*

Remark 6.39. Sections of Lebesgue-measurable sets are not necessarily Lebesgue-measurable.

Corollary 6.40. *Let f be measurable map on $(X \times Y, \mathcal{X} \times \mathcal{Y})$.T Then f_x is $\mathcal{Y}$-measurable for all $x \in X$, and f^y is $\mathcal{X}$-measurable for all $y \in Y$.*

Proof. Note that $(f_x)^{-1}(E) = (f^{-1}(E))_x$ and $(f^y)^{-1}(E) = (f^{-1}(E))^y$. Then the desired result follows from proposition 6.37. □

7. Measures

Now we endow measurable spaces with a measure, turning them into measure spaces. We will introduce the basic theory of measures in this section.

7.A. **Basic knowledge.**

Definition 7.1 (Measures). *A (non-negative)* ***measure*** *μ on a measurable space $(X, \mathcal{X})$ is a function $\mu : \mathcal{X} \to [0, +\infty]$ such that*

(1) *$\mu(\varnothing) = 0$;*
(2) (Countable additivity) $\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n)$ *whenever $(E_n)_{n=1}^{\infty}$ is a sequence of disjoint measurable sets.*

We refer to the triplet $(X, \mathcal{X}, \mu)$ as a ***measure space****.*

Definition 7.2. *A measure space $(X, \mathcal{X}, \mu)$ is* ***finite*** *if $\mu(X) < \infty$; it is* ***probability space*** *if $\mu(X) = 1$ (and then we call μ a* ***probability measure****). It is σ-****finite*** *if X can be covered by countably many subsets of finite measure.*

A measurable set E is a ***null set*** *if $\mu(E) = 0$. A property on points x in X is said to hold for* ***almost every*** *$x \in X$ (or* ***almost surely****, for probability spaces) if it holds outside of a null set. We abbreviate almost every and almost surely as a.e. and a.s. respectively. The complement of a null set is said to be a* ***co-null set*** *or to have* ***full measure****.*

A measure space is said to be ***complete*** *if every subset of a null set is measurable (and then we call μ a* ***complete*** *measure).*

If X is a topological space, then measures on $\mathcal{B}_X$ is called ***Borel measures****.*

Example 7.3. Let $(X, \mathcal{X})$ be a measurable space.

(1) (**Dirac measures**) Given a point $x \in X$, we define the Dirac measure δ_x to be the measure such that $\delta_x(E) = 1$ when $x \in E$ and $\delta_x(E) = 0$ otherwise. This is a probability measure.
(2) (**Counting measures**) We define the counting measure # by defining $\#(E)$ to be the cardinality $|E|$ of E when E is finite, or $+\infty$ otherwise.
(3) Any finite non-negative linear combination of measures is again a measure; any finite covex combination of probability measures is again a probability measure.

Example 7.4 (Push-forward). If $f : X \to Y$ is a measurable map from one measurable space $(X, \mathcal{X})$ to another $(Y, \mathcal{Y})$, and μ is a measure on $\mathcal{X}$, we define the **push-forward** $f_*\mu : \mathcal{Y} \to [0, +\infty]$ by the formula $f_*\mu(E) := \mu(f^{-1}(E))$; this is a measure on $(Y, \mathcal{Y})$.

Proposition 7.5. *Let $(X, \mathcal{X}, \mu)$ be a measure space. Then*

*(1) (****Monotonicity****) If $E \subset F$ are measurable sets, then $\mu(E) \le \mu(F)$.*
*(2) (****Countable subadditivity****) If $E_1, E_2, \cdots$ are a countable sequence of measurable sets, then $\mu(\bigcup_{n=1}^{\infty} E_n) \le \sum_{n=1}^{\infty} \mu(E_n)$. (Of course, one also has subadditivity for finite sequence.)*

*(3) (**Monotone convergence for sets**) If $E_1 \subset E_2 \subset \cdots$ are measurable, then $\mu(\bigcup_{n=1}^{\infty} E_n) = \lim_{n\to\infty} \mu(E_n)$.*
*(4) (**Dominated convergence for sets**) If $E_1 \supset E_2 \supset \cdots$ are measurable, and $\mu(E_1)$ is finite, then $\mu(\bigcap_{n=1}^{\infty} E_n) = \lim_{n\to\infty} \mu(E_n)$.*

Proof. Trivial. $\square$

Remark 7.6. (4) can fail if $\mu(E_1)$ is infinite. Just consider $E_n = (n, +\infty)$.

7.B. **Approximations — Egoroff's theorem, convergence in measure.** In the next we make a little more approximations of measurable maps via the measure. (We not only use pointwise limit as before.)

Theorem 7.7 (Egoroff's theorem). *Let $(X, \mathcal{X}, \mu)$ be a measure space with $\mu(X) < \infty$, let Y be a metric space, and let $f_1, f_2, \cdots$ and f be measurable functions from $(X, \mathcal{X})$ to $(Y, \mathcal{B}_Y)$ such that $f_n \to f$ $\mathcal{X}$-almost everywhere. Then for every $\varepsilon > 0$ there exists $E \subset X$ such that $\mu(E) < \varepsilon$ and $f_n \to f$ uniformly on E^c.*

Proof. For $j, k = 1, 2, \cdots$ define

$$C_{jk} := \bigcup_{n=k}^{\infty} \{x \in X : |f_n(x) - f(x)| > 2^{-j}\}$$

Then $C_{j,k+1} \subset C_{jk}$ for all j, k; and so, since $\mu(X) < \infty$,

$$\lim_{k\to\infty} \mu(C_{jk}) = \mu\left(\bigcap_{k=1}^{\infty} C_{jk}\right) = 0,$$

and hence there exists an integer $N(j)$ such that $\mu(C_{j,N(j)}) < \varepsilon 2^{-j}$. Putting $E := \bigcup_{j=1}^{\infty} C_{j,N(j)}$, then we have

$$\mu(E) \le \sum_{j=1}^{\infty} \mu(C_{j,N(j)}) < \varepsilon,$$

and for each j, each $x \in E^c$, and all $n \ge N(i)$, we have $|f_n(x) - f(x)| \le 2^{-j}$. Consequently $f_n \to f$ uniformly on E^c. $\square$

Now, via the measure, we introduce a new mode of convergence.

Definition 7.8 (Convergence in measure). *A sequence (f_n) of measurable complex-valued functions on a measure space $(X, \mathcal{X}, \mu)$ is called **Cauchy in measure** if for every $\varepsilon > 0$,*

$$\mu\left(\{x : |f_n(x) - f_m(x)| \ge \varepsilon\}\right) \to 0 \text{ as } m, n \to \infty,$$

*and that (f_n) **converges in measure** to f if for every $\varepsilon > 0$,*

$$\mu\left(\{x : |f_n(x) - f(x)| \ge \varepsilon\}\right) \to 0 \text{ as } n \to \infty.$$

Remark 7.9. Some people also talk about local convergence in measure, and the corresponding topology of (local) convergence in measure.

Theorem 7.10. *Let $(X, \mathcal{X}, \mu)$ be a measure space. Suppose that (f_n) is Cauchy in measure. Then there is a measurable function f such that $f_n \to f$ in measure, and there is a* *subsequence (f_{n_j}) that converges to f a.e. Moreover, if also $f_n \to g$ in measure, then $g = f$ a.e.*

Proof. We can choose a subsequence $(g_j) = (f_{n_j})$ of (f_n) such that $\mu(E_j) \le 2^{-j}$ where $E_j := \{x \in X : |g_j(x) - g_{j+1}(x)| \ge 2^{-j}\}$. Set $F_k = \bigcup_{j=k}^{\infty} E_j$ and $F = \bigcap_{k=1}^{\infty} F_k$. It follows that $\mu(F_k) \le \sum_{j\ge k} 2^{-j} = 2^{1-k}$, and hence $\mu(F) = 0$. Note that for $x \in F_k^c$ and for $h \ge j \ge k$ we have

$$|g_h(x) - g_j(x)| \le \sum_{l=j}^{h-1} |g_{l+1}(x) - g_l(x)| \le \sum_{l=j}^{h-1} 2^{-l} \le 2^{1-j}. \tag{7.1}$$

It follows that (g_j) is pointwise Cauchy on F^c.

Set $f(x) = \lim_{j\to\infty} g_j(x)$ for $x \in F^c$ and $f(x) = 0$ for $x \in F$. It easily follows from proposition 6.14 that f is measurable, and $g_j \to f$ a.e. It follows from (7.1) that $|g_j(x) - f(x)| \le 2^{1-j}$ for $x \in F_k^c$. Since $\mu(F_k) \to 0$ as $k \to 0$, $g_j \to f$ in measure. Note that

$$\{x : |f_n(x) - f(x)| \ge \varepsilon\} \subset \left\{x : |f_n(x) - g_j(x)| \ge \frac{1}{2}\varepsilon\right\} \bigcup \left\{x : |g_j(x) - f(x)| \ge \frac{1}{2}\varepsilon\right\}.$$

It follows that $f_n \to f$ in measure. Likewise, if $f_n \to g$ in measure, note that

$$\{x : |f(x) - g(x)| \ge \varepsilon\} \subset \left\{x : |f(x) - f_n(x)| \ge \frac{1}{2}\varepsilon\right\} \bigcup \left\{x : |f_n(x) - g(x)| \ge \frac{1}{2}\varepsilon\right\}.$$

Hence $\mu(\{x : |f(x) - g(x)| \ge \varepsilon\}) = 0$ for all $\varepsilon > 0$. It follows then $f = g$ a.e. □

Remark 7.11. One can refer to [For] for more relations among different modes of convergence, such as the convergence in L^1.

7.C. **Completion, completele spaces.** In the next we introduce the completion of a measure and the complete measure spaces.

Theorem 7.12 (Completion). *Let $(X, \mathcal{X}, \mu)$ be a measure space. Let $\mathcal{N} = \{N \in \mathcal{X} : \mu(N) = 0\}$ and $\overline{\mathcal{X}} = \{E \cup F : E \in \mathcal{X}$ and $F \subset N$ for some $N \in \mathcal{N}\}$. Then $\overline{\mathcal{X}}$ is a σ-algebra, and there is a unique extension μ of μ to a complete measure on $\overline{\mathcal{X}}$.*

Proof. To show that $\overline{\mathcal{X}}$ is a σ-algebra, it suffices to prove that $\overline{\mathcal{X}}$ is closed under countable unions and complements. Since $\mathcal{X}$ and $\mathcal{N}$ are closed under countable unions, so is $\overline{\mathcal{X}}$. If $E \cup F \in \overline{\mathcal{X}}$ where $E \in \mathcal{X}$ and $F \subset N \in \mathcal{N}$, we can assume that $E \cap N = \varnothing$ (otherwise, replace F and N by $F \setminus E$ and $N \setminus E$). Then $(E \cup F)^c = (E \cup N)^c \cup (N \setminus F)$ where $(E \cup N)^c \in \mathcal{X}$ and $N \setminus F \subset N \in \mathcal{N}$. Hence $(E \cup F)^c \in \overline{\mathcal{X}}$.

If $E \cup F \in \overline{\mathcal{X}}$ as above, we set $\overline{\mu}(E \cup F) = \mu(E)$. It's obvious that $\overline{\mu}$ is well-defined and is the only measure on $\overline{\mathcal{X}}$ that extends μ. □

Remark 7.13. The completion is the unique minimal complete refinement. See the next proposition. In particular, the completion of the Borel σ-algebra with respect to Lebesgue measure is known as the Lebesgue σ-algebra.

Proposition 7.14. *Let $(X, \mathcal{X}, \mu)$ and $(X, \mathcal{X}_1, \mu_1)$ be measure spaces satisfying $\mathcal{X} \subset \mathcal{X}_1 \subset \overline{\mathcal{X}}$ and $\mu_1 = \overline{\mu}|_{\mathcal{X}_1}$. Then the completion of $(X, \mathcal{X}_1, \mu_1)$ is still $(X, \overline{\mathcal{X}}, \overline{\mu})$, and $(X, \mathcal{X}_1, \mu_1)$ is complete iff $(X, \mathcal{X}_1, \mu_1) = (X, \overline{\mathcal{X}}, \overline{\mu})$.*

Proof. Trivial. □

Proposition 7.15. *Let $(X, \mathcal{X}, \mu)$ be a complete measure space, and let $f, g, f_n : X \to \mathbb{R}$ (or $\mathbb{C}$).*

(1) If f is measurable and $f = g$ a.e., then g is measurable.
(2) If f_n is measurable for $n \in \mathbb{N}$ and $f_n \to f$ a.e., then f is measurable.

Proof. Trivial. □

Proposition 7.16. *Let $(X, \mathcal{X}, \mu)$ be a measure space and let $(X, \overline{\mathcal{X}}, \overline{\mu})$ be its completion. If f is an $\overline{\mathcal{X}}$-measurable function on X, there is an $\mathcal{X}$-measurable function g such that $f = g$ $\overline{\mu}$-a.e.*

Proof. If $f = \chi_E$ where $E \in \overline{\mathcal{X}}$, then the conclusion just follows from the definition of $\overline{\mu}$. For the general case, choose a sequence (φ_n) of $\overline{\mathcal{X}}$-measurable simple functions that converges pointwise to f by theorem 6.22. For each n let ψ_n be an $\mathcal{X}$-measurable simple function with $\psi_n = \varphi_n$ except on a set $E_n \subset \overline{\mathcal{X}}$ with $\overline{\mu}(E_n) = 0$. Choose $N \in \mathcal{X}$ such that $\mu(N) = 0$ and $\bigcup_{n=1}^{\infty} E_n \subset N$, and set $g = \lim_{n\to\infty} \psi_n \chi_{X\setminus N}$. Then g is $\mathcal{X}$-measurable via proposition 6.14 and $g = f$ on N^c. □

7.D. **Outer measure, elementary sets.** In the next we introduce the concept of outer measure, which can be regarded as the predecessor of measure in the following sense:

Take $\mathbb{R}^2$ for example. One draws a grid of rectangles in the plane and approximates the area of E from below by the sum of the areas of the rectangles in the grid that are subsets of E, and from above by the sum of the areas of the rectangles in the grid that intersect E. The limits of these approximations as the grid is taken finer and finer give the "inner area" and "outer area" of E, and if they are equal, their common value is the "area" of E.

Definition 7.17 (Outer measure)**.** *An **outer measure** on a nonempty set X is a function $\mu^* : \mathcal{P}(X) \to [0, +\infty]$ that satisfies*

(1) $\mu^*(\varnothing) = 0$.
(2) (Monotonicity) $\mu^*(A) \le \mu^*(B)$ if $A \subset B$.
(3) (Countable subadditivity) $\mu^*(\bigcup_{n=1}^{\infty} A_n) \le \sum_{n=1}^{\infty} \mu^*(A_n)$.

The most common way to obtain outer measures is to start with a family $\mathcal{E}$ of "**elementary sets**" on which a notion of measure is defined (such as rectangles in the plane) and then to approximate arbitrary sets "from the outside" by countable unions of members of $\mathcal{E}$.

Proposition 7.18. *Given $\mathcal{E} \subset \mathcal{P}(X)$ and $\rho : \mathcal{E} \to [0, +\infty]$ that satisfies*

(1) $\varnothing \in \mathcal{E}$ and $\rho(\varnothing) = 0$;
(2) There exists $(E_n)_{n=1}^{\infty} \subset \mathcal{E}$ with $X = \bigcup_{n=1}^{\infty} E_n$.

For any $A \subset X$, define

$$\mu^*(A) = \inf\left\{\sum_{n=1}^{\infty} \rho(E_n) : E_n \in \mathcal{E} \text{ and } A \subset \bigcup_{n=1}^{\infty} E_n\right\}.$$

Then μ^* is an outer measure.

Proof. Obviously, $\mu^* : \mathcal{P}(X) \to [0, +\infty]$ makes sense, $\mu^*(\emptyset) = 0$, and $\mu^*(A) \leq \mu^*(B)$ if $A \subset B$. To prove the countable subadditivity, suppose that $(A_n)_{n=1}^\infty \subset \mathcal{P}(X)$ and $\varepsilon > 0$. For each n there exists $(E_n^k)_{k=1}^\infty \subset \mathcal{E}$ such that $A \subset \bigcup_{k=1}^\infty E_n^k$ and that $\sum_{k=1}^\infty \rho(E_n^k) \leq \mu^*(A_n) + \varepsilon 2^{-n}$. Since $\bigcup_{n=1}^\infty A_n \subset \bigcup_{n,k=1}^\infty E_n^k$, we have $\mu^*(A) \leq \sum_{n=1}^\infty \mu^*(A_n) + \varepsilon$. Since ε is arbitrary, we are done. □

7.E. **Correspondence between outer measures and measures.** The fundamental step that leads from outer measures to measures is as follows.

Theorem 7.19 (Carathéodory's theorem)**.** *If μ^* is an outer measure on X, the collection*

$$\mathcal{X} = \{A \subset X : \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c), \forall E \subset X\}$$

of μ^*-measurable sets is a σ-algebra, and the restriction of μ^* to $\mathcal{X}$ is a complete measure.

Remark 7.20. Some motivations for the notion of μ^*-measurability can be obtained by referring to the discussion above for $\mathbb{R}^2$. If E is "well-behaved" set such that $A \subset E$, then the derived equation $\mu^*(A) = \mu^*(E \cap A) = \mu^*(E) - \mu^*(E \cap A^c)$ says that the outer measure of A, $\mu^*(A)$, is equal to the "inner measure" of A, $\mu^*(E) - \mu^*(E \cap A^c)$.

Proof. Obviously $\mathcal{X}$ is closed under complements. By subadditivity it's clear that $\mathcal{X}$ is closed under finite intersections and unions and is finite additive. To show that $\mathcal{X}$ is a σ-algebra it suffices to show that $\mathcal{X}$ is closed under countable disjoint unions. If $(A_n)_{n=1}^\infty$ is a sequence of disjoint sets in $\mathcal{X}$, let $B_k = \bigcup_{n=1}^k A_n$ and $B = \bigcup_{n=1}^\infty A_n$. Then for $E \subset X$,

$$\begin{aligned} \mu^*(E \cap B_k) &= \mu^*(E \cap B_k \cap A_k) + \mu^*(E \cap B_k \cap A_k^c) \\ &= \mu^*(E \cap A_k) + \mu^*(E \cap B_{k-1}) \end{aligned}$$

and hence $\mu^*(E \cap B_k) = \sum_{n=1}^k \mu^*(E \cap A_n)$. Therefore,

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap B_k) + \mu^*(E \cap B_k^c) \\ &\geq \sum_{n=1}^k \mu^*(E \cap A_n) + \mu^*(E \cap B^c) \geq \sum_{n=1}^\infty \mu^*(E \cap A_n) + \mu^*(E \cap B^c) \\ &\geq \mu^*\left(\bigcup_{n=1}^\infty (E \cap A_n)\right) + \mu^*(E \cap B^c) \\ &= \mu^*(E \cap B) + \mu^*(E \cap B^c) \geq \mu^*(E) \end{aligned}$$

Hence $B \in \mathcal{X}$. Moreover, taking $E = B$ we get $\mu^*(B) = \sum_{n=1}^\infty \mu^*(A_n)$, so μ^* is countably additive on $\mathcal{X}$. Finally, if $\mu^*(A) = 0$, for any $E \subset X$ we have

$$\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c) = \mu^*(E \cap A^c) \leq \mu^*(E)$$

so that $A \in \mathcal{X}$. Therefore $\mu^*|_{\mathcal{X}}$ is a complete measure. □

Remark 7.21. Let $(X, \mathcal{X})$ be a measurable space. Via proposition 7.18 we see that any measure μ on $\mathcal{X}$ will also induce an outer measure μ^*, and we will see in proposition 7.24 that $\mu^*|_{\mathcal{X}} = \mu$. So some people regard measures and outer measures as the same things.

7.F. **Premeasure on algebra, from premeasure to outer measure.** More precisely, we can construct an outer measure from a premeasure on a algebra.

Definition 7.22 (Algebra). *An **algebra** of sets on X is a nonempty collection $\mathcal{A}$ of subsets of X that is closed under finite unions and complements (and hence is closed under finite intersections and contains $\varnothing$ and X).*

Definition 7.23 (Premeasure). *If $A \subset \mathcal{P}(X)$ is an algebra, a function $\mu_0 : \mathcal{A} \to [0, +\infty]$ is called a **premeasure** if*

(1) $\mu_0(\varnothing) = 0$;
(2) (Countable additivity) If $(A_n)_{n=1}^\infty$ is a sequence of disjoint sets in $\mathcal{A}$ such that $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$, then $\mu_0(\bigcup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mu_0(A_n)$.

Proposition 7.24 (Premeasure induces outer measures). *Let $A \subset \mathcal{P}(X)$ be an algebra, and let μ_0 be a premeasure on $\mathcal{A}$. Define*

$$\mu^* : \mathcal{P}(X) \to [0, +\infty], \quad E \mapsto \inf \left\{ \sum_{n=1}^\infty \mu_0(A_n) : A_n \in \mathcal{A}, E \subset \bigcup_{n=1}^\infty A_N \right\}.$$

Then

(1) $\mu^*|_\mathcal{A} = \mu_0$;
(2) Every set in $\mathcal{A}$ is μ^*-measurable.

Proof. Let $A \in \mathcal{A}$. If $A \subset \bigcup_{n=1}^\infty A_n$ with $A_n \in \mathcal{A}$, setting $B_k = A \cap (A_k \setminus \bigcup_{n=1}^{k-1} A_n)$ then the B_k's are disjoint members of $\mathcal{A}$ whose union is A, so $\mu_0(A) = \sum_{k=1}^\infty \mu(B_k) \le \sum_{k=1}^\infty \mu_0(A_k)$. It follows that $\mu_0(A) \le \mu^*(A)$, and the reverse inequality is obvious.

If $A \in \mathcal{A}$, $E \subset X$ and $\varepsilon > 0$, there is a sequence $(B_n)_{n=1}^\infty \subset \mathcal{A}$ with $E \subset \bigcup_{n=1}^\infty B_n$ and $\sum_{n=1}^\infty \mu_0(B_n) \le \mu^*(E) + \varepsilon$. Since μ_0 is additive on $\mathcal{A}$, we have

$$\mu^*(E) + \varepsilon \ge \sum_{n=1}^\infty \mu_0(B_n \cap A) + \sum_{n=1}^\infty \mu_0(B_n \cap A^c) \ge \mu^*(E \cap A) + \mu^*(E \cap A^c) \ge \mu^*(E).$$

Since ε is arbitrary, A is μ^*-measurable. ☐

7.G. **From premeasure to measure — extension comparison.** Since outer measure induces a measure, we get a extension for a premeasure. Moreover, we can compare the extensions, and in many cases the extension is unique.

Theorem 7.25 (Extension comparison). *Let $A \subset \mathcal{P}(X)$ be an algebra, let μ_0 be a premeasure on $\mathcal{A}$, and let $\mathcal{X}$ be the σ-algebra generated by $\mathcal{A}$. It follows that the outer measure μ^* given by proposition 7.24 extends μ_0 to a measure μ on $\mathcal{X}$. If ν is another measure on $\mathcal{X}$ that extends μ_0, then $\nu(E) \le \mu(E)$ for all $E \in \mathcal{X}$, with equality when $\mu(E) < \infty$. If μ_0 is σ-finite, then μ is the unique extension of μ_0 to a measure on $\mathcal{X}$.*

Proof. The first assertion follows from Carathéodory's theorem 7.19 and proposition 7.24 since the σ-algebra of μ^*-measurable sets includes $\mathcal{A}$ and hence $\mathcal{X}$.

As for the second assertion, note that if $E \in \mathcal{X}$ and $E \subset \bigcup_{n=1}^\infty A_n$ where $A_n \in \mathcal{A}$, then $\nu(E) \le \sum_{n=1}^\infty \nu(A_n) = \sum_{n=1}^\infty \mu_0(A_n)$. It follows that $\nu(E) \le \mu(E)$. Also, if we set $A = \bigcup_{n=1}^\infty A_n$, we have

$$\nu(A) = \lim_{k\to\infty} \nu\left(\bigcup_{n=1}^\infty A_n\right) = \lim_{k\to\infty} \mu\left(\bigcup_{n=1}^\infty A_n\right) = \mu(A).$$

If $\mu(E) < \infty$, we can choose the A_n's so that $\mu(A) < \mu(E) + \varepsilon$, and hence

$$\mu(E) \le \mu(A) = \nu(A) = \nu(E) + \nu(A \setminus E) \le \nu(E) + \mu(A \setminus E) \le \nu(E) + \varepsilon$$

Since ε is arbitrary, $\mu(E) = \nu(E)$.

Finally, suppose that $X = \bigcup_{n=1}^{\infty} A_n$ with $\mu_0(A_n) < \infty$, where we can assume that the A_n's are disjoint. Then for any $E \in \mathcal{X}$,

$$\mu(E) = \sum_{n=1}^{\infty} \mu(E \cap A_n) = \sum_{n=1}^{\infty} \nu(E \cap A_n) = \nu(E),$$

so $\mu = \nu$. □

Remark 7.26. The same statements of course apply for the completion $\overline{\mathcal{X}}$ of $\mathcal{X}$.

Remark 7.27. The uniqueness of extension in theorem 7.25 can convince us that the constructed measure via the above methods is unique.

7.H. **From elementary family to algebra, product measure.** In the next we introduce a basic method to generate an algebra.

Definition 7.28 (Elementary family)**.** *An elementary family $\mathcal{G}$ is a collection of subsets of X satisfying the following properties.*

(1) $\varnothing \in \mathcal{G}$;
(2) If $E, F \in \mathcal{G}$, then $E \cap F \in \mathcal{G}$;
(3) If $E \in \mathcal{G}$, then E^c is a finite disjoint union of elements in $\mathcal{G}$.

Theorem 7.29. *The collection $\mathcal{A}$ of finite disjoint unions of elements in an elementary family $\mathcal{G}$ forms an algebra.*

Proof. One can refer to [Rai]. □

Remark 7.30. Two premeasure μ and ν on an algebra $\mathcal{A}$ generated by an elementary family $\mathcal{G}$ coincide iff they coincide on $\mathcal{G}$.

Now we give a direct application, constructing the product measure.

Corollary 7.31 (Product measure)**.** *Let $(X_1, \mathcal{X}_1, \mu_1), \cdots, (X_n, \mathcal{X}_n, \mu_n)$ be measure spaces, and let $(\prod_{j=1}^n X_j, \prod_{j=1}^n \mathcal{X}_j)$ be the product measure space. Then*

*(1) (**Rectangle algebra**) We define a **rectangle** to be a set of the form $\prod_{j=1}^n A_j$ with $A_j \in \mathcal{X}_j$, and then the collection $\mathcal{G}$ of rectangles forms an elementary family, and hence the collection $\mathcal{A}$ of finite disjoint unions of rectangles is an algebra. Clearly, the σ-algebra generated by $\mathcal{A}$ is $\prod_{j=1}^n \mathcal{X}_j$;*
*(2) (**Premeasure**) Then we define a premeasure μ_0 on $\mathcal{A}$ by $\mu_0\left(\prod_{j=1}^n A_j\right) = \prod_{j=1}^n \mu_0(A_j)$ and linear extension.*
*(3) (**Product measure**) Finally, we get a product measure μ on $\prod_{j=1}^n \mathcal{X}_j$ as in theorem 7.25. The measure is referred to as the product measure of the $\mu_1, \cdots, \mu_n$ and is denoted by $\prod_{j=1}^n \mu_j$.*

Moreover, if the μ_j's are σ-finite so that the extension from $\mathcal{A}$ to $\prod_{j=1}^n \mathcal{X}_j$ is uniquely determined via theorem 7.25.

7.I. **From algebra to σ-algebra.** Finally, we introduce a lemma to show how to generate a σ-algebra via an algebra.

Lemma 7.32 (The monotone class lemma)**.** *We define a **monotone class** on a space X to be a subset $\mathcal{D}$ of $\mathcal{P}(X)$ that is closed under countable incresing unions and countable decreasing intersections. Clearly, the intersection of any family of monotone classes is a monotone class, so for any $\mathcal{E} \in \mathcal{P}(X)$ there is a unique smallest monotone class containing $\mathcal{E}$, called the monotone class **generated by** $\mathcal{E}$.*

If $\mathcal{A}$ if an algebra of subsets of X, then the monotone class $\mathcal{D}$ generated by $\mathcal{A}$ coincides with the σ-algebra generated by $\mathcal{A}$.

Proof. One can refer to [For] lemma 2.35. □

8. Lebesgue measure

8.A. **Construction of Lebesgue measure.** There are three methods to construct the Lebesgue measure on $\mathbb{R}^n$:

(1) Construct both outer measure m^* and inner measure m_*, and then we define

$$\begin{aligned}\mathcal{L}_0 &= \{A \subset \mathbb{R}^n : m^*(A) < \infty, m^*(A) = m_*(A)\}\\ \mathcal{L} &= \{A \subset \mathbb{R}^n : A \cap M \in \mathcal{L}_0, \forall M \in \mathcal{L}_0\}\end{aligned}$$

Then $\mathcal{L}$ is the collection of all Lebesgue-Measurable sets, and the Lebesgue measure can be defined as $m(A) = \sup\{m(A \cap M) : M \in \mathcal{L}_0\}$. One can refer to [Jon] for this method.

(2) Consturct the outer measure m^*, and then m^* induces the Lebesgue measure m. We will introduce this method in this subsection.

(3) Use the Riesz representation theorem for Radon measures, which will be introduced later.

In the next we introduce method 2. As in proposition 7.18, we start with a family of "elementary sets" $\mathcal{E}$ to define the outer measure. $\mathcal{E}$ can be the set of all cubes or be the set of all balls (see theorem 11.31). Here we choose cubes.

8.B. **Elementary sets — boxes and cubes, Lebesgue (outer) measure.**

Definition 8.1 (Box and cube)**.** *A **box** I in $\mathbb{R}^n$ is given by the product of n compact intervals $I = [a,b] := \prod_{j=1}^n [a_j, b_j]$ where $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$, and $a_j \le b_j$, $1 \le j \le n$, are real numbers. The **volume** $|I|$ of I is defined by $|I| = \prod_{j=1}^n |b_j - a_j|$. A box is called a **cube** if all its sides have the same length. the **interior** of a box I is given by $I^\circ = (a,b) := \prod_{j=1}^n (a_j, b_j)$. A union of boxes is said to be **almost disjoint** if the interiors of the boxes are disjoint. We denote by $\mathrm{dist}(E_1, E_2) = \inf\{|x_1 - x_2| : x_1 \in E_1, x_2 \in E_2\}$ the **distance** of two subsets $E_1, E_2 \subset \mathbb{R}^n$.*

Theorem 8.2 (Lebesgue measure)**.** *Let $m^* : \mathcal{P}(\mathbb{R}^n) \to [0, +\infty]$ be defined By*

$$m^*(E) := \inf\left\{\sum_{n=1}^{\infty} |Q_n| : (Q_n)_{n=1}^{\infty} \text{ is a countable cover of } E \text{ by cubes}\right\}$$

and set

$$\mathcal{L}(\mathbb{R}^n) := \{E \subset \mathbb{R}^n : m^*(A) = m^*(A \cap E) + m^*(A \cap E^c), \forall A \subset \mathbb{R}^n\}$$

Then

*(1) m^* is an outer measure; the so-called **Lebesgue outer measure**.*
(2) If $\mathrm{dist}(E_1, E_2) > 0$, then $m^*(E_1 \cup E_2) = m^*(E_1) + m^*(E_2)$.
(3) $\mathcal{L}(\mathbb{R}^n)$ is a σ-algebra that contains $\mathcal{B}_{\mathbb{R}^n}$.

Moreover, $m = m^*|_{\mathcal{L}(\mathbb{R}^n)} : \mathcal{L}(\mathbb{R}^n) \to [0, +\infty]$ is a complete measure.

Proof. (1) It follows from proposition 7.18 that m^* is an outer measure.

(2) Choose $\mathrm{dist}(E_1, E_2) > \delta > 0$ and fix $\varepsilon > 0$. There exists a cover $(Q_n)_{n=1}^{\infty}$ by cubes of $E := E_1 \cup E_2$ so that $\sum_{n=1}^{\infty} |Q_n| \le m^*(E) + \varepsilon$. We may assume that each Q_n has diameter less than δ, after possibly subdividing Q_n. Then each Q_n can intersect

at most one of E_1 or E_2. Hence setting $J_k = \{n : Q_n \cap E_k \neq \varnothing\}$, $k = 1, 2$, we have $J_1 \cap J_2 = \varnothing$, and $E_k \subset \bigcup_{n \in J_k} Q_n$. Therefore,

$$m^*(E) \leq m^*(E_1) + m^*(E_2) \leq \sum_{n \in J_1} |Q_n| + \sum_{n \in J_2} |Q_n| \leq \sum_{n=1}^{\infty} |Q_n| \leq m^*(E) + \varepsilon$$

Since ε is arbitrary, we are done.

(3) It follows from Carathéodory's theorem 7.19 that $\mathcal{L}(\mathbb{R}^n)$ is a σ-algebra. In order to prove that $\mathcal{B}_{\mathbb{R}^n} \subset \mathcal{L}(\mathbb{R}^n)$ it suffices to prove that $\mathcal{L}(\mathbb{R}^n)$ contains all closed subsets of $\mathbb{R}^n$. Let $F \subset \mathbb{R}^n$ be closed, and let $A \subset \mathbb{R}^n$. It suffices to prove

$$m^*(A) \geq m^*(A \cap F) + m^*(A \cap F^c).$$

WLOG, we assume that $m^*(A) < \infty$. Set

$$A_0 := \{x \in A : \operatorname{dist}(x, F) \geq 1\}$$
$$A_j := \left\{x \in A : (j+1)^{-1} \leq \operatorname{dist}(x, F) < j^{-1}\right\}, \quad j \geq 1$$

Then by (2) we have that for each $n \in \mathbb{N}$,

$$\sum_{j=0}^{n} m^*(A_{2j}) \leq m^*\left(\bigcup_{j=0}^{n} A_{2j}\right) \leq m^*(A)$$
$$\sum_{j=0}^{n} m^*(A_{2j+1}) \leq m^*\left(\bigcup_{j=0}^{n} A_{2j+1}\right) \leq m^*(A)$$

and hence $\sum_{j=1}^{\infty} m^*(A) < \infty$. Therefore $\sum_{j=n+1}^{\infty} m^*(A_j) \to 0$ as $n \to \infty$ and

$$\begin{aligned} m^*(A \cap F) + m^*(A \cap F^c) &\leq m^*(A \cap F) + m^*\left(\bigcup_{j=0}^{n} A_j\right) + \sum_{j=n+1}^{\infty} m^*(A_j) \\ &= m^*\left((A \cap F) \cup \bigcup_{j=0}^{n} A_j\right) + \sum_{j=n+1}^{\infty} m^*(A_j) \\ &\leq m^*(A) + \sum_{j=n+1}^{\infty} m^*(A_j) \end{aligned}$$

Hence $m^*(A) \geq m^*(A \cap F) + m^*(A \cap F^c)$. Done.

Finally, it follows from Carathéodory theorem 7.19 that m is a complete measure. ☐

Remark 8.3. The completion of $\left(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, m^*|_{\mathcal{B}(\mathbb{R}^n)}\right)$ is $\left(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), m^*|_{\mathcal{L}(\mathbb{R}^n)}\right)$.

Remark 8.4. We can choose the "elementary sets" to be the **open cubes** (the interiors of cubes) to define the Lebesgue measure. Namely, putting

$$\lambda^*(E) := \inf\left\{\sum_{n=1}^{\infty} |P_n| : (P_n)_{n=1}^{\infty} \text{ is a countable cover of } E \text{ by open cubes}\right\},$$

then $\lambda^*(E) = m^*(E)$.

Proof. It's clear that $\lambda^*(E) \geq m^*(E)$. On the other hand, fix $\varepsilon > 0$, and then given a countable cover by cubes $(Q_n)_{n=1}^\infty$ of E with $\sum_{n=1}^\infty |Q_n| \leq m^*(E) + \varepsilon$, we can find a countable sequence of open cubes (P_n) with $Q_n \subset P_n$ and $|P_n| < |Q_n| + \varepsilon 2^{-n}$, and hence $m^*(E) + \varepsilon \geq \sum_{n=1}^\infty |Q_n| > \sum_{n=1}^\infty (|P_n| - \varepsilon 2^{-n}) = \sum_{n=1}^\infty |P_n| - \varepsilon$. Since ε is arbitrary, we get that $\lambda^*(E) \leq m^*(E)$. We are done. □

Remark 8.5. In this construction, we begin with the "elementary sets" $\mathcal{E}$ and then define m^* and m. In fact we can also begin with a algebra $\mathcal{A}$, and then define the premeasure m_0, which implies m^* and m.

More precisely, the collection $\mathcal{A}$ of all finite disjoint unions of sets of the form $F \cap G$, where F is closed and G is open, is an algebra. Although $\mathcal{A}$ is a little complicated and we don't construct the Lebesgue measure via it, but we can show the uniqueness of Lebesgue measure via this algebra and corresponding premeasure $m|_{\mathcal{A}}$ by theorem 7.25.

8.C. Examples — generalized Cantor sets, Cantor-Lebesgue function.

Example 8.6. There are some basic examples.

(1) One-point sets are null sets.
(2) The Cantor set C is a null set. (Moreover, $\operatorname{card}(C) = \operatorname{card}(\mathbb{R})$, and C is compact, nowhere dense, totally disconnected and has no isolated points.)
(3) (**The generalized Cantor set**) If $(\alpha_j)_{j=1}^\infty$ is any sequence of members in $(0,1)$, then, we can define a decreasing sequence $(K_j)_{j=0}^\infty$ of closed sets as follows: $K_0 = [0,1]$, and K_j is obtained by removing the open middle α_j-th from each of the intervals that make up K_{j-1}. The resulting limiting set $K = \bigcap_{j=1}^\infty K_j$ is called a **generalized Cantor set**. It's clear that $m(K) = \lim_{k\to\infty} m(K_j) = \prod_{j=1}^\infty (1-\alpha_j)$, which can achieve any number in $[0,1)$.

    On the other hand, to achieve any number $a \in (0,1)$ directly, we can remove the middle $\frac{1-a}{3^{n+1}}$ from each closed interval at stage n, thereby removing a total of
$$(1-a)\sum_{n=0}^\infty \frac{2^n}{3^{n+1}} = 1-a.$$
(4) For a cube $I = [a,b]$, we have $m([a,b]) = \big|[a,b]\big|$.
(5) Let $E = \bigcup_{j=1}^\infty Q_j$ be an almost disjoint union of cubes, then $m(E) = \sum_{j=1}^\infty |Q_j|$.

Remark 8.7. $\mathcal{L}(\mathbb{R}^n) \neq \mathcal{P}(R^n)$ if we admit the axiom of choice.

Example 8.8 (Cantor-Lebesgue function)**.** We construct the Cantor-Lebesgue function f in the following steps.

(1) Each $x \in [0,1]$ has a base-3 decimal expansion $x = \sum_{j=1}^\infty a_j 3^{-j}$ where $a_j = 0, 1$ or 2. This expansion is unique unless $x = p3^{-k}$ for some intergers p, k, in which case x has two expansions: one with $a_j = 0$ for $j > k$ and one with $a_j = 2$ for $j > k$. Assuming p is not divisible by 3, one of these expansions will have $a_k = 1$ and the other will have $a_0 = 0$ or 2. If we agree *always* to use the latter expansion, we see

that

$$a_1 = 1 \quad \text{iff} \quad \frac{1}{3} < x < \frac{2}{3}$$
$$a_1 \neq 1 \text{ and } a_2 = 1 \quad \text{iff} \quad \frac{1}{9} < x < \frac{2}{9} \text{ or } \frac{7}{9} < x < \frac{8}{9}$$

and so forth. Also note that if $x = \sum_{j=1}^{\infty} a_j 3^{-j}$ and $y = \sum_{j=1}^{\infty} b_j 3^{-j}$, then $x < y$ iff there exists an n such that $a_n < b_n$ and $a_j = b_j$ for $j < n$.

(2) The **Cantor set** C is the set of all $x \in [0,1]$ that have a base-3 expansion $x = \sum_{j=1}^{\infty} a_j 3^{-j}$ with $a_j \neq 1$ for all j. Thus C is obtained from $[0,1]$ by removing the open middle third $(\frac{1}{3}, \frac{2}{3})$, then removing the open middle thirds $(\frac{1}{9}, \frac{2}{9})$ and $(\frac{7}{9}, \frac{8}{9})$ of the two remaining intervals, and so forth.

(3) Define the Cantor-Lebesgue function $f : [0,1] \to [0,1]$ by

$$f(x) = \begin{cases} \sum_{j=1}^{\infty} a_j 2^{-j}, & x = \sum_{j=1}^{\infty} 2a_j 3^{-j} \in C \text{ for } a_j \in \{0,1\}; \\ \sup_{y \leq x, y \in C} f(y), & x \in [0,1] \setminus C. \end{cases}$$

Remark 8.9. The Cantor function challenges naive intuitions about continuity and measure; though it is continuous everywhere and has zero derivative almost everywhere, $f(x)$ goes from 0 to 1 as x goes from 0 to 1, and takes on every value in between. The Cantor function is the most frequently cited example of a real function that is uniformly continuous (precisely, it is Hölder continuous of exponent $\alpha = \log 2 / \log 3$) but not absolutely continuous. It is constant on intervals of the form $(0.x_1x_2x_3 \cdots x_n022222 \cdots, 0.x_1x_2x_3 \cdots x_n200000 \cdots)$, and every point not in the Cantor set is in one of these intervals, so its derivative is 0 outside of the Cantor set. On the other hand, it has no derivative at any point in an uncountable subset of the Cantor set containing the interval endpoints described above.

8.D. **Filling problems.**

Proposition 8.10. *Every open set $U \subset \mathbb{R}^n$ is a countable almost disjoint union of cubes.*

Proof. Consider the collection $\mathcal{A}_0$ of cubes of side length 1 defined by the lattice $\mathbb{Z}^n$. Set

$$\mathcal{U}_0 := \{Q \in \mathcal{A}_0 : Q \subset U\}, \quad \text{and} \quad \mathcal{D}_0 := \{Q \in \mathcal{A}_0 : Q \cap U \neq \varnothing, Q \cap U^c \neq \varnothing\}.$$

Let $\mathcal{A}_1$ be the collection of cubes that we obtain by subdividing each cube in $\mathcal{D}_0$ into 2^n cubes of side length $\frac{1}{2}$, and set

$$\mathcal{U}_1 := \{Q \in \mathcal{A}_1 : Q \subset U\}, \quad \text{and} \quad \mathcal{D}_1 := \{Q \in \mathcal{A}_1 : Q \cap U \neq \varnothing, Q \cap U^c \neq \varnothing\}.$$

Continue this process. Then $U = \bigcup_{Q \in \mathcal{U}} Q$ where $\mathcal{U} := \bigcup_{j=1}^{\infty} U_j$ is a countable almost disjoint union of cubes. □

Proposition 8.11. *Let $U \subset \mathbb{R}^n$ be open, $\delta > 0$. There exists a countable collection $\mathcal{G}$ of disjoint closed cubes in U such that $|Q| < \delta$ for each $B \in \mathcal{G}$ and*

$$m\left(U \setminus \bigcup_{Q \in \mathcal{G}} Q\right) = 0$$

Proof. Fix $0 < \theta_1 < \theta_2 < 1$. WLOG we assume that $m(U) < \infty$ (otherwise we apply the conclusion to $U_m = \{x \in U : m < |x| < m+1\}$ for $m = 0, 1, \cdots$). Via proposition 8.10 and subdividing, there exists a finite collection $(Q_j)_{j=1}^{M_1}$ of almost disjoint cubes in U such that $|Q_j| < \delta$ for $j = 1, \cdots, M_1$, and

$$m\left(U \setminus \bigcup_{j=1}^{M_1} Q_j\right) \le \theta_2 m(U)$$

Via a method like in remark 8.4, we can assume that Q_j is disjoint and

$$m\left(U \setminus \bigcup_{j=1}^{M_1} Q_j\right) \le \theta_1 m(U)$$

Now letting $U_2 := U \setminus \bigcup_{j=1}^{M_1} Q_j$, there exists, for the same reason, a finite collection $(Q_j)_{j=M_1+1}^{M_2}$ of isjoint cubes in U_2 such that

$$m\left(U \setminus \bigcup_{j=1}^{M_2} Q_j\right) = m\left(U_2 \setminus \bigcup_{j=M_1+1}^{M_2} Q_j\right) \le \theta_1 m\left(U_2\right) \le \theta_1^2 m\left(U\right)$$

Continue this process to obtain a countable collection of disjoint balls such that

$$m\left(U \setminus \bigcup_{j=1}^{M_k} Q_j\right) \le \theta_1^k m\left(U\right)$$

since $\theta_1^k \to 0$ as $k \to \infty$, the theorem is proved if $m(U) < \infty$. We are done. □

8.E. **Regularity, Radon measures on $\mathbb{R}^n$.** Our next aim is to show the regularity of Lebesgue measure on $\mathbb{R}^n$. First, we build the framework of Radon measures on $\mathbb{R}^n$.

Definition 8.12 (Regularity). *Let X be a topological space. A measure μ on a σ-algebra $\mathcal{X} \supset \mathcal{B}_X$ is called **outer regular** if*

$$\mu(E) = \inf\left\{\mu(U) : E \subset U, U \text{ is open}\right\}, \quad \forall E \in \mathcal{X}$$

*and **inner regular** if*

$$\mu(E) = \sup\left\{\mu(K) : K \subset E, K \text{ is compact}\right\}, \quad \forall E \in \mathcal{X}$$

*If μ is both outer and inner regular, it is called **regular**.*

*An outer measure μ^* on $\mathbb{R}^n$ is call **Borel regular** if for each $E \subset \mathbb{R}^n$, there exists a Borel set $B \supset E$ such that $\mu^*(E) = \mu^*(B)$.*

Definition 8.13 (Radon measure on $\mathbb{R}^n$). *A **Radon measure** on $\mathbb{R}^n$ is a Borel measure that is finite on compact sets.*

Remark 8.14. As in definition 11.2, more generally, a Radon measure on a LCH space is a Borel measure that not only has finiteness on compact sets, but also has regularity. In fact, on $\mathbb{R}^n$, finiteness on compact sets implies regularity. See theorem 8.16.

Remark 8.15. Moreover, as pointed out in remark 7.21, some people regard outer measures and measures as the same things. Actually, under that identification, a Radon measure μ on $\mathbb{R}^n$ is equivalent to a **Radon outer measure** μ^* on $\mathbb{R}^n$, where a **Radon outer measure** is an outer measure μ^* on $\mathbb{R}^n$ satisfying:

(1) μ^* is Borel regular;
(2) All Borel sets are μ^*-measurable;
(3) $\mu^*(K) < \infty$ for each compact set $K \subset \mathbb{R}^n$.

See proposition 8.17.

Now we give detailed explanations of the above remarks.

Theorem 8.16 (Regularity of Radon measure on $\mathbb{R}^n$). *Each Radon measure μ on $\mathbb{R}^n$ is σ-finite and regular. For each Borel set A and each $\varepsilon > 0$ there is an open set U and a closed set F so that*

$$F \subset A \subset U, \quad \textit{and} \quad \mu(U \setminus F) \leq \varepsilon. \tag{8.1}$$

Proof. Obviously μ is σ-finite, since $\mathbb{R}^n = \bigcup_{n=1}^{\infty} \overline{B_n(0)}$. Let we prove (8.1).

First we assume that μ is finite. Let $\mathcal{A}$ be the set of all Borel sets A that satisfy (8.1). It suffices to prove that $\mathcal{A}$ is a σ-algebra that contains all closed subsets.

It's clear that $\mathcal{A}$ is closed under complements. To show that $\mathcal{A}$ is a σ-algebra it suffices to show that $\mathcal{A}$ is closed under countable unions. Suppose that $A_j \in \mathcal{A}$, $j \geq 1$, and $\varepsilon > 0$. So there exist open U_j's and closed F_j's such that $F_j \subset A_j \subset U_j$ and $\mu(U_j \setminus F_j) \leq \varepsilon 2^{-j-1}$. Then $U := \bigcup_{j=1}^{\infty} U_j$ is open and $F_k := \bigcup_{j=1}^{k} F_j$ is closed for finite k. Note that $F_k \subset \bigcup_{j=1}^{\infty} A_j \subset U$ and

$$U \setminus F_k \subset \left(\bigcup_{j=1}^{\infty} (U_j \setminus F_j) \right) \bigcup \left(\bigcup_{j=1}^{\infty} (F_j \setminus F_k) \right).$$

Since μ is finite, for k suffices large, we have $\mu\left(\bigcup_{j=1}^{\infty} F_j\right) - \mu(F_k) < \frac{\varepsilon}{2}$, and hence

$$\mu(U \setminus F_k) \leq \sum_{j=1}^{\infty} \mu(U_j \setminus F_j) + \mu\left(\bigcup_{j=1}^{\infty} F_j\right) - \mu(F_k) < \varepsilon.$$

Thus $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$.

Let F be a closed subset. Setting $U_j = \left\{x : \operatorname{dist}(x, F) < \frac{1}{j}\right\}$, then U_j is open for each j and $\bigcup_{j=1}^{\infty} U_j = F$. Since μ is finite, $\lim_{j\to\infty} \mu(U_j) = \mu(F)$. It follows that $F \in \mathcal{A}$.

Assume that μ is not finite. Let A be a Borel set and let $\varepsilon > 0$ be given. It's clear that $\nu_F(E) := \mu(E \cap F)$ is a finite Radon measure on $\mathbb{R}^n$ if $F \in \mathcal{B}_{\mathbb{R}^n}$ and $F \subset K$ for some compact subset K. By the above, for $j = 1, 2, \cdots$, there exists a closed set $C_j \subset (B_j(0) \setminus A)$ with $\nu_{B_j(0)}\left(B_j(0) \setminus (A \cup C_j)\right) = \mu\left(B_j(0) \setminus (A \cup C_j)\right) \leq \varepsilon 2^{-j}$. Then $U := \bigcup_{j=1}^{\infty} (B_j(0) \setminus C_j)$ is open, $A \subset U$, and

$$\mu(U \setminus A) \leq \sum_{j=1}^{\infty} \mu\left(B_j(0) \setminus (A \cup C_j)\right) \leq \varepsilon.$$

Similarly, there exists a closed set $F_j \subset A_j := A \cap \{x \in \mathbb{R}^n : i \le |x| < i+1\}$ with $\mu(A_j \setminus F_j) \le \varepsilon 2^{-j-1}$. Then $F := \bigcup_{j=0}^{\infty} F_j \subset \bigcup_{j=0}^{\infty} A_j = A$, and

$$\mu(A \setminus F) \le \sum_{j=0}^{\infty} \mu(A_j \setminus F_j) \le \varepsilon$$

It remains to show that F is closed. It's well known that for a locally finite collection $(S_i)_{i \in I}$ of subsets of a topological space X, we have $\overline{\bigcup_{i\in I} S_i} = \bigcup_{i \in I} \overline{S_i}$. The conclusion follows.

Finally we prove that μ is regular. The outer regular easily follows from (8.1). Moreover, from the above we know that

$$\mu(A) = \sup\{\mu(F) : F \subset A, F \text{ is closed}\}, \quad \forall A \in \mathcal{B}_{\mathbb{R}^n}. \tag{8.2}$$

Note that for any closed $F \subset \mathbb{R}^n$ the sets $K_j = F \cap \overline{B_j(0)}$ are compact and $\mu(F) = \lim_{j\to\infty} \mu(K_j)$. It follows that μ is inner regular. □

Proposition 8.17. *Let μ be a Radon measure on $\mathbb{R}^n$. Then the outer measure μ^* induced by proposition 7.18 is a Radon outer measure (see remark 8.15 for definition). Conversely, if an outer measure μ^* on $\mathbb{R}^n$ is a Radon outer measure, then $\mu^*|_{\mathcal{B}_{\mathbb{R}^n}}$ is a Radon measure.*

Proof. Suppose that μ is a Radon measure on $\mathbb{R}^n$. As pointed out in remark 7.21, $\mu^*|_{\mathcal{B}_{\mathbb{R}^n}} = \mu$. Then (2) and (3) in remark 8.15 follows. If $\mu^*(E) = \infty$, take $B = \mathbb{R}^n$. Suppose that $\mu^*(E) < \infty$. For each $k \ge 1$, choose a countable collection $\mathcal{A}_k$ of Borel sets so that

$$E \subset \bigcup_{A \in \mathcal{A}_k} A =: B_k, \quad \text{and} \quad \sum_{A \in \mathcal{A}_k} \mu(A) \le \mu^*(E) + \frac{1}{k}$$

Then $B = \bigcap_{k=1}^{\infty} B_k$ is a Borel set that satisfies $E \subset B$ and

$$\mu^*(B) \le \mu(B_k) \le \sum_{A \in \mathcal{A}_k} \mu(A) \le \mu^*(E) + \frac{1}{k}, \quad \forall k$$

Hence $\mu^*(E) = \mu^*(B)$.

The converse statement is obvious. □

For more properties of Randon measures on $\mathbb{R}^n$, one can refer to [Evaa].

8.F. Regularity of Lebesgue measure, F_σ and G_δ sets.
In the next we come back the regularity of Lebesgue measure.

Definition 8.18 (F_σ and G_δ sets)**.** *An F_σ set is a countable union of closed sets, and a G_δ set is a countable intersection of open sets.*

Theorem 8.19 (Properties of the Lebesgue measure)**.** *Let m denote the Lebesgue measure and let m^* denote the Lebesgue outer measure.*

(1) m^* is Borel regular.
(2) $m|_{\mathcal{B}_{\mathbb{R}^n}}$ is a Radon measure, and m is regular and σ-finite.
(3) A set $E \subset \mathbb{R}^n$ is Lebesgue measurable iff there is an F_σ set A and a G_δ set B satisfying $A \subset E \subset B$ and $m(B \setminus A) = 0$.

Moreover, the completion of $\left(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, m|_{\mathcal{B}(\mathbb{R}^n)}\right)$ *is* $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), m)$*. In particular,* $\mathcal{L}(\mathbb{R}^n)$ *is the completion of the Borel* σ*-algebra* $\mathcal{B}_{\mathbb{R}^n}$ *with respect to the Lebesgue measure.*

Proof. (1) If $m^*(E) = \infty$, take $B = \mathbb{R}^n$. Suppose that $m^*(E) < \infty$. For each $k \geq 1$, choose a countable collection $\mathcal{A}_k$ of cubes so that

$$E \subset \bigcup_{Q \in \mathcal{A}_k} Q =: B_k, \quad \text{and} \quad \sum_{Q \in \mathcal{A}_k} |Q| \leq m^*(E) + \frac{1}{k}$$

Then $B = \bigcap_{k=1}^{\infty} B_k$ is a Borel set that satisfies $E \subset B$ and

$$m^*(B) \leq m^*(B_k) \leq \sum_{Q \in \mathcal{A}_k} |Q| \leq m^*(E) + \frac{1}{k}, \quad \forall k$$

Hence $m^*(E) = m^*(B)$.

(2) It's clear that $m|_{\mathcal{B}_{\mathbb{R}^n}}$ has finiteness on compact sets and hence is a Randon measure. Thus m is also σ-finite. Via theorem 8.16,

$$\begin{aligned} \forall B \in \mathcal{B}_{\mathbb{R}^n} : m(B) &= \inf\{m(U) : B \subset U, U \text{ is open}\} \\ &= \sup\{m(K) : K \subset B, K \text{ is compact}\}. \end{aligned}$$

If $E \in \mathcal{L}(\mathbb{R}^n)$, by (1) there exists a Borel set B with $E \subset B$ and $m(E) = m^*(E) = m^*(B) = m(B)$, and hence

$$\begin{aligned} m(E) = m(B) &= \inf\{m(U) : B \subset U, U \text{ is open}\} \\ &\geq \inf\{m(U) : E \subset U, U \text{ is open}\} \geq m(E) \end{aligned}$$

which implies that m is outer regular.

To see that m is inner regular, let $E \subset \mathbb{R}^n$ be measurable, and suppose first that E is contained in a cube Q. Let $\varepsilon > 0$. Since $m(Q \setminus E) < \infty$ and m is outer regular, there exist an open set $U \supset (Q \setminus E)$ with $m(U) \leq m(Q \setminus E) + \varepsilon$. The set $K := Q \setminus U \subset E$ is compact and satisfies

$$\begin{aligned} m(E) = m(Q) - m(Q \setminus E) &\leq m(Q) - m(U) + \varepsilon \\ &\leq m(Q) - m(Q \cap U) + \varepsilon = m(K) + \varepsilon \end{aligned}$$

If E is not contained in a cube, for each $j \geq 1$, there is a compact $K_j \subset [-j, j]^n$ so that $m(K_j) \geq m(E \cap [-j, j]^n) - \frac{1}{j}$. Hence $m(K_j) \to m(E)$ as $k \to \infty$ and hence m is inner regular.

(3) Assume that E is Lebesgue measurable. By (2) there exist open sets G_j's and closed sets F_j's satisfying $F_j \subset E \subset G_k$ and $m(G_j \setminus F_j) \leq \frac{1}{j}$. Then sets $F = \bigcup_{j=1}^{\infty} F_j$ and $G = \bigcap_{j=1}^{\infty} G_j$ are as required.

Conversely, suppose that there exist such F and G for E. It suffices to show that for any $A \subset \mathbb{R}^n$ we have

$$m^*(A \cap E) + m^*(A \cap E^c) = m^*(A).$$

Since we have $m^*(A \cap F) + m^*(A \cap F^c) = m^*(A)$, it suffices to show that $m^*(A \cap F) = m^*(A \cap E)$ and $m^*(A \cap E^c) = m^*(A \cap F^c)$. Note that $A \cap F \subset A \cap E \subset A \cap G$ and

$A \cap G^c \subset A \cap E^c \subset A \cap F^c$. Also note that

$$m^* ((A \cap G) \setminus (A \cap F)) = m^* (A \cap (G \setminus F)) \leq m^* (G \setminus F) = 0,$$

and similarly $m^* ((A \cap F^c) \setminus (A \cap G^c)) = 0$. It follows that $m^*(A \cap E) = m^*(A \cap F)$ and $m^*(A \cap E^c) = m^*(A \cap F^c)$. We are done.

Finally, it follows from (3) and theorem 7.12 that $\mathcal{L}(\mathbb{R}^n) \subset \overline{\mathcal{B}_{\mathbb{R}^n}}$ and $\overline{m}|_{\mathcal{L}(\mathbb{R}^n)} = m|_{\mathcal{L}(\mathbb{R}^n)}$, where $\left(\mathbb{R}^n, \overline{m}|_{\mathcal{L}(\mathbb{R}^n)}, \overline{m}\right)$ is the completion of $\left(\mathbb{R}^n, m|_{\mathcal{L}(\mathbb{R}^n)}, m\right)$. Then it follows from proposition 7.14 that the completion of $\left(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, m|_{\mathcal{B}(\mathbb{R}^n)}\right)$ is $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), m)$, since $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), m)$ is complete. □

8.G. **Existence of non-measurable sets.**

Theorem 8.20 (Existence of non-measurable sets)**.** *On $\mathbb{R}^n$ consider the equivalence relation $x \sim y$ iff $x - y \in \mathbb{Q}^n$. The axiom of choice allows us to choose exactly one element in each equivalence class and to gather these elements in one set N; such a set is called a **Vitali set**. N is not Lebesgue measurable.*

Moreover, given $E \subset \mathbb{R}^n$, then $m(E) = 0$ iff every subset of E is Lebesgue measurable.

Proof. Note that

$$\mathbb{R}^n = \bigcup_{x \in N} (x + \mathbb{Q}^n) = \bigcup_{j=1}^{\infty} (r_j + N)$$

where $\mathbb{Q}^n = \{r_1, r_2, \cdots\}$ and these are two disjoint union. To show that $N \notin \mathcal{L}(\mathbb{R}^n)$, via theorem 8.16, it suffices to prove that $m^*(N) > 0$ and $m_*(N) := \sup \{\mu(F) : F \subset A, F \text{ is compact}\} = 0$.

First note that $m^*(N) > 0$; otherwise via theorem 8.19 we have $m(E) = m^*(E) = 0$, and hence $m(r_k + N) = 0$ for each k, which implies that $m(\mathbb{R}^n) = 0$, a contradiction.

Then we show that $m_*(E) = 0$. Given any compact set $K \subset E$. Then setting $D = B_1(0) \cap \mathbb{Q}^n$ we know that $\bigcup_{r \in D} (r + K)$ is a disjoint union and is bounded. Hence $m\left(\bigcup_{r \in D} (r + K)\right) = \sum_{r \in D} m(K) < \infty$, and it follows that $m(K) = 0$. Thus $m_*(E) = 0$.

Now we prove the second assertion. The "only if" implication follows from theorem 8.19. On the other hand, suppose that A is a Lebesgue measurable set with $m(A) > 0$. Note that

$$A = \bigcup_{j=1}^{\infty} \left((r_j + N) \cap A\right),$$

and

$$0 < m(A) = m^*(A) \leq \sum_{j=1}^{\infty} m^* \left((r_j + N) \cap A\right),$$

Hence there exist $k \geq 1$ such that $B := (r_k + N) \cap A$ has positive outer measure. Also note that

$$\begin{aligned} m_*(B) &:= \sup \{\mu(F) : F \subset B, F \text{ is compact}\} \\ &\leq \sup \{\mu(F) : F \subset (r_k + N), F \text{ is compact}\} = 0 \end{aligned}$$

Hence B is not Lebesgue measurable. We are done. □

Remark 8.21. In the previous proof the axiom of choice plays an essential role. In fact, Solovay constructed a model in which all axioms of Zermelo–Frankel set theory, except the axiom of choice, hold and in which every subset of R is Lebesgue measurable.

8.H. **Uniqueness, product property.** In the next we introduce the uniqueness of Lebesgue measure.

Definition 8.22 (Translation invariant). *We call a measure μ on $\mathbb{R}^n$* ***translation invariant*** *if it satisfies that if E is measurable and $x \in \mathbb{R}^n$, then $x + E$ is measurable and $\mu(x + E) = \mu(E)$.*

Theorem 8.23 (Uniqueness of Lebesgue measure). *There are some editions.*

(1) $m|_{\mathcal{B}_{\mathbb{R}^n}}$ is the unique measure on $\mathcal{B}_{\mathbb{R}^n}$ satisfying $m([a, b]) = |[a, b]|$.
(2) m^* is the unique Borel regular outer measure such that all Borel sets are measurable and $m^*([a, b]) = |[a, b]|$.
(3) $m|_{\mathcal{B}_{\mathbb{R}^n}}$ is the unique translation invariant Randon measure on $\mathcal{B}_{\mathbb{R}^n}$ up to a scaling.

Proof. For (1), let $\mathcal{G}$ be the collection of sets of the form $F \cap G$, where F is closed and G is open. Then $\mathcal{G}$ is an elementary family. Via theorem 7.29, letting $\mathcal{A}$ be the collection of all finite disjoint unions of sets in $\mathcal{G}$, then $\mathcal{A}$ is an algebra. Via theorem 7.25 and remark 7.30, it suffices to show that any Borel measure μ on $\mathbb{R}^n$ satisfying $\mu([a, b]) = |[a, b]|$ will coincide with m on $\mathcal{G}$, since the σ-algebra generated by $\mathcal{A}$ is $\mathcal{B}_{\mathbb{R}^n}$.

Via proposition 8.10, μ coincide with m on all open sets. If F is closed and G is open, set $G_j := \left\{x \in \mathbb{R}^n : \operatorname{dist}(x, F) < \frac{1}{j}\right\}$. Then G_j is open, $G_j \supset G_{j+1}$, and $F = \bigcap_{j=1}^{\infty} G_j$. If $\mu(G) < \infty$, then $m(G) = \mu(G) < \infty$ and hence

$$\mu(F \cap G) = \mu\left(\bigcap_{j=1}^{\infty}(G_j \cap G)\right) = \lim_{j\to\infty} \mu(G_j \cap G) = \lim_{j\to\infty} m(G_j \cap G) = m\left(\bigcap_{j=1}^{\infty}(G_j \cap G)\right) = m(F \cap G)$$

If $\mu(G) = \infty$, then

$$\mu\left(F \cap G \cap (-k, k)^n\right) = m\left(F \cap G \cap (-k, k)^n\right)$$

and letting $k \to \infty$ we have $\mu(F \cap G) = m(F \cap G)$. Then (1) follows.

Let μ^* be a Borel regular outer measure such that all Borel sets are measurable and $\mu^*([a, b]) = |[a, b]|$. By (1) and Carathéodory's theorem 7.19, μ^* coincide with m^* on all Borel sets. Let $E \subset \mathbb{R}^n$. Since they are Borel regular, there exist two Borel sets $B_1, B_2 \supset E$ such that $\mu^*(B_1) = \mu^*(E)$ and $m^*(B_2) = m^*(E)$. Setting $B = B_1 \cap B_2$, it's clear that $m^*(E) = m^*(B) = \mu^*(B) = \mu^*(E)$. Then (2) follows.

Let μ be a translation invariant Randon measure on $\mathcal{B}_{\mathbb{R}^n}$. Set $\mu([0, 1)^n) := C < \infty$. Consider the grid of dyadic cubes of the form $[a_1, b_1) \times \cdots \times [a_n, b_n)$ defined by the lattice $2^{-k}\mathbb{Z}^n$. Since these cubes are all translates of each other, we have

$$2^{kn}\mu(Q) = \mu([0, 1)^n) = Cm([0, 1)^n) = C2^{kn}m(Q),$$

for each such cube Q. We may infer by the regularity that μ vanishes on degenerate boxes, and so $\mu(Q) = Cm(Q)$ for each closed dyadic cube $Q = [a_1, b_1] \times \cdots \times [a_n, b_n]$. Also, We may infer by proposition 8.10 that $\mu(E) = C\lambda(E)$ for each open set E, and thus for each Borel set E by the regularity of μ and m. Then (3) follows. $\square$

Corollary 8.24. *The* $(m+n)$*-dimensional Lebesgue measure space*

$$\left(\mathbb{R}^{m+n}, \mathcal{L}(\mathbb{R}^{m+n}), m|_{\mathcal{L}(\mathbb{R}^{m+n})}\right)$$

is the completion of $(\mathbb{R}^m \times \mathbb{R}^n, \mathcal{L}(\mathbb{R}^m) \times \mathcal{L}(\mathbb{R}^n), m|_{\mathcal{L}(\mathbb{R}^m)} \times m|_{\mathcal{L}(\mathbb{R}^n)})$.

Proof. First we show that

$$\mathcal{B}_{\mathbb{R}^{m+n}} \subset \mathcal{L}(\mathbb{R}^m) \times \mathcal{L}(\mathbb{R}^n) \subset \mathcal{L}(\mathbb{R}^{m+n}).$$

The first inclusion follows from the fact that each cube in $\mathbb{R}^{m+n}$ belongs to $\mathcal{L}(\mathbb{R}^n) \times \mathcal{L}(\mathbb{R}^m)$, and $\mathcal{B}_{\mathbb{R}^{m+n}}$ is the σ-algebra generated by the cubes in $\mathbb{R}^{m+n}$. It follows from theorem 8.19 (3) that $E \times \mathbb{R}^n$ and $\mathbb{R}^m \times F$ belong to $\mathcal{L}(\mathbb{R}^{m+n})$ if $E \in \mathcal{L}(\mathbb{R}^m)$ and $F \in \mathcal{L}(\mathbb{R}^n)$; then the second inclusion follows via proposition 6.29.

Then note that $m|_{\mathcal{L}(\mathbb{R}^{m+n})}$ and $m|_{\mathcal{L}(\mathbb{R}^m)} \times m|_{\mathcal{L}(\mathbb{R}^n)}$ coincide on boxes, and hence coincide on $\mathcal{B}_{\mathbb{R}^n}$ via theorem 8.23 (1). Via theorem 8.19, the completion of $\left(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, m|_{\mathcal{B}(\mathbb{R}^n)}\right)$ is $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), m)$. Hence, via proposition 7.14, it suffices to show that $m|_{\mathcal{L}(\mathbb{R}^{m+n})}$ and $m|_{\mathcal{L}(\mathbb{R}^m)} \times m|_{\mathcal{L}(\mathbb{R}^n)}$ coincide on $\mathcal{L}(\mathbb{R}^m) \times \mathcal{L}(\mathbb{R}^n)$. For all $E \in \mathcal{L}(\mathbb{R}^m) \times \mathcal{L}(\mathbb{R}^n) \subset \mathcal{L}(\mathbb{R}^{m+n})$, it follows from theorem 8.19 (3) that $A \subset E \subset B$ for some $A, B \in \mathcal{B}_{\mathbb{R}^n}$ with $m(B \setminus A) = 0$. Note that

$$m(A) \leq m|_{\mathcal{L}(\mathbb{R}^m)} \times m|_{\mathcal{L}(\mathbb{R}^n)}(A) \leq m|_{\mathcal{L}(\mathbb{R}^m)} \times m|_{\mathcal{L}(\mathbb{R}^n)}(E) \leq m|_{\mathcal{L}(\mathbb{R}^m)} \times m|_{\mathcal{L}(\mathbb{R}^n)}(B) = m(B).$$

Then the desired coincidence follows. We are done. $\square$

For more properties such as *invariance properties*, one can refer to [Jon] or [Rai].

Moreover, since the completion of $\left(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, m|_{\mathcal{B}(\mathbb{R}^n)}\right)$ is $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), m)$, some conclusions of the Radon measure $\left(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, m|_{\mathcal{B}(\mathbb{R}^n)}\right)$ also apply to $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), m)$. Hence, one can refer to the section 11, which is about Radon measures, for more properties of Lebesgue measures.

9. Integration on a measure space

In this section we introduce the integration on a measure space and its properties. First we fix the arithmetic in $[0, \infty]$. We define

$$a + \infty = \infty + a = \infty \quad \text{if } a \in [0, \infty];$$
$$a \cdot \infty = \infty \cdot a = \begin{cases} \infty, & a \in (0, \infty] \\ 0, & a = 0. \end{cases}$$

Then addition and multiplication in $[0, \infty]$ are commutative, associative, and distributive. The cancellation laws have to be treated with some care; $a + c = b + c$ implies $a = b$ only if $c \in [0, \infty)$, and $ac = bc$ implies $a = b$ only if $c \in (0, \infty)$.

In the next we define the integration in three steps:

(1) Define the integration of *non-negative simple functions*;
(2) Induce the integration of *non-negative measurable functions*;
(3) Induce the integration of *measurable real-valued and complex-valued functions*.

9.A. Integration on L^+.

Definition 9.1 (L^+ space)**.** *Let $(X, \mathcal{X}, \mu)$ be a measure space. Then we define*

$$L^+ = L^+(X) := \textit{the space of all measurable functions from } X \textit{ to } [0, \infty].$$

Definition 9.2 (Integration of non-negative simple functions)**.** *Let $(X, \mathcal{X}, \mu)$ be a measure space. The **integration** $\int \varphi \, d\mu$ with respect to the measure μ of a simple function $\varphi : X \to [0, \infty]$ with standard representation $\varphi = \sum_{j=1}^N a_j \chi_{E_j}$ is defined by*

$$\int \varphi \, d\mu = \int_X \varphi \, d\mu := \sum_{j=1}^N a_j \mu(E_j).$$

Moreover, if $E \subset \mathcal{X}$, then we set

$$\int_E \varphi \, d\mu = \int \varphi \chi_E \, d\mu = \sum_{j=1}^N a_j \mu(E_j \cap E).$$

Proposition 9.3. *Let φ and ψ be simple functions in L^+.*

(1) If $c \in [0, \infty)$, $\int c\varphi \, d\mu = c \int \varphi \, d\mu$.
(2) $\int (\varphi + \psi) \, d\mu = \int \varphi \, d\mu + \int \psi \, d\mu$.
(3) If $\varphi \le \psi$, then $\int \varphi \, d\mu \le \int \psi \, d\mu$.
(4) The map $A \mapsto \int_A \varphi \, d\mu$ is a measure on $\mathcal{X}$.

Proof. Trivial. Just use the common refinement. □

Remark 9.4. In the next we will see that in (4) we can change φ to any $f \in L^+$.

Remark 9.5. The integration itself is of course of great significance. But on the other hand, a new perspective of studying measures is inspired: the measures form functionals on a function space via the integration. (They finally form functionals on L^1.)

More precisely, under good conditions, we have the Riesz representation theorem 11.6.

Now we induce the integration of functions in L^+.

Definition 9.6 (Integration of non-negative measurable functions)**.** *Let $(X, \mathcal{X}, \mu)$ be a measure space. The* ***integration*** *$\int f\, d\mu$ with respect to the measure μ of a function $f \in L^+$ is defined by*

$$\int f\, d\mu = \sup\left\{\int_X \varphi\, d\mu \,:\, 0 \le \varphi \le f, \varphi \text{ is simple}\right\}.$$

Moreover, if $E \subset \mathcal{X}$, then we set

$$\int_E f\, d\mu = \int f\chi_E\, d\mu.$$

Remark 9.7. By proposition 9.3 (3) we know that two definitions $\int f\, d\mu$ agree when f is simple. Moreover, it's obvious from the definition that for $f, g \in L^+$,

$$\int f\, d\mu \le \int g\, d\mu \text{ whenever } f \le g, \quad \text{and} \quad \int cf\, d\mu = c\int f\, d\mu \text{ if } c \ge 0.$$

Also, the integration has additivity. See corollary 9.14 for the general cases.

Proposition 9.8. *Let $(X, \mathcal{X}, \mu)$ be a measure space. If $f \in L^+$, then $\int f\, d\mu = 0$ iff $f = 0$ a.e.*

Proof. If f is simple, the conclusion is trivial. In general, if $f = 0$ a.e. and φ is simple with $0 \le \varphi \le f$, then $\varphi = 0$ a.e. and hence $\int f\, d\mu = \sup_{0\le\varphi\le f} \int \varphi\, d\mu = 0$. On the other hand, if it false that $f = 0$ a.e., then it follows that $\mu(E_n) > 0$ for some n where $E_n := \left\{x : f(x) < \frac{1}{n}\right\}$, since $\{x \in X : f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n$. Note that $f > \frac{1}{n}\chi_{E_n}$; hence $\int f\, d\mu \ge \frac{1}{n}\mu(E_n) > 0$. Then the conclusion follows. □

Proposition 9.9. *Let $(X, \mathcal{X}, \mu)$ be a measure space. If $f \in L^+$ and $\int f\, d\mu < \infty$, then $A := \{x : f(x) = \infty\}$ is a null set and $B := \{x : f(x) > 0\}$ is σ-finite.*

Proof. Trivial. □

9.B. **(Improved) monotone convergence theorem, Fatou's lemma.** Theorem 6.22 shows that f can be approximated pointwisely by a sequence of monotonely incresing simple functions (which is uniformly when f is bounded). Actually this monotone approximation leads to the approximation in the sense of integration. More precisely, we have the monotone convergence theorem.

Theorem 9.10 (The monotone convergence theorem)**.** *Let $(X, \mathcal{X}, \mu)$ be a measure space. If (f_n) is a sequence in L^+ such that $f_j \le f_{j+1}$ for all j, and $f = \lim_{j\to\infty} f_j$, then $f \in L^+$, and $\int f\, d\mu = \lim_{j\to\infty} \int f_j\, d\mu$.*

Proof. Since $f = \sup_{j\ge1} f_j$, f is measurable via proposition 6.18. It's clear that $\int f\, d\mu \ge \int f_j\, d\mu$ for all j, and hence $\int f\, d\mu \ge \lim_{j\to\infty} \int f_j\, d\mu$, where $(\int f_j\, d\mu)_{j=1}^{\infty}$ is incresing and hence the limit makes sense.

For the converse inequality, fix $\alpha \in (0,1)$. It suffices to show that for any simple function φ with $0 \leq \varphi \leq f$ we have

$$\lim_{j\to\infty} \int f_j \, d\mu \geq \alpha \int \varphi \, d\mu.$$

Setting $E_j := \{x : f_j(x) \geq a\varphi(x)\}$, then $E_1 \subset E_2 \subset \cdots$ and $X = \bigcup_{j=1}^{\infty} E_j$. Hence

$$\int f_j \, d\mu \geq \int_{E_j} f_j \, d\mu \geq a \int_{E_j} \varphi \, d\mu.$$

It follows that

$$\lim_{j\to\infty} \int f_j \, d\mu \geq a \lim_{j\to\infty} \int_{E_j} \varphi \, d\mu = a \int \varphi \, d\mu,$$

since $E \mapsto \int_E \varphi \, d\mu$ is a measure. We are done. □

Remark 9.11. Theorem 9.10 and theorem 6.22 give us a new method to define the integration of $f \in L^+$. But in this way we need to show that the defined integration is independent from the choice of (φ_n).

Remark 9.12. In fact, in the monotone incresing theorem, if f_n increses to f a.e., we already have $\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu$. We will show this later.

Remark 9.13. The hypothesis that the sequence (f_n) be incresing, at least a.e., is essential for the monotone convergence theorem. For example, consider $(\mathbb{R}, \mathcal{L}(\mathbb{R}), m)$ and $f_n = \chi_{(n,n+1)}$.

Corollary 9.14 (Additivity). *Let $(X, \mathcal{X}, \mu)$ be a measure space. If (f_n) is a finite or infinite sequence in L^+ and $f = \sum_{n=1}^{\infty} f_n$, then $\int f \, d\mu = \sum_{n=1}^{\infty} \int f_n \, d\mu$.*

Proof. First we prove the statement for the sum of two functions f_1 and f_2. By theorem 6.22, there exist two sequences $(\varphi_j)_{j=1}^{\infty}$ and $(\psi_j)_{j=1}^{\infty}$ of non-negative simple functions that increse to f_1 and f_2. Then $(\varphi_j+\psi_j)_{j=1}^{\infty}$ increses to f_1+f_2, so by the monotone convergence theorem 9.10 and proposition 9.3,

$$\int (f_1 + f_2) \, d\mu = \lim_{j\to\infty} \int (\varphi_j + \psi_j) \, d\mu = \lim_{j\to\infty} \int \varphi_j \, d\mu + \lim_{j\to\infty} \int \psi_j \, d\mu = \int f_1 \, d\mu + \int f_1 \, d\mu.$$

Hence, by induction, $\int \left(\sum_{j=1}^{N} f_j\right) d\mu = \sum_{j=1}^{N} \int f_j \, d\mu$ for any finite N. Letting $N \to \infty$ and applying the monotone convergence theorem 9.10 again, we obtain the conclusion. □

Corollary 9.15. *Let $(X, \mathcal{X}, \mu)$ be a measure space. If $f \in L^+$ and set $\nu(E) = \int_E f \, d\mu$ for $E \in \mathcal{X}$, then ν is a measure on $\mathcal{X}$, and for any $g \in L^+$,*

$$\int g \, d\nu = \int fg \, d\mu. \tag{9.1}$$

Proof. Let $(E_n)_{n=1}^{\infty}$ is a sequence of disjoint sets in $\mathcal{X}$. Then via corollary 9.14,

$$\nu\left(\bigcup_{n=1}^{\infty} E_n\right) = \int \sum_{n=1}^{\infty} f\chi_{E_n} \, d\mu = \sum_{n=1}^{\infty} \int f\chi_{E_n} \, d\mu = \sum_{n=1}^{\infty} \nu(E_n)$$

It follows that ν is a measure on $\mathcal{X}$. By definition, (9.1) holds for $g = \chi_E$, $E \in \mathcal{X}$. Moreover, via corollary 9.14, for each non-negative simple function we have

$$\int \sum_{j=1}^N a_j \chi_{E_j} \, d\nu = \sum_{j=1}^N a_j \int \chi_{E_j} \, d\nu = \sum_{j=1}^N a_j \int f \chi_{E_j} \, d\mu = \int f \sum_{j=1}^N a_j \chi_{E_j} \, d\mu,$$

and hence (9.1) holds for non-negative simple functions. The general case follows from theorem 6.22 and the monotone convergence theorem 9.10. □

Remark 9.16. This corollary inspires us that we may represent a measure via a give measure. Actually, we have the Radon-Nikodym-Lebesgue theorem 10.18.

Now we come back to remark 9.12 and improve the monotone convergence theorem.

Corollary 9.17 (Improved monotone convergence theorem). *Let $(X, \mathcal{X}, \mu)$ be a measure space. If $(f_n) \subset L^+$, $f \in L^+$, and f_n increses to f a.e., then $\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu$.*

Proof. If $f_n(x)$ increses to $f(x)$ for $x \in E$ and $\mu(E^c) = 0$, then $f - f\chi_E = 0$ a.e. and $f_n - f_n\chi_E = 0$ a.e., so by the monotone convergence theorem 9.10, we have $\int f \, d\mu = \int f \chi_E \, d\mu = \lim_{n\to\infty} \int f_n \chi_E \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu$. □

Remark 9.18. Actually under the purposes of integration, we can alter functions on null sets. Thus via proposition 6.13 and proposition 6.14, $\lim_{j\to\infty} f_j$ can be regarded as an a.e.-defined measurable function under the purposes of integration, and hence we can remove the hypothesis that f is in L^+ in the above corollary.

As remark 9.13 said, being incresing is essential for the monotone convergence theorem. If we remove this condition, we also have Fatou's lemma.

Theorem 9.19 (Fatou's lemma). *Let $(X, \mathcal{X}, \mu)$ be a measure space. If $(f_n) \subset L^+$, then*

$$\int \liminf_{n\to\infty} f_n \, d\mu \le \liminf_{n\to\infty} \int f_n \, d\mu.$$

Proof. Setting $g_k := \inf_{n\ge k} f_n \in L^+$, then g_k increses to $\liminf_{n\to\infty} f_n$, and hence

$$\int \liminf_{n\to\infty} f_n \, d\mu = \lim_{k\to\infty} \int g_k \, d\mu.$$

Therefore, it suffices to show that

$$\int g_k \, d\mu \le \inf_{j\ge k} \int f_j \, d\mu.$$

Note that $g_k = \inf_{n\ge k} f_n \le f_j$ for all $j \ge k$. It follows that $\int g_k \, d\mu \le \int f_j \, d\mu$ for all $j \ge k$. Then the conclusion follows. □

Corollary 9.20. *Let $(X, \mathcal{X}, \mu)$ be a measure space. If $(f_n) \subset L^+$, $f \in L^+$ and $f_n \to f$ a.e., then $\int f \, d\mu \le \liminf_{n\to\infty} \int f_n \, d\mu$.*

9.C. **Integration of measurable functions, the dominated convergence theorem.** Now we induce the integration of measure real-valued and complex-valued functions.

Definition 9.21 (Integration of measurable functions)**.** *Let $(X, \mathcal{X}, \mu)$ be a measure space. If $f : X \to [-\infty, +\infty]$ is measurable, then the* ***integration*** *$\int f \, d\mu$ with respect to the measure μ of f is defined by*

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu.$$

We shall be mainly concerned with the case where $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ are both finite; we then say that f is ***integrable****. Since $|f| = f^+ + f^-$, it's clear that f is integrable iff $\int |f| \, d\mu < \infty$.*

Next, if $f : X \to \mathbb{C}$ is measurable, then the ***integration*** *$\int f \, d\mu$ with respect to the measure μ of f is defined by*

$$\int f \, d\mu = \int \Re f \, d\mu + i \int \Im f \, d\mu.$$

We say that f is ***integrable*** *if $\int |f| \, d\mu < \infty$. More generally, if $E \in \mathcal{X}$, f is* ***integrable on*** *E if $\int_E |f| < \infty$. Since $|f| \leq |\Re f| + |\Im f| \leq 2|f|$, f is integrable iff $\Re f$ and $\Im f$ are both integrable. We denote the space of complex-valued integrable functions by $L^1(\mu)$* ***(or*** *$L^1(X, \mu)$****, or*** *$L^1(\mu)$****, or simply*** *L^1****.)****.*

Remark 9.22. Sometimes we also write the integration in the following form:

$$\int f(x) \, d\mu(x) = \int f \, d\mu.$$

Proposition 9.23. *Let $(X, \mathcal{X}, \mu)$ be a measure space.*

(1) (Linearity) *The set of integrable real-valued (or complex-valued) functions on X is a real (or complex) vector space, and the integration is a linear functional on it.*

(2) (Monotony) *If f and g are real-valued and $f \leq g$, then*

$$\int f \, d\mu \leq \int g \, d\mu.$$

(3) (Triangle inequality) *If $f \in L^1$, then*

$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu.$$

(4) $(\sigma - additivity)$ If $(E_n)_{n=1}^\infty$ is a sequence of disjoint sets in $\mathcal{X}$ and $f \in L^1$, then

$$\int_{\bigcup_{n=1}^\infty E_n} f \, d\mu = \sum_{n=1}^\infty \int_{E_n} f \, d\mu$$

(5) If $f \in L^1$, then $\{x : f(x) \neq 0\}$ is σ-finite.

(6) If $f, g \in L^1$, then $\int_E f \, d\mu = \int_E g \, d\mu$ for all $E \in \mathcal{X}$ iff $\int |f - g| \, d\mu = 0$ iff $f = g$ a.e.

Proof. Trivial. □

Remark 9.24. This proposition shows that for the purposes of integration it makes no difference if we alter functions on null sets (this is feasible via proposition 6.13). In this fashion we can treat $\overline{\mathbb{R}}$-valued functions that are finite a.e. as $\mathbb{R}$-valued functions for the purposes of integration.

With this in mind, we shall find it more convenient to **redefine $L^1(\mu)$ to be the set of equivalence classes of a.e.-defined integrable functions on X**, where f and g are considered equivalent iff $f = g$ a.e. This new $L^1(\mu)$ is still a complex vector space, and we shall still employ the notation "$f \in L^1(\mu)$" to mean that f is an a.e.-defined integrable function.

Moreover, under the new definition, $L^1(\mu)$ is a metric space with distance function $\rho(f, g) = \int |f - g|\, d\mu$. We shall refer to the convergence with respect to this metric as *convergence in L^1*; thus $f_n \to f$ in L^1 iff $\int |f_n - f|\, d\mu \to 0$.

Now we present the last of the three basic convergence theorems.

Theorem 9.25 (The dominated convergence theorem). *Let $(X, \mathcal{X}, \mu)$ be a measure space. Let (f_n) be a sequence in L^1 such that*

(1) $f_n \to f$ a.e., and
(2) there exists a non-negative $g \in L^1$ such that $|f_n| \leq g$ a.e. for all n.

Then $f \in L^1$ and $\int f\, d\mu = \lim_{n\to\infty} f_n\, d\mu$.

Proof. It follows from proposition 6.18 and proposition 6.13 that f is measurable (perhaps after redefinition on a null set). Since $|f| \leq g$ a.e., we have $f \in L^1$. By taking real and imaginary parts it suffices to assume that f_n and f are real-valued, in which case we have $g + f_n \geq 0$ a.e. and $g - f_n \geq 0$ a.e. Thus by Fatou's lemma 9.19,

$$\int g\, d\mu + \int f\, d\mu \leq \liminf_{n\to\infty} \int (g + f_n)\, d\mu = \int g\, d\mu + \liminf_{n\to\infty} \int f_n\, d\mu,$$
$$\int g\, d\mu - \int f\, d\mu \leq \liminf_{n\to\infty} \int (g - f_n)\, d\mu = \int g\, d\mu - \limsup_{n\to\infty} \int f_n\, d\mu,$$

Therefore, $\liminf_{n\to\infty} \int f_n\, d\mu \geq \int f\, d\mu \geq \limsup_{n\to\infty} \int f_n\, d\mu$, and the result follows. □

Corollary 9.26 (Improved additivity). *Let $(X, \mathcal{X}, \mu)$ be a measure space. Suppose that (f_j) is a sequence in L^1 such that $\sum_{j=1}^{\infty} \int |f_j|\, d\mu < \infty$. Then $\sum_{j=1}^{\infty} f_j$ converges a.e. to a function in L^1, and $\int \left(\sum_{j=1}^{\infty} f_j\right) d\mu = \sum_{j=1}^{\infty} \int f_j\, d\mu$.*

Proof. By corollary 9.14, we have

$$\int \sum_{j=1}^{\infty} |f_j|\, d\mu = \sum_{j=1}^{\infty} \int |f_j|\, d\mu < \infty,$$

so the function $g = \sum_{j=1}^{\infty} |f_j|$ is in L^1. In particular, by proposition 9.9, g is finite for a.e. x, and for each such x the series $\sum_{j=1}^{\infty} f_j(x)$ converges. Moreover, $|\sum_{j=1}^{n} f_j| \leq g$ for all n, so we can apply the dominated convergence theorem 9.25 to the sequence of partial sum to obtain $\int \left(\sum_{j=1}^{\infty} f_j\right) d\mu = \sum_{j=1}^{\infty} \int f_j\, d\mu$. □

For more application, such as the conclusions about the integration depending on parameters, one can refer to [For] or [Rai].

9.D. **Product measure, Fubini-Tonelli theorem.** In the next we introduce the integration theory under the product measures, in which we keep the notations in section 6. First we show that the product measure can be represented by integration.

Theorem 9.27 (Product measure). *Let $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ be σ-fintie measure spaces. If $E \in \mathcal{X} \times \mathcal{Y}$, then the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable on X and Y, respectively, and*

$$(\mu \times \nu)(E) = \int_X \nu(E)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y).$$

Proof. First suppose that μ and ν are finite, and let $\mathcal{D}$ be the set of all $E \in \mathcal{X} \times \mathcal{Y}$ for which the conclusions of the theorem are true. If $E = A \times B$, then $\nu(E_x) = \chi_A(x)\nu(B)$ and $\mu(E^y) = \mu(A)\chi_B(y)$, so clearly $E \in \mathcal{D}$. By additivity it follows that finite disjoint unions of rectangles are in $\mathcal{D}$, so by corollary 7.31 lemma 7.32 it suffices to show that $\mathcal{A}$ is a monotone class. If (E_n) is an incresing sequence in $\mathcal{D}$ and $E = \bigcup_{n=1}^\infty E_n$, then the functions $f_n(y) = \mu\left((E_n)^y\right)$ are measurable and increse pointwise to $f(y) = \mu\left(E^y\right)$. Hence f is measurable via proposition 6.18, and by the monotone convergence theorem 9.10,

$$\int \mu(E^y)\, d\nu(y) = \lim_{n\to\infty} \int \mu\left((E_n)^y\right)\, d\nu(y) = \lim_{n\to\infty} \mu \times \nu(E_n) = \mu \times \nu(E).$$

Likewise $\mu \times \nu(E) = \int \nu(E_x) d\mu(x)$, so $E \in \mathcal{D}$. Similarly, if (E_n) is a decreasing sequence in $\mathcal{D}$ and $E = \bigcap_{n=1}^\infty E_n$, then we apply the dominated convergence theorem 9.25 to derive that $E \in \mathcal{D}$, where the function $y \mapsto \mu\left((E_n)^y\right)$ is in $L^1(\nu)$ because $\mu\left((E_n)\right) \le \mu(X) < \infty$ and $\nu(Y) \le \infty$, and the majorant function can be chosen as $y \mapsto \mu\left((E_1)^y\right)$. Thus $\mathcal{F}$ is a monotone class, and the proof is complete for the case of fintie measure spaces.

Finally, if μ and ν are σ-finite, we can write $X \times Y$ as the union of an increasing sequence $(X_j \times Y_j)$ of rectangles of fintie measure. If $E \in \mathcal{X} \times \mathcal{Y}$, the preceding argument applies to $E \cap (X_j \times Y_j)$ for each j to give

$$\mu \times \nu\left(E \cap (X_j \times Y_j)\right) = \int \chi_{X_j}(x)\nu(E_x \cap Y_j)\, d\mu(x) = \int \chi_{Y_j}(y)\mu(E^y \cap X_j)\, d\nu(y),$$

and we apply the monotone convergence theorem 9.10 to yield the desired result. □

Moreover, we have the Fubini-Tonelli theorem.

Theorem 9.28 (Fubini-Tonelli theorem). *Let $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ be σ-fintie measure spaces.*

(1) If $f \in L^+(X \times Y)$, then the functions

$$\varphi(x) = \int_Y f_x\, d\nu \quad \text{and} \quad \psi(y) = \int_X f^y\, d\mu$$

are in $L^+(X)$ and $L^+(Y)$ respectively, and

$$\int_{X\times Y} f\, d(\mu \times \nu) = \int_X \varphi\, d\mu = \int_Y \psi\, d\nu. \tag{9.2}$$

(2) If $f : X \to \mathbb{C}$ and

$$\int_X \varphi^* \, d\mu < \infty, \quad \textit{where} \quad \varphi^*(x) := \int_Y |f|_x \, d\nu,$$

then $f \in L^1(\mu \times \nu)$.

(3) If $f \in L^1(\mu \times \nu)$, then $f_x \in L^1(\nu)$ for μ-a.e. $x \in X$, $f^y \in L^1(\mu)$ for ν-a.e. $y \in Y$, the a.e.-defined functions φ and ψ are in $L^1(\mu)$ and $L^1(\nu)$ respectively, and (9.2) holds.

Proof. The definitions in (1) of φ and ψ make sense by corollary 6.40. Theorem 9.27 implies (1) in the case that $f = \chi_E$ for $E \in \mathcal{X} \times \mathcal{Y}$, and hence (1) holds for all non-negative simple functions. In the general case, let (f_n) be a sequence of simple functions that increse pointwise to f as in theorem 6.22. The monotone convergence theorem 9.10 implies, first, that the corresponding φ_n and ψ_n increse to φ and ψ respectively, and, second that

$$\int_X \varphi \, d\mu = \lim_{n\to\infty} \int_X \varphi_n \, d\mu = \lim_{n\to\infty} \int_{X\times Y} f_n \, d(\mu \times \nu) = \int_{X\times Y} f \, d(\mu \times \nu),$$
$$\int_Y \psi \, d\nu = \lim_{n\to\infty} \int_Y \psi_n \, d\nu = \lim_{n\to\infty} \int_{X\times Y} f_n \, d(\mu \times \nu) = \int_{X\times Y} f \, d(\mu \times \nu),$$

which is (9.2). This establishes (1).

(2) follows by applying (1) to $|f|$.

For (3), by taking the positive and negative parts of real and imaginary parts of f, WLOG we may assume that $f \in L^+ \cap L^1$. Then (3) follows from (1) and proposition 9.9. □

Remark 9.29. Theorem 9.28 can false if one of the measure spaces if not σ-fintie. Consider the following example:

If $X = Y = [0,1]$, μ the Lebesgue measure, ν the counting measure, and $f(x,y) = 1$ for $x = y$ and $f(x,y) = 0$ otherwise, then

$$\int_X f(x,y) \, d\mu(X) = 0 \quad \text{and} \quad \int_Y f(x,y) \, d\nu(y) = 1$$

for all $x, y \in [0,1]$ so that

$$\int_X \left(\int_Y f(x,y) \, d\nu(y) \right) d\mu(x) = 1 \neq 0 = \int_Y \left(\int_X f(x,y) \, d\mu(x) \right) d\nu(y).$$

(The function $f = \chi_{x=y}$ is measurable since $\{x = y\} = \bigcap_{n=1}^{\infty} Q_n$ where $Q_n := \left(\left[\frac{0}{n}, \frac{1}{n}\right] \times \left[\frac{0}{n}, \frac{1}{n}\right]\right) \cup \cdots \cup \left(\left[\frac{n-1}{n}, \frac{n}{n}\right] \times \left[\frac{n-1}{n}, \frac{n}{n}\right]\right)$ is measurable.)

Theorem 9.30 (Fubini-Tonelli theorem for complete measures). *Let $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ be σ-fintie measure spaces, and let $(X \times Y, \mathcal{L}, \lambda)$ be the completion of $(X \times Y, \mathcal{X} \times \mathcal{Y}, \mu \times \nu)$.*

(1) If $f \in L^+(\lambda)$, then f_x is $\mathcal{Y}$-measurable for μ-a.e. $x \in X$, f^y is $\mathcal{X}$-measurable for ν-a.e. $y \in Y$, and the functions

$$\varphi(x) = \int_Y f_x \, d\nu \quad \textit{and} \quad \psi(y) = \int_X f^y \, d\mu$$

are a.e.-defined and measurable on their domains respectively, and

$$\int_{X\times Y} f\, d(\mu\times\nu) = \int_X \varphi\, d\mu = \int_Y \psi\, d\nu. \tag{9.3}$$

(2) If $f : X \to \mathbb{C}$ *and*

$$\int_X \varphi^*\, d\mu < \infty, \quad \text{where} \quad \varphi^*(x) := \int_Y |f|_x\, d\nu,$$

then $f \in L^1(\mu\times\nu)$.

(3) If $f \in L^1(\mu\times\nu)$*, then* $f_x \in L^1(\nu)$ *for* μ*-a.e.* $x \in X$*,* $f^y \in L^1(\mu)$ *for* ν*-a.e.* $y \in Y$*, the a.e.-defined functions* φ *and* ψ *are in* $L^1(\mu)$ *and* $L^1(\nu)$ *respectively, and (9.2) holds.*

Proof. This easily follows from Fubini-Tonelli theorem 9.28 and proposition 7.16. □

Remark 9.31. In this theorem we regard an a.e.-defined function as a function defined on a co-null set E, and we call E the **domain** of it.

9.E. **The transformation of integration, polar coordinates.** In the next, we introduce the transformation of integration. For the general cases, we have the following proposition: (Recall the concept of push-forward in example 7.4.)

Proposition 9.32. *Let* $(X, \mathcal{X}, \mu)$ *and* $(Y, \mathcal{Y}, \nu)$ *be measure spaces, and let* $f : X \to Y$ *be a measurable map. Then* $g\circ f \in L^1(\mu)$ *iff* $g \in L^1(f_*\mu)$*, and*

$$\int_Y g\, d(f_*\mu) = \int_X g\circ f\, d\mu. \tag{9.4}$$

Proof. For $E \in \mathcal{Y}$ and $g = \chi_E$, (9.4) follows from $\chi_E\circ f = \chi_{f^{-1}(E)}$. So (9.4) holds for simple function and hence for L^+ functions by theorem 6.22 and the monotone convergence theorem 9.10. In particular, (9.4) holds for $|g|$ and so $g\circ f \in L^1(\mu)$ iff $g \in L^1(f_*\mu)$. Finally by taking the positive and negative parts of real and imaginary parts of g, (9.4) holds for complex-valued functions. □

In the next we focus on the Lebesgue measure m.

Theorem 9.33. *Suppose that* $T \in GL(n, \mathbb{R})$.

(1) If f *is a Lebesgue measurable function (or Borel measurable) on* $\mathbb{R}^n$*, so is* $f\circ T$*. If* $f \geq 0$ *or* $f \in L^1(m)$*, then*

$$\int f(x)\, dm(x) = |\det T| \int f\circ T(x)\, dm(x). \tag{9.5}$$

(2) If $E \in \mathcal{L}(\mathbb{R}^n)$ *(or* $\mathcal{B}_{\mathbb{R}^n}$*), then* $T(E) \in \mathcal{L}(\mathbb{R}^n)$ *(or* $\mathcal{B}_{\mathbb{R}^n}$*), and* $m(T(E)) = |\det T|\, m(E)$*.*

Proof. First suppose that f is Borel measurable. Then $f\circ T$ is Borel measurable since T is continuous and hence Borel measurable. If (9.5) holds for transformations T and S, it also holds for $T\circ S$ since

$$\begin{aligned}\int f(x)\, dm(x) = |\det T| \int f\circ T(x)\, dm(x) &= |\det T|\, |\det S| \int (f\circ T)\circ S(x)\, dm(x) \\ &= |\det(T\circ S)| \int f\circ(T\circ S)(x)\, dm(x).\end{aligned}$$

Hence it suffices to prove (9.5) when T is of the following three types:

$$\begin{aligned} T_1(x_1,\cdots,x_j,\cdots,x_n) &= (x_1,\cdots,cx_j,\cdots,x_n) \quad (c\neq 0),\\ T_2(x_1,\cdots,x_j,\cdots,x_n) &= (x_1,\cdots,x_j+cx_k,\cdots,x_n) \quad (k\neq j),\\ T_3(x_1,\cdots,x_j,\cdots,x_k,\cdots,x_n) &= (x_1,\cdots,x_k,\cdots,x_j,\cdots,x_n). \end{aligned}$$

This is a simple consequence of Fubini-Tonelli theorem 9.30: for T_3 interchange the order of integration in the variables x_j and x_k, and for T_1 and T_2 we integrate first with respect to x_j and use the one-dimensional formulas

$$\int f(t)\,dm(t) = |c|\int f(ct)\,dm(t), \qquad \int f(t+a)\,dm(t) = \int f(t)\,dm(t).$$

Since $\det T_1 = c$, $\det T_2 = 1$, $\det T_3 = -1$, in this case (9.5) is proved. Moreover, if E is a Borel set, so is $T(E)$ since T^{-1} is continuous and hence Borel measurable. By taking $f=\chi_{T(E)}$, we obtain $m(T(E)) = |\det T|\, m(E)$.

The result for Lebesgue measurable functions and sets now follows from theorem 8.19. □

Corollary 9.34. *Lebesgue measure is invariant under orthogonal transformations.*

Next we shall generalize theorem 9.33 to differentiable maps.

Theorem 9.35 (Transformation formula)**.** *Let $U, V \subset \mathbb{R}^n$ be open and let $f\in C^1(U,V)$ be bijective. If g is a Lebesgue measurable (or Borel measurable) function on V, then $g\circ f$ is Lebesgue measurable (or Borel measurable) on U. If $g\geq 0$ or $g\in L^1(V)$, then we have the* ***transformation formula***

$$\int_U g(f(x))\left|J_f(x)\right|\,dm(x) = \int_V g(y)\,dm(y),$$

where $J_f = \det\left(\frac{\partial f}{\partial x}\right)$ is the Jacobi determinant of f. In particular, for Lebesgue measurable (or Borel measurable) $E\subset U$, $f(E)$ is Lebesgue measurable (or Borel measurable), and

$$m(f(E)) = \int_E \left|J_f(x)\right|\,dm(x)$$

Proof. One can refer to [For] or [Rai]. □

In the next we introduce the integration under the polar coordinates.

Definition 9.36 (Polar coordinates)**.** *Let $S^{n-1} = \{x\in\mathbb{R}^n : |x|=1\}$ denote the unit sphere in $\mathbb{R}^n$. The map*

$$\varphi : \mathbb{R}^n\setminus\{0\} \to (0,\infty)\times S^{n-1}, \qquad x\mapsto \left(|x|, \frac{x}{|x|}\right)$$

defines a diffeomorphism with inverse $(r,y)\mapsto ry$; we call $(r,y)=\varphi(x)$ the ***polar coordinates*** *of x.*

Let ρ be the measure on $(0,\infty)$ defined by $\rho(E) = \int_E r^{n-1}\,dm(r)$.

Theorem 9.37 (Polar coordinates). *There is a unique Borel measure σ on S^{n-1} such that $\varphi_* m = \rho \times \sigma$. If f is Borel measurable on $\mathbb{R}^n$ and $f \geq 0$ or $f \in L^1(m)$, then*

$$\int_{\mathbb{R}^n} f(x)\, dm(x) = \int_{(0,\infty)} \int_{S^{n-1}} f(ry) r^{n-1}\, d\sigma(y)\, dm(r).$$

Proof. By proposition 9.32 and Fubini-Tonelli theorem 9.28, it suffices to show that there is a unique Borel measure σ on S^{n-1} such that $\varphi_* m = \rho \times \sigma$. For Borel sets $E \in S^{n-1}$ we define

$$\sigma(E) := n \cdot m\left(\varphi^{-1}\left((0,1] \times E\right)\right).$$

It suffices to show that σ is a Borel measure σ on S^{n-1} such that $\varphi_* m = \rho \times \sigma$.

Since the map $E \mapsto \varphi^{-1}((0,1] \times E)$ maps Borel sets to Borel sets and commutes with unions, intersections, and complements, σ is a Borel measure on S^{n-1}.

Note the following points:

(1) For $N \in \mathbb{N}$ and a fixed Borel set $E \subset S^{n-1}$, the collection $\mathcal{G}_{N,E}$ of the form $(a,b] \times E$, where $b \leq N$, forms an elementary family. Via theorem 7.29, letting $\mathcal{A}_{N,E}$ be the collection of all finite disjoint unions of sets in $\mathcal{G}$, then $\mathcal{A}_{N,E}$ is an algebra;

(2) Borel rectangles in $(0,\infty) \times S^{n-1}$ are disjoint countable unions of sets in $\bigcup_{N\in\mathbb{N}, E\in\mathcal{B}_{S^{n-1}}} \mathcal{A}_{N,E}$;

(3) The collection $\mathcal{G}$ of rectangles forms an elementary family, and hence the collection $\mathcal{A}$ of finite disjoint unions of rectangles is an algebra via theorem 7.29. Clearly, the σ-algebra generated by $\mathcal{A}$ is $\prod_{j=1}^n \mathcal{X}_j$.

Thus to show that $\varphi_* m = \rho \times \sigma$, it suffices to show that $\varphi_* m = \rho \times \sigma$ holds on $(a,b] \times E$ where $E \subset S^{n-1}$, which follows from theorem 9.33. We are done. □

Remark 9.38. The formula of previous theorem can be extended to Lebesgue measurable functions by considering the completion of σ.

Remark 9.39. In particular, if $f(x) = g(|x|)$, it yields

$$\int_{\mathbb{R}^n} f(x)\, dm(x) = \sigma(S^{n-1}) \int_{(0,\infty)} g(r) r^{n-1}\, dm(r).$$

Example 9.40. We have the following basic examples.

$$\int_{\mathbb{R}^n} e^{-a|x|^2}\, dm(x) = \left(\frac{\pi}{a}\right)^{n/2}, \quad a > 0,$$

$$\sigma(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}, \quad \text{and} \quad m(B^n) = \frac{\pi^{n/2}}{\Gamma(n/2+1)}.$$

9.F. **Riemann integration v.s. Lebesgue intergration.** In the next we introduce the relation between Riemann integration and Lebesgue intergration. One can refer to [Mei] or [Jon] for a detailed introduction to Riemann integration, which is relatively elementary and so we won't give the whole details.

First we introduce the semicontinuity to help us study continuity.

Definition 9.41 (Semicontinuity). *Let X be a topological space, and let $f : X \to [-\infty, +\infty]$. Then f is called **lower semicontinuous** if $\{x : f(x) > a\}$ is open for all $a \in \mathbb{R}$, and f is called **upper semicontinuous** if $\{x : f(x) < a\}$ is open for all $a \in \mathbb{R}$.*

In particular, let $X = \mathbb{R}^n$ *and let* $x \in \mathbb{R}^n$. *Then* f *is called* ***lower semicontinuous at*** x *if for any* $t < f(x)$ *there exists* $\delta > 0$ *such that for all* $y \in B_\delta(x)$, $t < f(y)$*; and* f *is called* ***upper semicontinuous at*** x *if for any* $t > f(x)$ *there exists* $\delta > 0$ *such that for all* $y \in B_\delta(x)$, $t > f(y)$.

We abbreviate lower semicontinuous and upper semicontinuous as LSC and USC respectively.

Remark 9.42. It follows from remark 6.6 that LSC or USC functions are Borel measurable.

Proposition 9.43. *Let* $f : \mathbb{R}^n \to [-\infty, +\infty]$ *and define*

$$\underline{f}(x) = \liminf_{y \to x} f(y), \quad \text{and} \quad \overline{f}(x) = \limsup_{y \to x} f(y).$$

Then

(1) $\underline{f} \le f \le \overline{f}$, *where* $\underline{f}$ *is LSC and* $\overline{f}$ *is USC.*

(2) f *is continuous at* $x \iff f$ *is LSC and USC at* $x \iff \underline{f}(x) = f(x) = \overline{f}(x)$.

(3) f *is LSC* $\iff \underline{f}(x) = f(x)$.

Proof. Easy. One can refer to [Jon]. □

Remark 9.44. See proposition 11.14 for more properties of LSC functions.

Definition 9.45 (Step functions and Riemann integration)**.** *Let* I *be a box in* $\mathbb{R}^n$. *A Lebesgue measurable function* $\sigma : I \to [-\infty, +\infty]$ *is called a* ***step function*** *if there exists an almost disjoint collection* $\{I_j\}_{j=1}^N$ *of subboxes if* I *satisfying* $\bigcup_{j=1}^N I_j = I$ *and that* $\sigma|_{I_j^\circ}$ *is constant for each* j.

Let $f : I \to \mathbb{R}$ *be a bounded function. We all* f ***Riemann integrable*** *if for any* $\varepsilon > 0$, *there exist step functions* σ *and* τ *on* I *such that*

$$\sigma \le f \le \tau, \quad \int_I (\tau - \sigma)\, dm < \varepsilon.$$

If f *is Riemann integrable, then we define*

$$\begin{aligned} (R)\int_I f(x)\,dx &= \sup\left\{\int_I \sigma\,dm : \sigma \le f, \sigma \text{ is a step function}\right\} \\ &= \inf\left\{\int_I \tau\,dm : f \le \tau, \tau \text{ is a step function}\right\}. \end{aligned}$$

Remark 9.46. For any step function τ we have

$$(R)\int_I \tau(x)\,dx = \int_I \tau\,dm.$$

Theorem 9.47 (Riemann integration)**.** *Let* f *be a bounded function on a box* I. *Then*

(1) f *is Riemann integrable* $\iff f$ *is continuous a.e.*

(2) If f is Riemann integrable, then f is measurable and its Riemann integration and Lebesgue integration on I are equal:

$$(R)\int_I f(x)\,dx = \int_I f\,dm.$$

Proof. (1) is well-known. One can refer to [Mei] or [Jon]. In the next we give a direct proof of (2).

Suppose that f is Riemann integrable. Then there exist two sequences of step functions $(\sigma_j)_{j=1}^{\infty}$ and $(\tau_j)_{j=1}^{\infty}$ on I such that

$$\sigma_j \le f \le \tau_j, \qquad \int_I (\tau_j - \sigma_j)\,dm < \frac{1}{j}, \qquad \forall j.$$

It follows that $\underline{\sigma_j} \le \underline{f}$ and $\overline{f} \le \overline{\tau_j}$, and hence via proposition 9.43 (2) we have

$$\sigma_j \le \underline{f} \le f \le \overline{f} \le \tau_j, \quad a.e.$$

since σ_j and τ_j are continuous a.e. Setting $g = \sup_{j\ge 1}\sigma_j$ and $h = \inf_{j\ge 1}\tau_j$ then g and h are Lebesgue measurable with

$$g \le \underline{f} \le f \le \overline{f} \le h, \quad a.e.$$

Note that

$$\forall j : \int_I (h-g)\,dm \le \int_I (\tau_j - \sigma_j) dm < \frac{1}{j} \implies \int_I g\,dm = \int_I h\,dm.$$

It follows form proposition 9.23 (6) that

$$g = \underline{f} = f = \overline{f} = h, \quad a.e.$$

(This also implies f is continuous a.e. via proposition 9.43 (2).) Since $f = h$ a.e. and h is Lebesgue measurable, via proposition 7.15, f is also Lebesgue measurable. Note that

$$(R)\int_I f(x)\,dx \le \int_I \tau_j\,dm < \int_I \sigma_j\,dm + \frac{1}{j} \le \int_I f\,dm + \frac{1}{j}, \quad \forall j;$$
$$(R)\int_I f(x)\,dx \ge \int_I \sigma_j\,dm > \int_I \tau_j\,dm - \frac{1}{j} \ge \int_I f\,dm - \frac{1}{j}, \quad \forall j.$$

The desired result follows. □

For more properties of Riemann integration, one can refer to [Jon].

10. Signed measures and differentiation

The principal theme of this section is the concept of differentiating a measure ν with respect to another measure μ on the same σ-algebra.

As pointed out in [For], in developing the theory of differentiation, it is useful to generalize the concept of measures so as to allow measures to assume negative or even complex values. There are three reasons for this.

(1) First, in applications such "signed measures" can represent things such as electric charge that can be either positive or negative.
(2) Second, the differentiation theory proceeds more naturally in the more general seting.
(3) Finally, complex measures have a functional-analytic significance in section 11.

10.A. Signed measure.

Definition 10.1 (Signed measure). *Let $(X, \mathcal{X})$ be a measurable space. A **signed measure** on $(X, \mathcal{X})$ is a function $\nu : \mathcal{X} \to [-\infty, +\infty]$ such that*

(1) $\nu(\varnothing) = 0$;
(2) ν assumes at most one of the values $\pm\infty$;
(3) If (E_j) is a sequence of disjoint sets in $\mathcal{X}$, then $\nu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \nu(E_j)$, where the latter sum converges absolutely if $\nu\left(\bigcup_{j=1}^{\infty} E_j\right)$ is finite.

Definition 10.2. *Every measure is a signed measure; for emphasis we shall sometimes refer to measures as **positive measures**.*

*If ν is a signed measure on $(X, \mathcal{X})$, a set $E \in \mathcal{X}$ is called **positive** (resp. **negative**, **null**) for ν if $\nu(F) \geq 0$ (resp. $\nu(F) \leq 0$, $\nu(F) = 0$) for all $F \in \mathcal{X}$ such that $F \subset E$.*

*Two signed measures μ and ν on $(X, \mathcal{X})$ are called **mutually singular**, or that ν is **singular with respect to** μ, or vice versa, if there exist $E, F \in \mathcal{X}$ such that $E \cap F = \varnothing$, $E \cup F = X$, E is null for μ, and F is null for ν, and we denote this relation by $\mu \perp \nu$.*

Remark 10.3. Informally speaking, mutual singularity means that μ and ν live on disjoint sets.

The following are some basic properties of signed measures.

Proposition 10.4. *Let ν be a signed measure on a measurable space $(X, \mathcal{X})$.*

(1) If (E_j) is an incresing sequence in $\mathcal{X}$, then $\nu\left(\bigcup_{j=1}^{\infty} E_j\right) = \lim_{j\to\infty} \nu(E_j)$.
(2) If (E_j) is a decreasing sequence in $\mathcal{X}$ and $\nu(E_1)$ is finite, then $\nu\left(\bigcap_{j=1}^{\infty} E_j\right) = \lim_{j\to\infty} \nu(E_j)$.
(3) Any measurable subset of a positive set is positive.
(4) The union of any countable family of positive sets is positive.

Proof. Trivial. □

Example 10.5. Let $(X, \mathcal{X})$ be a measurable space.

(1) If μ_1 and μ_2 are measures on $\mathcal{X}$, and at least one of them is fintie, then $\nu = \mu_1 - \mu_2$ is a signed measure.

(2) If μ is a measure on $\mathcal{X}$ and $f : \mathcal{X} \to [-\infty, +\infty]$ is a measurable function such that at least one of $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ is finite (in which case we shall call f an **extended μ-integrable** function), then $\nu : \mathcal{X} \to [-\infty, +\infty]$, $E \mapsto \int_E f \, d\mu$ is a signed measure.

We will see that every signed measure can be represented in either of these two forms. See the next subsection.

10.B. **Decomposition theorems, standard integration representation.** Now we prove the above claims.

Theorem 10.6 (Decomposition theorems)**.** *There are two editions of decomposition theorems.*

*(1) (**The Hahn decomposition theorem**) If ν is a signed measure on a measurable space $(X, \mathcal{X})$, there exist a positive set P and a negative set N for ν such that $P \cup N = X$ and $P \cap N = \varnothing$. If P' and N' an another such pair, then $P \Delta P' = N \Delta N'$ is null for ν.*

*(2) (**The Jordan decomposition theorem**) If ν is a signed measure on a measurable space $(X, \mathcal{X})$, there exist unique positive measures ν^+ and ν^- such that $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$.*

Proof. WLOG, we assume that ν does not assume the value $+\infty$. (Otherwise, consider $-\nu$.) Set

$$M = \sup\{\nu(E) : E \text{ is positive}\}.$$

Thus there is a sequence (P_j) of positive sets such that $\nu(P_j) \to M$. Set

$$P = \bigcup_{j=1}^{\infty} P_j \quad \text{and} \quad N = X \setminus P.$$

Then P is positive via proposition 10.4. In the next we show that N is negative.

N has the following basic properties: If $E \subset N$ is positive and $\nu(E) > 0$, then $E \cup P$ is positive and $\nu(E \cup P) = \nu(E) + \nu(P) > M$, which is impossible; it follows that if $A \subset N$ and $\nu(A) > 0$, there exists $B \subset A$ with $\nu(B) > \nu(A)$. Indeed, since A cannot be positive, there exists $C \subset A$ with $\nu(C) < 0$; thus if $B = A \setminus C$, we have $\nu(B) = \nu(A) - \nu(C) > \nu(A)$.

Now we prove that N is negative. Suppose for contradiction, and then via the properties above we can specify a sequence of subsets (A_j) of N and a sequence (n_j) of positive integers as follows: n_1 is the smallest integer for which there exist a set $B \subset N$ with $\nu(B) > n_1^{-1}$, and A_1 is such a set. Proceeding inductively, n_j is the smallest integer for which there exists a set $B \subset A_{j-1}$ with $\nu(B) > \nu(A_{j-1}) + n_j^{-1}$, and A_j is such a set. Setting $A = \bigcap_{j=1}^{\infty}$, then

$$+\infty > \nu(A) = \lim_{j \to \infty} \nu(A_j) > \sum_{j=1}^{\infty} n_j^{-1},$$

so $n_j \to \infty$ as $j \to \infty$. But once again, there exists $B \subset A$ with $\nu(B) > \nu(A) + n^{-1}$ for some integer n. For j sufficiently large we have $n < n_j$, and $B \subset A_{j-1}$, which contradicts the construction of n_j and A_j. Thus the assumption that N is not negative is untenable.

Moreover, if P', N' is another pair of sets as in the statement of the theorem, we have $P \subset P' \subset P$ and $P \subset P' \subset N'$, so that $P \setminus P'$ is both positive and negative, hence null; likewise for $P' \setminus P$. Then (1) follows.

For (2). Let $X = P \cup N$ be a Hahn decomposition as above, and define

$$\nu^+(E) = \nu(E \cap P) \quad \text{and} \quad \nu^-(E) = -\nu(E \cap N).$$

Then clearly $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$. If also $\nu = \mu^+ - \mu^-$ and $\mu^+ \perp \mu^-$, let $E, F \in \mathcal{X}$ be such that $E \cap F = \varnothing$, $E \cup F = X$, and $\mu^+(F) = \mu^-(E) = 0$. Then $X = E \cup F$ is another Hahn decomposition for ν, so $P \Delta E$ is a ν-null set. Therefore, for any $A \in \mathcal{X}$, $\mu^+(A) = \mu^+(A \cap E) = \nu(A \cap E) = \nu(A \cap P) = \nu^+(A)$, and likewise $\nu^- = \mu^-$. □

Corollary 10.7 (Standard integration representation)**.** *If ν is a signed measure on a measurable space $(X, \mathcal{X})$, then*

$$\nu(E) = \int_E f \, d\mu,$$

where $\mu = \nu^+ + \nu^-$ and $f = \chi_P - \chi_N$, $X = P \cup N$ being a Hahn decomposition for ν.

10.C. **Variation, integration, derivative, absolutely continuous.** In the next we introduce the following concepts, which are vital in differentiation theory.

Definition 10.8 (Variation, integration)**.** *The measures ν^+ and ν^- in theorem 10.6 are called the **positive** and **negative variations** of ν, and $\nu = \nu^+ - \nu^-$ is called the **Jordan decomposition** of ν.*

*Furthermore, we define the **total variation** of ν to be the measure $|\nu|$ defined by $|\nu| = \nu^+ + \nu^-$.*

*A signed measure ν is called **finite** (resp. σ-**finite**) if $|\nu|$ is finite (resp. σ-finite).*

*Moreover, **integration** with respect to a signed measure ν is defined in the obvious way:*

$$L^1(\nu) := L^1(\nu^+) \cap L^1(\nu^-); \qquad \int f \, d\nu := \int f \, d\nu^+ - \int f \, d\nu^-, \quad \forall f \in L^1(\nu).$$

Remark 10.9. It's clear that any $E \in \mathcal{X}$ is ν-null iff $|\nu|(E) = 0$, and $\nu \perp \mu$ iff $|\nu| \perp \mu$ iff $\nu^+ \perp \mu$ and $\nu^- \perp \mu$.

Definition 10.10 (Derivative)**.** *Let μ, ν be two signed measures on a measurable space $(X, \mathcal{X})$. A μ-measurable function $f : X \to [-\infty, +\infty]$ is called the **derivative** of ν with respect to ν, if we have*

$$\nu(E) = \int_E f \, d\mu, \quad \forall E \in \mathcal{X},$$

which is denoted by $d\nu = f \, d\mu$.

Definition 10.11 (Absolutely continuous)**.** *Let $(X, \mathcal{X})$ be a measurable space, let ν be a signed measure on $\mathcal{X}$, and let μ be a positive measure on $\mathcal{X}$. We say that ν is **absolutely continuous** with respect to μ and write*

$$\nu \ll \mu$$

if $\nu(E) = 0$ for every $E \in \mathcal{X}$ for which $\mu(E) = 0$.

Remark 10.12. It's clear that any $\nu \ll \mu$ iff $|\nu| \ll \mu$ iff $\nu^+ \ll \mu$ and $\nu^- \ll \mu$.

One can extend the notion of absolute continuity to the case where μ is a signed-measure, but we shall have no need of this more general definition.

Remark 10.13. The term "absolute continuity" has the background of real-variable theory. Actually the names of some other above notations come from the differentiation theory on $\mathbb{R}$ as well.

Example 10.14. Let $(X, \mathcal{X})$ be a measurable space.

(1) If μ is a measure on $\mathcal{X}$ and f is an extended μ-integrable function, the signed measure ν defined by $\nu(E) = \int_E f\, d\mu$ is clearly *absolutely continuous* with respect to μ.
(2) If $\mu_1, \cdots, \mu_n$ are measure on $\mathcal{X}$, then $\mu_j \ll \sum_{j=1}^n \mu_j$ for all j.

Proposition 10.15. *Let $(X, \mathcal{X})$ be a measurable space, let ν be a signed measure on $\mathcal{X}$, and let μ be a positive measure on $\mathcal{X}$.*

(1) (Absolute continuity is in a sense the antithesis of mutual singularity) *If $\nu \perp \mu$ and $\nu \ll \mu$, then $\nu = 0$.*
(2) $\nu \ll \mu$ iff for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|\nu(E)| < \varepsilon$ whenever $\mu(E) < \delta$.

Proof. Trivial. □

Corollary 10.16. *Let μ be a measure on a measurable space $(X, \mathcal{X})$. If $f \in L^1(\mu)$, then for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\left|\int_E f\, d\mu\right| < \varepsilon$ whenever $\mu(E) < \delta$.*

Proof. Apply proposition 10.15 (2) to $\Re f$ and $\Im f$. □

10.D. **Radon-Nikodym-Lebesgue theorem.** Now we come back to the differentiation theory. First, we give an intuitive and precise characterization of the antithesis of being mutually singular in a special case.

Lemma 10.17. *Suppose that ν and μ are finite measures on a measurable space $(X, \mathcal{X})$. Either $\nu \perp \mu$, or there exist $\varepsilon > 0$ and $E \in \mathcal{X}$ such that $\mu(E) > 0$ and E is a positive set for $\nu - \varepsilon\mu$.*

Proof. Let $X = P_n \cup N_n$ be a Hahn decomposition for $\nu - n^{-1}\mu$, and let $P = \bigcup_{j=1}^\infty P_j$ and $N = \bigcap_{j=1}^\infty N_j = P^c$. Then N is a negative set for $\nu - n^{-1}\mu$ for all n, i.e., $0 \le \nu(N) \le n^{-1}\nu(N)$ for all n, so $\nu(N) = 0$. If $\mu(P) = 0$, then $\nu \perp \mu$. If $\mu(P) > 0$, then $\mu(P_n) > 0$ for some n, and P_n is a positive set for $\nu - n^{-1}\mu$. □

Theorem 10.18 (The Radon-Nikodym-Lebesgue Theorem)**.** *Let ν be a σ-finite signed measure and let μ be a σ-finite positive measure on a measurable space $(X, \mathcal{X})$. There exist unique σ-finite signed measures λ, ρ on $\mathcal{X}$ such that*

$$\nu = \lambda + \rho, \quad \lambda \perp \mu, \quad \rho \ll \mu$$

Moreover, there is an extended μ-integrable function $f : X \to \mathbb{R}$ such that $d\rho = f\, d\mu$, and any such two functions are equal μ-a.e.

Proof. Case 1: Suppose that ν and μ are finite positive measures. Our framework is that we first find f and then define $d\lambda = d\nu - f\,d\mu$. Set

$$\mathcal{F} = \left\{ f : X \to [0, +\infty] : \int_E f\,d\mu \le \nu(E),\ \forall E \in \mathcal{X} \right\}.$$

It's clear that $\mathcal{F}$ is nonempty, and if $f, g \in \mathcal{F}$, then $\max\{f, g\} \in \mathcal{F}$. Hence

$$M := \sup\left\{ \int f\,d\mu : f \in \mathcal{F} \right\} \le \nu(X) < \infty.$$

Choose a sequence $(f_n) \subset \mathcal{F}$ such that $\int f_n\,d\mu \to M$. Setting $g_n = \max\{f_1, \cdots, f_n\}$ and $f = \sup_{n\ge1} f_n$, then $g_n \in \mathcal{F}$, g_n increses pointwise to f, and $\int g_n\,d\mu \ge \int f_n\,d\mu$. It follows that $\lim_{n\to\infty} \int g_n\,d\mu \to M$ and hence, by the monotone convergence theory 9.10, $f \in \mathcal{F}$ and $\int f\,d\mu = M$.

We calim that the measure $d\lambda = d\nu - f\,d\mu$, which is positive since $f \in \mathcal{F}$, is singular with respect to μ. If not, by lemma 10.17, there exist $E \in \mathcal{X}$ and $\varepsilon > 0$ such that $\mu(E) > 0$ and E is a positive set for $\lambda - \varepsilon\mu$. But then

$$\varepsilon\chi_E\,d\mu \le d\lambda = d\nu - f\,d\mu \implies f + \varepsilon\chi_E \in \mathcal{F},\ \text{where} \int (f + \varepsilon\chi_E)\,d\mu = M + \varepsilon\mu(E) > M.$$

This contradicts the definition of M.

Thus the existence of λ, f and $d\rho = f\,d\mu$ is proved. As for the uniqueness, if also $d\nu = d\lambda_1 + f_1 d\mu$, we have $d\lambda - d\lambda_1 = (f_1 - f)\,d\mu$ and hence

$$(\lambda - \lambda_1) \perp \mu \quad \text{and} \quad (\lambda - \lambda_1) \ll \mu$$

By proposition 10.15 (1) we know $\lambda = \lambda_1$, and hence $f_1 = f$ μ-a.e.

Case 2: Suppose that ν and μ are σ-finite positive measures. Then X is a countable disjoint union of μ-finite sets and a countable disjoint union of ν-disjoint sets; by taking intersections of these we obtain a disjoint sequence $(A_j) \subset \mathcal{X}$ such that $\mu(A_j)$ and $\nu(A_j)$ are finite for all j and $X = \bigcup_{j=1}^\infty A_j$. Set

$$\mu_j(E) = \mu(E \cap A_j) \quad \text{and} \quad \nu_j(E) = \nu(E \cap A_j).$$

By the reasoning above, for each j we have $d\nu_j = d\lambda_j + f_j\,d\mu_j$ where $\lambda_j \perp \mu_j$. Obviously, WLOG we may assume that $f_j = 0$ on A_j^c. Setting $\lambda = \sum_{j=1}^\infty \lambda_j$ and $f = \sum_{j=1}^\infty f_j$, it's clear that $d\nu = d\lambda + f\,d\mu$, $\lambda \perp \mu$, and $d\lambda$ and $f\,d\mu$ are σ-finite, as desired. Uniqueness follows as before.

The General Case: If ν is a signed measure, we apply the preceding argument to ν^+ and ν^- and subtract the results. $\square$

10.E. **Lebesgue decomposition, Radon-Nikodym derivative, Chain rule.** Based on Radon-Nikodym-Lebesgue theorem 10.18, we give the following new definitions.

Definition 10.19 (Lebesgue decomposition and Radon-Nikodym derivative)**.** *The decomposition $\nu = \lambda + \rho$ where $\lambda \perp \mu$ and $\rho \ll \mu$ in the Radon-Nikodym-Lebesgue theorem 10.18 is called the **Lebesgue decomposition** of ν with respect to μ.*

*In the case where $\nu \ll \mu$, Radon-Nikodym-Lebesgue theorem 10.18 says that $d\mu = f\,d\mu$ for some f, where f is called the **Radon-Nikodym derivative** of ν with respect to μ. We* *denote it by $d\nu/d\mu$:*

$$d\nu = \frac{d\nu}{d\mu}\, d\mu.$$

Proposition 10.20 (Chain rules). *Suppose that ν is a σ-finite signed measure and μ, λ are σ-finite measures on $(X, \mathcal{X})$ such that $\nu \ll \mu$ and $\mu \ll \lambda$.*

(1) If $g \in L^1(\nu)$, then $g(d\nu/d\mu) \in L^1(\mu)$ and

$$\int g\, d\nu = \int g \frac{d\nu}{d\mu}\, d\mu.$$

(2) We have $\nu \ll \lambda$ and

$$\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda} \quad \lambda - a.e.$$

Proof. Trivial. □

10.F. **Complex measure, Radon-Nikodym-Lebesgue theorem, total variation.** In the next we generalize the conclusions to $\mathbb{C}$.

Definition 10.21 (Complex measure). *A* ***complex measure*** *on a measurable space $(X, \mathcal{X})$ is a map $\nu : M \to \mathbb{C}$ such that*

(1) $\nu(\emptyset) = 0$;

(2) If (E_j) is a sequence of disjoint sets in $\mathcal{X}$, then $\nu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \nu(E_j)$, where the series converges absolutely.

Definition 10.22. *Let ν be a complex measure on a measurable space $(X, \mathcal{X})$. We shall write ν_r and ν_i for the* ***real and imaginary parts of*** *ν, and then we set*

$$L^1(\nu) := L^1(\nu_r) \cap L^1(\nu_i), \qquad \int f\, d\nu := \int f\, d\nu_r + \int f\, d\nu_i, \quad \forall f \in L^1(\nu).$$

If λ and μ are complex measures, we say that $\nu \perp \mu$ if $\nu_a \perp \mu_b$ for $a, b = r, i$, and if λ is a positive measure, we say that $\nu \ll \lambda$ if $\nu_r \ll \lambda$ and $\nu_i \ll \lambda$.

Remark 10.23. If ν is a complex measure, then ν_r and ν_i are finite signed measures, and hence the range of a complex measure is a bounded subset of $\mathbb{C}$.

One has merely to apply the preceding theorems to the real and imaginary parts separately to generalize them to complex measures. In particular:

Theorem 10.24 (The Radon-Nikodym-Lebesgue Theorem). *Let ν be a complex measure and let μ be a σ-finite positive measure on a measurable space $(X, \mathcal{X})$. There exist a complex measure λ and an $f \in L^1(\mu)$ such that $\lambda \perp \mu$ and $d\nu = d\lambda + f\, d\mu$. If also $\lambda_1 \perp \mu$ and $d\nu = d\lambda_1 + f_1\, d\mu$, then $\lambda = \lambda_1$ and $f = f_1$ μ-a.e.*

Based on this we can show that the total variation of a complex measure is well-defined.

Definition 10.25 (Total variation). *The* ***total variation*** *of a complex measure ν is the positive measure $|\nu|$ determined by the property that if $d\nu = f\, d\mu$ where μ is a positive measure, then $d|\nu| = |f|\, d\mu$.*

Proposition 10.26. *Let ν be a complex measure on a measurable space $(X, \mathcal{X})$.*
(1) Setting $\mu = |\nu_r| + |\nu_i|$, then $d\nu = f\, d\mu$ for some $f \in L^1(\mu)$.
(2) Definition 10.25 is well-defined.
(3) Definition 10.25 agrees with the previous definition of ν when ν is a signed measure.

Proof. (1) follows from Radon-Nikodym-Lebesgue theorem 10.24. To show (2), it suffices to show that $|\nu|$ is independent of the choice of μ and f, which easily follows from chain rules 10.20. Finally, (3) follows from decomposition theorems 10.6. These are easy, one can refer to [For] for details. □

For more properties of complex measures, one can refer to [For].

10.G. **Pointwise Differentiation on $\mathbb{R}^n$, coincidence of two derivatives.** In the next, we come back to the differentiation on $\mathbb{R}^n$.

Definition 10.27 (Pointwise Differentiation on $\mathbb{R}^n$). *Let ν be a signed or complex Borel measure on $\mathbb{R}^n$. Then the **pointwise derivative** of ν with respect to m is defined by*

$$F(x) = \lim_{r\to 0} \frac{\nu(B(x,r))}{m(B(x,r))},$$

when it exists.

Remark 10.28. We can change balls into any family of sets that *shrink nicely*. See the Lebesgue differentiation theorem 10.36.

Remark 10.29. In the next when we apply the Radon-Nikodym-Lebesgue theorem 10.24, then m denotes the measure $m|_{\mathcal{B}_{\mathbb{R}^n}}$.

We have the natural question:

Question 10.30. *For the nice case that $\nu \ll m$ where ν is a complex Borel measure on $\mathbb{R}^n$, by Radon-Nikodym-Lebesgue theorem 10.24 we have $d\nu = f\, dm$ for some $f \in L^1(m)$. It follows that the pointwise derivative satisfies:*

$$F(x) = \lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f\, dm.$$

Then people will think, is it true that $F = f$ almost everywhere (with respect to m)?

Remark 10.31. From the point of view of the function f, this may be regarded as a generalization of the fundamental theorem of calculus: the derivative of the indefinite integration (namely, ν) is f. (Actually f is the derivative of ν with respect to m.)

The answer is yes! We summarize the (stronger) conclusions as follows. (We will use the technical lemma corollary 11.10 in section 11.)

Theorem 10.32. *Let $f \in L^1_{loc}$, then*

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)\, dm(y) = f(x), \quad \text{for } m\text{-a.e. } x.$$

Moreover,

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dm(y) = 0, \quad \text{for } m\text{-a.e. } x.$$

Proof. It suffices to show that for $N \in N$, $A_r f(x) \to f(x)$ for m-a.e. x with $|x| \leq N$. But for $|x| \leq N$ and $r \leq 1$ the values $A_r f(x)$ depend only on the values $f(y)$ for $|y| \leq N+1$, so by replacing f with $f\chi_{B(0,N+1)}$ we may assume that $f \in L^1$.

Via corollary 11.10 we can find a continuous integrable function g such that

$$\int |g(y) - f(y)|\, dm(y) < \varepsilon. \tag{10.1}$$

Continuity of g implies that

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} g(y)\, dm(y) = g(x), \quad \forall x \in \mathbb{R}^n.$$

Therefore

$$\begin{aligned} & \limsup_{r\to 0} \left| \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)\, dm(y) - f(x) \right| \\ \leq\ & \limsup_{r\to 0} \left| \frac{1}{m(B(x,r))} \int_{B(x,r)} [f(y) - g(y)]\ dm(y) \right| + |f-g|(x) \\ \leq\ & \limsup_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - g(y)|\ dm(y) + |f-g|(x). \end{aligned}$$

Hence, if we set

$$\begin{aligned} A_\alpha &:= \left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} \left| \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)\, dm(y) - f(x) \right| > \alpha \right\}; \\ B_{\alpha,g} &:= \left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - g(y)|\ dm(y) > \alpha \right\}; \\ C_{\alpha,g} &:= \{x \in \mathbb{R}^n : |f-g|(x) > \alpha\}. \end{aligned}$$

Then

$$A_\alpha \subset B_{\alpha/2,g} \cup C_{\alpha/2,g}, \quad \text{and} \quad A_0 = \bigcup_{n=1}^{\infty} A_{1/n},$$

and hence it suffices to show that $m(A_{1/n}) = 0$ for each n. Note that

$$m\left(C_{\alpha,g}\right) \leq \frac{1}{\alpha} \int_{C_{\alpha,g}} |f-g|\, dm \leq \frac{\varepsilon}{\alpha}.$$

It suffices to show that $m\left(B_{\alpha,g}\right) = O(\varepsilon)$, which is derived in the following lemma. Hence the first assertion follows.

For the second assertion, let D be a countable dense subset of $\mathbb{C}$, and for each $d \in D$ we set

$$E_d = \left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} \left| \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - d|\, dm(y) - |f(x) - d| \right| > 0 \right\}.$$

By the first assertion it follows that $m(E_d) = 0$ for each $d \in D$, and hence $E := \bigcup_{d\in D} E_d$ also satisfies $m(E) = 0$. Then if $x \notin E$, for any $\varepsilon > 0$, we can choose $d \in D$ with

$|f(x) - d| < \varepsilon$; then

$$\begin{aligned}
&\limsup_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\, dm(y) \\
\leq\ &\limsup_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - d|\, dm(y) + \varepsilon \\
\leq\ &|f(x) - d| + \varepsilon < 2\varepsilon.
\end{aligned}$$

Since ε is arbitrary, the second assertion follows. □

10.H. **Hardy-Littlewood maximal function, Lebesgue differentiation theorem.** Moreover, via using the technical lemma, Vitali's covering theorem 11.33, we will get the Lebesgue differentiation theorem.

Lemma 10.33. *For every $f \in L^1_{loc}$, we define the **Hardy-Littlewood maximal function** Hf by*

$$Hf(x) = \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y)|\, dm(y).$$

Then Hf is Borel measurable, and there is a constant $C > 0$ such that for all $f \in L^1$ and all $\alpha > 0$, we have

$$m(\{x \in \mathbb{R}^n : Hf(x) > \alpha\}) \leq \frac{C}{\alpha} \int |f(x)|\, dm(x). \tag{10.2}$$

Proof. The first assertion is easy, and one can refer to [For] for details. In the next, we prove the second assertion (10.2). Set

$$E = \{x \in \mathbb{R}^n : Hf(x) > \alpha\}.$$

Note that

$$\begin{aligned}
x \in E &\iff \exists r_x > 0 : \frac{1}{m(B(x,r_x))} \int_{B(x,r_x)} |f(y)|\, dm(y) > \alpha \\
&\iff \exists r_x > 0 : m(B(x,r_x)) < \frac{1}{\alpha} \int_{B(x,r_x)} |f(y)|\, dm(y)
\end{aligned}$$

Then the desired result easily follows from Vitali's covering theorem 11.33. □

Remark 10.34. In fact, this lemma is stronger than what we need, since we change "$\limsup_{r\to 0}$" to "$\sup_{r>0}$". For the general case, the corresponding proposition is lemma 11.35.

Definition 10.35 (Shrink nicely)**.** *A family $(E_r)_{r>0}$ of Borel subsets of $\mathbb{R}^n$ is said to **shrink nicely** to $x \in \mathbb{R}^n$ if*

(1) $E_r \subset B(x,r)$ for each r;

(2) There is a constant $\alpha > 0$, independent of r, such that $m(E_r) > \alpha m(B(x,r))$.

Theorem 10.36 (The Lebesgue differentiation theorem)**.** *Suppose that $f \in L^1_{loc}$. Then*

$$\lim_{r\to 0} \frac{1}{m(E_r)} \int_{E_r} |f(y) - f(x)|\, dm(y) = 0, \quad \textit{for } m\textit{-a.e. } x.$$

for every family $(E_r)_{r>0}$ that shrinks nicely to x.

Proof. For some $\alpha > 0$ we have

$$\begin{aligned}\frac{1}{m(E_r)}\int_{E_r}|f(y)-f(x)|\,dm(y) &\leq \frac{1}{E_r}\int_{B(x,r)}|f(y)-f(x)|\,dm(y)\\ &\leq \frac{1}{\alpha m(B(x,r))}\int_{B(x,r)}|f(y)-f(x)|\,dm(y).\end{aligned}$$

Then the conclusion follows from theorem 10.32. □

For more conclusions, such as the differentiation theory on $\mathbb{R}$, one can refer to [For].

10.I. **A New framework of derivatives.** We will see in the next section that we can build a new framework of derivatives; that is, a framework of differentiating Radon measures with respect to Radon measures.

The key points are theorem 11.36, theorem 11.37 and theorem 11.38. Now we show how these theorems imply the classical conclusions. In the next we admit all conclusions in section 11. New readers can skip this subsection first

Proposition 10.37. *Every complex Borel measure ν on $\mathbb{R}^n$ has a Randon-Nikodym-Lebesgue representation*

$$d\nu = d\lambda + f\,dm, \tag{10.3}$$

where $\lambda \perp m$ and $f \in L^1(m)$. Then, we have

$$f(x) = \lim_{r\to 0}\frac{\nu(B(x,r))}{m(B(x,r))} = \lim_{r\to 0}\frac{1}{m(B(x,r))}\int_{B(x,r)} f\,dm, \quad \text{for } m\text{-a.e. } x, \tag{10.4}$$

and

$$\lim_{r\to 0}\frac{\lambda(B(x,r))}{m(B(x,r))} = 0 \quad \text{for } m\text{-a.e. } x. \tag{10.5}$$

Moreover, for any $g \in L^1(m)$, we have

$$\lim_{r\to 0}\frac{1}{m(B(x,r))}\int_{B(x,r)} g\,dm = g(x), \quad \text{for } m\text{-a.e. } x. \tag{10.6}$$

Proof. By remark 11.26 we know that every complex Borel measure ν on $\mathbb{R}^n$ is a complex Radon measure. Applying theorem 11.36 and theorem 11.37 to the positive and negative parts of real and imaginary parts of ν, we find the Randon-Nikodym-Lebesgue representation 10.3, and then formulas (10.4) and (10.5) easily follows.

Since $g\,dm$ is Radon measure (which is absolutely continuous with respect to m), formula (10.6) follows from (10.4). □

Lemma 10.38. *If $f \in L^+(\mathbb{R}^n)$, then $f\,dm$ is Radon iff $f \in L^1_{loc}(m)$, where*

$$L^1_{loc}(m) := \left\{ f : \mathbb{R}^n \to \mathbb{C} : \int_K |f|\,dm < \infty \text{ for all bounded measurable set } K \subset \mathbb{R}^n \right\}.$$

Moreover, for a signed measure ν on $\mathbb{R}^n$ with

$$d\nu = d\lambda + f\,dm$$

where $\lambda \perp m$ and f is an extended m-integrable function, ν is Radon iff $f \in L^1_{loc}$.

Proposition 10.39. *If $f \in L^1_{loc}$ and $d\nu = d\lambda + f\,dm$ where $\lambda \perp m$, we have*

$$f(x) = \lim_{r\to 0} \frac{\nu(B(x,r))}{m(B(x,r))} = \lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f\,dm, \quad \text{for } m\text{-a.e. } x,$$

and

$$\lim_{r\to 0} \frac{\lambda(B(x,r))}{m(B(x,r))} = 0 \quad \text{for } m\text{-a.e. } x.$$

Moreover, if $g \in L^1_{loc}(m)$, then

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} g\,dm = g(x), \quad \text{for } m\text{-a.e. } x.$$

Proof. Note that ν is Radon via the lemma 10.38. Similarly, by applying theorem 11.36 and theorem 11.37 to ν, the results easily follow. ☐

Corollary 10.40 (Coincidence of two derivatives)**.** *If ν is a signed or complex Randon measure with $\nu \ll m$, then the derivative of ν with respect to m, which exists via Randon-Nikodym-Lebesgue theorem 10.18 or 10.24, equals to the pointwise derivative of ν with respect to m for m-a.e. x.*

11. Radon measures

The subject of this section is Radon measures and integration theory on LCH spaces. A great significance of Radon measure is that we can approximate functions by continuous functions in Sobolev spaces.

11.A. LCH space, Radon measure, positive linear functionals on $C_c(X, \mathbb{C})$.

Definition 11.1 (LCH space). *We call X an **LCH space** if it is locally compact and Hausdorff.*

Definition 11.2 (Radon measure). *Let X be an LCH space. A Randon measure on X is a Borel measure that is finite on all compact sets, outer regular on all Borel sets, and inner regular on all open sets.*

Remark 11.3. Randon measures are also inner regular on σ-finite sets. See theorem 11.8.

Definition 11.4 (Positive linear functionals on $C_c(X, \mathbb{C})$). *Let $C_c(X, \mathbb{C})$ be the space of complex-valued continuous functions on X with compact support. A linear functional I on $C_c(X, \mathbb{C})$ will be called **positive** if $I(f) \geq 0$ whenever $f \geq 0$.*

Definition 11.5. *If U is open in X and $f \in C_c(X, \mathbb{C})$, we shall write $f \prec U$ to mean that $0 \leq f \leq 1$ and $\operatorname{supp} f \subset U$.*

11.B. Riesz Representation Theorem.

As we said in remark 9.5, we will establish the relation between Radon measures and functionals.

Theorem 11.6 (Riesz Representation Theorem). *Let X be an LCH space. If I is a positive linear functional on $C_c(X, \mathbb{C})$, there is a unique Radon measure μ on X such that $I(f) = \int f \, d\mu$ for all $f \in C_c(X, \mathbb{C})$. Moreover, μ satisfies*

$$\mu(U) = \sup \{I(f) : f \in C_c(X, \mathbb{C}), f \prec U\}, \quad \forall \text{ open subset } U \subset X, \tag{11.1}$$

and

$$\mu(K) = \inf \{I(f) : f \in C_c(X, \mathbb{C}), f \geq \chi_K\}, \quad \forall \text{ compact subset } K \subset X. \tag{11.2}$$

Proof. Let us begin by establishing uniqueness. If μ is a Radon measure such that $I(f) = \int f \, d\mu$ for all $f \in C_c(X, \mathbb{C})$, and $U \subset X$ is open, then clearly $I(f) \leq \mu(U)$ whenever $f \prec U$. On the other hand, if $K \subset U$ is compact, by Urysohn's lemma there is an $f \in C_c(X, \mathbb{C})$ such that $f \prec U$ and $f = 1$ on K, and hence $\mu(K) \leq \int f \, d\mu \leq I(f) \leq \mu(U)$. Since μ is inner regular, it follows that (11.1) is satisfied. Thus μ is determined by I on open sets, and hence on all Borel sets because of outer regularity.

This argument proves the uniqueness of μ and also suggests how to go about proving existence. The outline is as follows.

(1) We define $\mu(U)$ for any open subset U by (11.1), and then define

$$\mu^*(E) = \inf \{\mu(U) : U \supset E, U \text{ is open}\}.$$

(2) μ^* is an outer measure, and $\mu^*(U) = \mu(U)$ for any open subset U.

(3) Every open set is μ^*-measurable, and hence $\mu := \mu^*|_{\mathcal{B}_X}$ is a Borel measure.

(4) μ satisfies (11.2).
(5) $I(f) = \int f\, d\mu$ for all $f \in C_c(X, \mathbb{C})$.

One can refer to [For] for details. $\square$

Remark 11.7. Obviously, for all Radon measure μ, the map $f \mapsto \int f\, d\mu$ is a positive linear functional on $C_c(X, \mathbb{C})$. Hence we derive the correspondence between Radon measures and positive linear functionals.

More generally, if μ is a Borel measure on X such that $\mu(K) < \infty$ for every compact $K \subset X$, then clearly $C_c(X, \mathbb{C}) \subset L^1(\mu)$, so the map $f \mapsto \int f\, d\mu$ is a positive linear functional on $C_c(X, \mathbb{C})$. Then people will think, is such measure already the corresponding Radon measure? The answer is no, but if we add some new condition to the space, the answer will be yes. See theorem 11.13.

11.C. Conclusions of regularity — σ-finite, F_σ and G_δ, σ-compact.

Theorem 11.8. *Every Radon measure on an LCH space X is inner regular on all of its σ-finite sets.*

Proof. Suppose that μ is Radon and E is σ-finite Borel set. If $\mu(E) < \infty$, for any $\varepsilon > 0$ we can choose an open $U \supset E$ such that $\mu(U) < \mu(E) + \varepsilon$ and a compact $F \subset U$ such that $\mu(F) > \mu(U) - \varepsilon$. Since $\mu(U \setminus E) < \varepsilon$, we can also choose an open $V \supset U \setminus E$ such that $\mu(V) < \varepsilon$. Setting $K = F \setminus V$, then K is compact, $K \subset E$, and

$$\mu(K) = \mu(F) - \mu(F \cap V) > \mu(E) - \varepsilon - \mu(V) > \mu(E) - 2\varepsilon.$$

Thus μ is inner regular on E.

On the other hand, if $\mu(E) = \infty$, E is an increasing union of sets E_j with $\mu(E_j) < \infty$ and $\mu(E_j) \to \infty$. Thus for any $N \in \mathbb{N}$ there exists j such that $\mu(E_j) > N$, and hence, by the preceding argument, there exists a compact $K \subset E_j$ with $\mu(K) > N$. Hence μ is inner regular on E. $\square$

Corollary 11.9. *Every σ-finite Radon measure on an LCH space X is regular. If an LCH space X is σ-compact, every Radon measure on X is regular.*

Corollary 11.10. *Let X be an LCH space. If μ is a Radon measure on X, then $C_c(X, \mathbb{C})$ is dense in $L^p(\mu)$ for $1 \le p < \infty$.*

Proof. Via proposition 12.7 (2), it suffices to show that for any Borel set E with $\mu(E) < \infty$, χ_E can be approximated in the L^p norm by elements of $C_c(X, \mathbb{C})$.

Given $\varepsilon > 0$, by theorem 11.8 there exist a compact $K \subset E$ and an open $U \supset E$ such that $\mu(U \setminus K) < \varepsilon$. and by Urysohn's lemma we can choose $f \in C_c(X, \mathbb{C})$ such that $\chi_K \le f \le \chi_U$. Then $\|\chi_E - f\|_p \le \mu(U \setminus K)^{1/p} < \varepsilon^{1/p}$. We are done. $\square$

Remark 11.11. It follows from corollary 11.10, theorem 8.19, proposition 7.16 and proposition 9.23 that $C_c(\mathbb{R}^n)$ is dense in $L^p(m)$.

Theorem 11.12. *Let X be an LCH space. Suppose that μ is a σ-finite Radon measure on X and E is a Borel set in X.*

(1) For every $\varepsilon > 0$ there exist an open U and a closed F with $F \subset E \subset U$ and $\mu(U \setminus F) < \varepsilon$.

(2) There exist an F_σ set A and a G_δ set B such that $A \subset E \subset B$ and $\mu(B \setminus A) = 0$.

Proof. Write $E = \bigcup_{j=1}^{\infty} E_j$ where the E_j's are disjoint and have finite measure. For each j, choose an open $U_j \supset E_j$ with $\mu(U_j) < \mu(E_j) < \varepsilon 2^{-j-1}$, and set $U = \bigcup_{j=1}^{\infty} U_j$. Then U is open, and $\mu(U \setminus E) \le \sum_{j=1}^{\infty} \mu(U_j \setminus E_j) < \frac{\varepsilon}{2}$. Similarly, for E^c, we obtain an open $V \supset E^c$ with $\mu(V \setminus E^c) < \frac{\varepsilon}{2}$. Let $F = V^c$. Then F is closed, $F \subset E$, and

$$\mu(U \setminus F) = \mu(U \setminus E) + \mu(E \setminus F) = \mu(U \setminus E) + \mu(V \setminus E^c) < \varepsilon$$

Then (1) follows.

By (1) there exist open sets G_j's and closed sets F_j's satisfying $F_j \subset E \subset G_k$ and $m(G_j \setminus F_j) \le \frac{1}{j}$. Then sets $F = \bigcup_{j=1}^{\infty} F_j$ and $G = \bigcap_{j=1}^{\infty} G_j$ are as required. □

Theorem 11.13. *Let X be an LCH space in which every open set is σ-**compact** (which is the case, for example, if X is second countable). Then every Borel measure on X that is finite on compact sets is regular and hence Radon.*

Proof. If μ is a Borel measure on X such that $\mu(K) < \infty$ for every compact $K \subset X$, then clearly $C_c(X, \mathbb{C}) \subset L^1(\mu)$, so the map $f \mapsto \int f\, d\mu$ is a positive linear functional on $C_c(X, \mathbb{C})$. Let ν be the associated Radon measure according to Riesz representation theorem 11.6.

If $U \subset X$ is open, let $U = \bigcup_{j=1}^{\infty} K_j$ where each K_j is compact. Choose $f_1 \in C_c(X, \mathbb{C})$ such that $f \prec U$ and $f = 1$ on K_1. Proceeding inductively, for $n > 1$ choose $f_n \in C_c(X, \mathbb{C})$ such that $f_n \prec U$ and $f_n = 1$ on $\bigcup_{j=1}^{n} K_j$ and on $\bigcup_{j=1}^{n-1} \operatorname{supp}(f_j)$. Then f_n increses pointwise to χ_U as $n \to \infty$, and hence

$$\mu(U) = \lim_{n\to\infty} \int f_n\, d\mu = \lim_{n\to\infty} \int f_n\, d\nu = \nu(U)$$

by the monotone convergence theorem 9.10.

Next, if E is any Borel set and $\varepsilon > 0$, by theorem 11.12 there exists an open $V \supset E$ and a closed $F \subset E$ with $\nu(V \setminus F) < \varepsilon$. Since $V \setminus F$ is open, $\mu(V \setminus F) = \nu(V \setminus F) < \varepsilon$. In particular, $\mu(V) \le \mu(E) + \varepsilon$ and hence μ is outer regular. Also, $\mu(F) \ge \mu(E) - \varepsilon$. Note that F is σ-compact since X is σ-compact, so there exist compact $K_j \subset F$ with $\mu(K_j) \to \mu(F)$. It follows that μ is inner regular. Thus μ is regular, and equal to ν by Riesz representation theorem 11.6. □

11.D. **LSC functions, integration approximation.** In the next we introduce some properties of LSC functions, and then use them to derive a new method of integration approximation.

Proposition 11.14. *Let X be a topological space.*

(1) If U is open in X, then χ_U is LSC.
(2) If f is LSC and $c \in [0, \infty)$, then cf is LSC.
(3) If $\mathcal{S}$ is a family of LSC functions and $f(x) = \sup\{g : g \in \mathcal{S}\}$, then f is LSC.
(4) If f_1 and f_2 are LSC, so is $f_1 + f_2$.
(5) If X is an LCH space and f is LSC and non-negative, then

$$f(x) = \sup\{g(x) : g \in C_c(X, \mathbb{C}), 0 \le g \le f\}.$$

Proof. One can refer to [For] proposition 7.11. □

Proposition 11.15. *If μ is a Radon measure and f is non-negative Borel measurable function, then*

$$\int f\,d\mu = \inf\left\{\int g\,d\mu : g \geq f, g \text{ is } LSC\right\}.$$

If $\{x : f(x) > 0\}$ is σ-finite, then

$$\int f\,d\mu = \sup\left\{\int g\,d\mu : 0 \leq g \leq f, g \text{ is } USC\right\}.$$

Proof. Let (φ_n) be a sequence of non-negative simple functions that increse pointwise to f. Then $f = \varphi_1 + \sum_{n=2}^{\infty}(\varphi_n - \varphi_{n-1})$, and each term in this series is a non-negative simple function, so we can write $f = \sum_{j=1}^{\infty} a_j \chi_{E_j}$ where $a_j > 0$. Given $\varepsilon > 0$, for each j choose an open $U_j \supset E_j$ such that $\mu(U_j) \leq \mu(E_j) + \varepsilon 2^{-j} a_j^{-1}$. Then $g = \sum_{j=1}^{\infty} a_j \chi_{U_j}$ is LSC by proposition 11.14. Note that $g \geq f$, and $\int g\,d\mu \geq \int f\,d\mu + \varepsilon$; this establishes the first assertion.

For the second, if $a > \int f\,d\mu$, let N be large enough so that $\sum_{j=1}^{N} a_j \mu(E_j) > a$. Since the E_j's are σ-finite, by theorem 11.8 there are compact sets $K_j \subset E_j$ such that $\sum_{j=1}^{N} a_j \mu(K_j) > a$. Thus if $g = \sum_{j=1}^{N} a_j \chi_{K_j}$, then g is USC, $g \leq f$, and $\int g\,d\mu > a$. □

Proposition 11.16. *Let $\mathcal{S}$ be a family of non-negative LSC functions on an LCH space X that is direct by $\leq$ (that is, for every $g_1, g_2 \in \mathcal{S}$ there exists $g \in \mathcal{S}$ such that $g_1 \leq g$ and $g_2 \leq g$.) Set*

$$f(x) = \sup\{g(x) : g \in \mathcal{S}\}.$$

If μ is any Radon measure on X, then

$$\int f\,d\mu = \sup\left\{\int g\,d\mu : g \in \mathcal{S}\right\}.$$

Proof. It easily follows from proposition 11.14 and the monotone convergence theorem 9.10. One can refer to [For] for details. □

Corollary 11.17. *If μ is Radon and f is non-negative and LSC, then*

$$\int f\,d\mu = \sup\left\{\int g\,d\mu : g \in C_c(X,\mathbb{C}), 0 \leq g \leq f\right\}.$$

Proof. It follows from proposition 11.14 (5) and proposition 11.16. □

11.E. $C_c(X,\mathbb{C})^* = C_0(X,\mathbb{C})^*$**, Jordan decomposition of** $C_0(X,\mathbb{R})^*$**.** Now we generalize the Riesz representation theorem 11.6 to the case of bounded linear functionals. First we still consider the positive linear functional.

Proposition 11.18. *Let X be an LCH space, and let I be a positive linear functional on $C_c(X,\mathbb{C})$. Then I is bounded with respect to the uniform norm $\|\cdot\|_\infty$ iff the associated Radon measure μ satisfies $\mu(X) < \infty$.*

Proof. By Riesz representation theorem 11.6 (11.1),

$$\mu(X) = \sup\{I(f) : f \in C_c(X,\mathbb{C}), 0 \leq f \leq 1\}$$

It follows that

$$\sup_{\substack{f\in C_c(X,\mathbb{C})\\ \|f\|_\infty\le 1}} \frac{I(f)}{\|f\|_\infty} < \infty \implies \mu(X) < \infty.$$

The converse statement is obvious. We are done. $\square$

Remark 11.19. We have identitied the positive bounded linear functionals on $C_c(X,\mathbb{C})$: they are given by integration against finite Radon measures.

In the next, we generalize positive bounded linear functionals on $C_c(X,\mathbb{C})$ to all $I \in C_c(X,\mathbb{C})^*$. The key fact is that $C_c(X,\mathbb{C})^*$ can be reduced to $C_c(X,\mathbb{R})^*$ and that $C_c(X,\mathbb{R})^*$ has a "Jordan decomposition".

Before doing this generalization, we do some small technical treatments. Note that $C_c(X,\mathbb{C})$ is not always complete, and $C_c(X,\mathbb{C})$ and its completion have the same dual space. Sometimes it brings convenience if we realize this.

Definition 11.20 (Vanishes at infinity and $C_0(X,\mathbb{C})$)**.** *Let X be a topological space and let $f \in C(X,\mathbb{C})$. We say that f **vanishes at infinity** if for every $\varepsilon > 0$ the set $\{x : |f(x)| \ge \varepsilon\}$ is compact, and we define*

$$C_0(X,\mathbb{C}) = \{f \in C(X,\mathbb{C}) : f \text{ vanishes at infinity}\}.$$

Proposition 11.21. *If X is an LCH space, then $C_0(X,\mathbb{C})$ is the closure of $C_c(X,\mathbb{C})$ in the space $(C(X,\mathbb{C}), \|\cdot\|_\infty)$.*

Proof. One can refer to [For] proposition 4.35. $\square$

Remark 11.22. Therefore, if X is an LCH space, then $C_0(X,\mathbb{C})$ is the completion of $C_c(X,\mathbb{C})$, and hence $C_0(X,\mathbb{C})^* = C_c(X,\mathbb{C})^*$.

Now we focus on $C_0(X,\mathbb{C})^*$. As mentioned above, we note that any $I \in C_0(X,\mathbb{C})^*$ is uniquely determined by its restriction J to $C_0(X,\mathbb{R})$, and we have $J = J_1 + iJ_2$ where J_1 and J_2 are real linear functionals. Moreover, $C_c(X,\mathbb{R})^*$ has the following "*Jordan decomposition*":

Proposition 11.23. *If $f \in C_0(X,\mathbb{R})^*$, there exist positive functionals $I^\pm \in C_0(X,\mathbb{R})^*$ such that $I = I^+ - I^-$.*

Proof. Define

$$I_1 : C_0(X,[0,\infty)) \mapsto \mathbb{R}, \quad f \mapsto \sup\{I(g) : g \in C_0(X,\mathbb{R}), 0 \le g \le f\}.$$

and then define

$$I^+ : C_0(X,\mathbb{R}) \to \mathbb{R}, \quad f \mapsto I_1(f^+) - I_1(f^-), \quad \text{and} \quad I^- = I^+ - I.$$

It's easy to see that $I^\pm \in C_0(X,\mathbb{R})^*$. One can refer to [For] for details. $\square$

Corollary 11.24. *For any $I \in C_0(X,\mathbb{C})^*$ there are finite Radon measures $\mu_1, \cdots, \mu_4$ such that $I(f) = \int f\, d\mu$ where $\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4)$.*

11.F. **Signed (or complex) Radon measure, Riesz representation theorem.** To establish a better correspondence, we introduce some new concepts.

Definition 11.25 (Signed Radon measure and complex Radon measure)**.** *A* ***signed Radon measure*** *is a signed Borel measure whose positive and nagative variations are Radon, and a* ***complex Radon measure*** *is a complex Borel measure whose real and imaginary parts are signed Radon measures. We denote* ***the space of complex Radon measures*** *on X by $M(X)$, and for $\mu \in M(X)$ we define*

$$\|\mu\| = |\mu|(X),$$

where $|\mu|$ is the total variation of μ.

Remark 11.26. Note that if ν is a complex measure, then ν_r and ν_i are finite signed measures. Hence, via theorem 11.13, if X is an LCH space in which every open set is σ-compact, then every complex Borel measure is Radon.

Proposition 11.27. *If μ is a complex Borel measure, then μ is Radon iff $|\mu|$ is Radon. Moreover, $M(X)$ is a vector space and $\mu \mapsto \|\mu\|$ is a norm on it.*

Proof. Trivial. □

Theorem 11.28 (Riesz representation theorem)**.** *Let X be an LCH space, and for $\mu \in M(X)$ and $f \in C_0(X, \mathbb{C})$ let $I_\mu(f) = \int f\, d\mu$. Then the map $\Phi : M(X) \to C_0(X, \mathbb{C})^*$, $\mu \mapsto I_\mu$, is an isometric isomorphism.*

Proof. By corollary 11.24, the linear map Φ is surejective. It suffices to show that $\|\mu\| = \|I_\mu\|$ for all $\mu \in M(X)$. Note that

$$\left| \int f\, d\mu \right| \le \int |f|\, d\mu \le \|f\|_\infty \|\mu\|,$$

so $I_\mu \in C_0(X)^*$ and $\|I_\mu\| \le \|\mu\|$. Moreover, if $h = d\mu/d|\mu|$, then $|h| = 1$, and hence by Lusin's theorem, for any $\varepsilon > 0$ there exists $f \in C_c(X)$ such that $\|f\|_\infty = 1$ and $f = \overline{h}$ except on a set E with $|\mu|(E) < \varepsilon/2$. Then

$$\begin{aligned} \|\mu\| = \int |h|^2\, d|\mu| = \int \overline{h}\, d\mu &\le \left| \int f\, d\mu \right| + \left| \int \left(f - \overline{h} \right) d\mu \right| \\ &\le \left| \int f\, d\mu \right| + 2|\mu|(E) \\ &\le \|I_\mu\| + \varepsilon \end{aligned}$$

It follows that $\|\mu\| \le \|I_\mu\|$. So the proof is complete. □

Remark 11.29. One can refer to [For] or [Evaa] for Lusin's theorem.

Corollary 11.30. *If X is a compact Hausdorff space, then $C(X, \mathbb{C})^*$ is isometrically isomorphism to $M(X)$.*

For more properties of Radon measures, such as the products of Radon measures, one can refer to [For].

11.G. **Filling problems and covering theorems for $\mathbb{R}^n$, Vitali, Besicovitch.** Now we come back to $\mathbb{R}^n$. We will introduce the differentiation of Radon measures with respect to Radon measures, and the weak convergence of Radon measures.

Conclusions of filling propblems are important tools of the differentiation theory. Generally speaking, filling problems are highly related to the covering theorems.[5] In the next we introduce filling problems and covering theorems for $\mathbb{R}^n$ first.

Theorem 11.31 (Filling theorem)**.** *There are two editions.*

(1) (Filling open sets with balls – Lebesgue measure) *Let $U \subset \mathbb{R}^n$ be open, $\delta > 0$. There exists a countable collection $\mathcal{G}$ of disjoint closed balls in U such that* $\operatorname{diam} B < \delta$ *for all $B \in \mathcal{G}$ and*

$$m\left(U \setminus \bigcup_{B\in\mathcal{G}} B\right) = 0.$$

(2) (Filling open sets with balls – Borel outer measure) *Let μ^* be an outer measure on $\mathbb{R}^n$ such that all Borel sets are μ^*-measurable, and let $\mathcal{F}$ be any collection of nondegenerate closed balls. Let A denote the set of centers of the balls in $\mathcal{F}$. Assume that*

$$\mu^*(A) < \infty,$$

and

$$\inf\{r : B(a,r) \in \mathcal{F}\} = 0, \quad \forall a \in A.$$

Then for each open set $U \subset \mathbb{R}^n$, there exists a countable collection $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

$$\bigcup_{B\in\mathcal{G}} B \subset U,$$

and

$$\mu^*\left((A \cap U) \setminus \bigcup_{B\in\mathcal{G}} B\right) = 0.$$

Remark 11.32. We can set $A = U$ in (2) to derive the normal filling theorem for open sets.

Proof. The framework and idea of the first assertion are as follows.

(1) WLOG we can assume that $m(U) < \infty$; otherwise we apply the finite conclusion to

$$U_m := \{x \in U : m < |x| < m+1\}, \quad m = 0, 1, \cdots.$$

(2) It suffices to cover a fixed percentage of measures at a time by finite disjoint closed balls.

(3) Since for each ball $B \subset \mathbb{R}^n$ we have

$$\frac{m(\widehat{B})}{m(B)} = 5^n. \tag{11.3}$$

[5]We also use Vitali's covering theorem in the proof of Lebesgue differentiation theorem.

where $\widehat{B} = B(x, 5r)$ for $B = B(x, r)$. It suffices to show that there exists a finite collection $(B_j)_{j=1}^M$ of disjoint closed balls with

$$\bigcup_{j=1}^{M} \widehat{B}_j \supset U.$$

This is the Vitali's covering theorem, for which we will introduce later.

The framework and idea of the second assertion are as follows.

(1) It suffices to cover a fixed percentage of measures at a time by finite disjoint closed balls.
(2) In this case we don't (11.3), we turn to the Pigeonhole principle. This is the Besicovitch's covering theorem, which we will introduce later.

The left is trivial. It suffices to show the following covering theorems. □

Theorem 11.33 (Covering theorems)**.** *There are two editions of covering theorems.*

*(1) (**Vitali's covering theorem**) Let $\mathcal{F}$ be any collection of nondegenerate closed balls in $\mathbb{R}^n$ with*

$$\sup\{\operatorname{diam}(B) : B \in \mathcal{F}\} < \infty.$$

Then there exists a countable familiar $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

$$\bigcup_{B\in F} B \subset \bigcup_{B\in\mathcal{G}} \widehat{B},$$

where $\widehat{B} = B(x, 5r)$ for $B = B(x, r)$.

*(2) (**Besicovitch's covering theorem**) There exists a constant N_n, depending only on the dimension n, with the following property:*

If $\mathcal{F}$ is any collection of nondegenerate closed balls in $\mathbb{R}^n$ with

$$\sup\{\operatorname{diam}(B) : B \in \mathcal{F}\} < \infty,$$

and if A is the set of centers of balls in $\mathcal{F}$, then there exists N_n countable collections $\mathcal{G}_1, \cdots, \mathcal{G}_n$ of disjoint balls in $\mathcal{F}$ such that

$$A \subset \bigcup_{j=1}^{N_n} \bigcup_{B\in\mathcal{G}_j} B.$$

Proof. The framework and idea of Vitali's covering theorem are as follows.

(1) It's natural to find the maximal disjoint collection, which is certainly countable. The difficulty is to ensure $(\widehat{B}_j)$ covers U.
(2) It suffices to show that for each $B \in \mathcal{F}$, $B \subset \widehat{B}_j$ for some j. By the maximal property of $(\widehat{B}_j)$, it follows that $B \cap B_j \neq \varnothing$ for some j. It suffices to modify the construction of $(\widehat{B}_j)$ such that there exists B_j with $B \cap B_j \neq \varnothing$ and $\operatorname{diam}(B_j) > 2\operatorname{diam}(B)$.
(3) According to the requirements, we make a division of diameter:

$$\mathcal{F}_j := \left\{B \in \mathcal{F} : \frac{D}{2^j} < \operatorname{diam}(B) < \frac{D}{2^{j-1}}\right\} \quad \text{where} \quad D := \sup\{\operatorname{diam}(B) : B \in \mathcal{F}\},$$

and choose $\mathcal{G}_k$ as any maximal disjoint subcollection of

$$\left\{B \in \mathcal{F}_k : B \cap B' = \varnothing, \forall B' \in \bigcup_{j=1}^{k-1} \mathcal{G}_j\right\}.$$

(4) The left is trivial.

The framework and idea of Vitali's covering theorem are as follows.

(1) Our basic framework is to find a covering model that covers A in countable steps, and then divide it.
(2) For the covering model, a natural idea is to eliminate the gaps in order to ensure the covering, and another natural idea is to ensure the basic separation, which ensures that we can make reasonable division. We can create some basic models, in which some parameters can be adjusted.
(3) An intuitive observation is that with a certain degree of separation, there must be few "near balls", so the division of "near balls" is actually easy to talk about. The key is the handling of "distant balls".
(4) Given a "distant ball", we can make a rough estimate of the number of balls in front that intersect it. Then, we adjust the parameters of the covering model to ensure that the intersection numbers have a fixed upper bound. Then our work can be successful.

For a detailed proof of Besicovitch's covering theorem, one can refer to [Evaa]. □

11.H. **Differentiating Radon measures with respect to Radon measures.** In the next we introduce the differentiation theory of Radon measures. It's the generalization of that we introduced before. One can refer to [Evaa] for the proofs.

Definition 11.34 (Differentiation of Radon measures). *Let μ and ν be Radon measures on $\mathbb{R}^n$. For each $x \in \mathbb{R}^n$, define*

$$\overline{D}_\mu \nu(x) = \begin{cases} \limsup\limits_{r\to 0} \dfrac{\nu(B(x,r))}{\mu(B(x,r))}, & \text{if } \mu(B(x,r)) > 0 \text{ for all } r > 0; \\ +\infty, & \text{if } \mu(B(x,r)) = 0 \text{ for some } r > 0. \end{cases}$$

and

$$\underline{D}_\mu \nu(x) = \begin{cases} \liminf\limits_{r\to 0} \dfrac{\nu(B(x,r))}{\mu(B(x,r))}, & \text{if } \mu(B(x,r)) > 0 \text{ for all } r > 0; \\ +\infty, & \text{if } \mu(B(x,r)) = 0 \text{ for some } r > 0. \end{cases}$$

*If $\overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x) < \infty$, we say ν is **differentiable** with respect to μ at x and write*

$$D_\mu \nu(x) := \overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x).$$

*$D_\mu \nu(x)$ is the **derivative** of ν with respect to μ. We also call $D_\mu \nu$ the **density** of ν with respect to μ.*

Lemma 11.35. *Let μ and ν be Radon measures on $\mathbb{R}^n$. Fix $0 < \alpha < \infty$. Then*

(1) $A \subset \left\{x \in \mathbb{R}^n : \underline{D}_\mu \nu \le \alpha\right\}$ implies $\nu(A) \le \alpha\mu(A)$.

(2) $A \subset \left\{x \in \mathbb{R}^n : \overline{D}_\mu \nu \leq \alpha\right\}$ *implies* $\nu(A) \geq \alpha\mu(A)$.

(If A is not a Borel set, we regard μ *and* ν *as Radon outer measures.)*

Theorem 11.36 (Differentiating Radon measures). *Let* μ *and* ν *be Radon measures on* $\mathbb{R}^n$. *Then*

(1) $D_\mu \nu$ *exists and is finite* μ*-a.e.;*
(2) $D_\mu \nu$ *is* μ*-measurable.*

Theorem 11.37 (Randon-Nikodym-Lebesgue theorem). *Let* μ *and* ν *be Radon measures on* $\mathbb{R}^n$.

(1) Then

$$\nu = \nu_{ac} + \nu_s$$

where ν_{ac} *and* ν_s *are Radon measures on* $\mathbb{R}^n$ *with*

$$\nu_{ac} \ll \mu \quad \text{and} \quad \nu_s \perp \mu.$$

(2) Furthermore,

$$D_\mu \nu = D_\mu \nu_{ac}, \quad \text{and} \qquad D_\mu \nu_s = 0 \quad \mu - a.e.$$

and for any Borel set A we have

$$\nu(A) = \int_A D_\mu \nu \, d\mu + \nu_s(A).$$

Theorem 11.38 (Average properties). *Let* μ *be a Radon measure on* $\mathbb{R}^n$.

(1) If $f \in L^1_{loc}(\mathbb{R}, \mu)$*, then*

$$\lim_{r \to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f \, d\mu = f(x), \quad \text{for } \mu - a.e. \ x \in \mathbb{R}^n.$$

(2) If $f \in L^p_{loc}(\mathbb{R}, \mu)$ *for some* $1 \leq p < \infty$*, then*

$$\lim_{r \to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f - f(x)|^p \, d\mu = 0, \quad \text{for } \mu - a.e. \ x \in \mathbb{R}^n.$$

(3) If $f \in L^p_{loc}(\mathbb{R}, m)$ *for some* $1 \leq p < \infty$*, then*

$$\lim_{B \to \{x\}} \frac{1}{\mu(B)} \int_B |f - f(x)|^p \, dm = 0, \quad \text{for } m - a.e. \ x \in \mathbb{R}^n.$$

where the limit is taken over all closed balls B containing x, as $\operatorname{diam}(B) \to 0$.

Theorem 11.39 (Points of density 1 and density 0). *Let* $E \subset \mathbb{R}^n$ *be Lebesgue-measurable. Then*

$$\lim_{r \to 0} \frac{m\,(B(x,r) \cap E)}{m\,(B(x,r))} = 1, \quad \text{for } m - a.e. \ x \in E$$

and

$$\lim_{r \to 0} \frac{m\,(B(x,r) \cap E)}{m\,(B(x,r))} = 0 \quad \text{for } m - a.e. \ x \in \mathbb{R}^n \setminus E$$

11.I. **Weak convergence of Radon measures.** Finally we introduce the properties of weak convergence of Radon measures and their corollaries. These are not difficult, one can refer to [Evaa] for their proofs.

Definition 11.40 (Weak convergence of Radon measures). *Let μ, μ_k ($k = 1, 2, \cdots$) be Radon measures on $\mathbb{R}^n$. Then we say that the measures $(\mu_k)_{k=1}^{\infty}$* ***converge weakly*** *to the measure μ, written*

$$\mu_k \rightharpoonup \mu,$$

if we have

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f \, d\mu_k = \int_{\mathbb{R}^n} f \, d\mu, \quad \forall f \in C_c(\mathbb{R}^n).$$

Theorem 11.41 (Weak convergence). *Let $(\mu_k)_{k=1}^{\infty}$ be a sequence of Radon measures on $\mathbb{R}^n$ satisfying*

$$\sup_{k\geq 1} \mu_k(K) < \infty \quad \text{for each compact set } K \subset \mathbb{R}^n.$$

Then there exists a subsequence $(\mu_{k_j})_{j=1}^{\infty}$ and a Radon measure μ such that

$$\mu_{k_j} \rightharpoonup \mu.$$

Corollary 11.42. *Let $U \subset \mathbb{R}^n$ be open, let $1 \leq p < \infty$, and let $(f_k)_{k=1}^{\infty}$ be a sequence of functions in $L^p(U, m)$ satisfying*

$$\sup_{k\geq 1} ||f_k||_{L^p(U,m)} < \infty$$

Then we have the following conclusions about weak convergence:

(1) If $1 < p < \infty$, then there exists a subsequence $(f_{k_j})_{j=1}^{\infty}$ and a function $f \in L^p(U, m)$ such that

$$f_{k_j} \rightharpoonup f \quad \text{in } L^p(U, m).$$

(2) If $p = 1$, and suppose also

$$\lim_{l\to\infty} \sup_{k\geq 1} \int_{\{|f_k|\geq l\}} |f_k| \, dm = 0,$$

then there exists a subsequence $(f_{k_j})_{j=1}^{\infty}$ and a function $f \in L^1(U, m)$ such that

$$f_{k_j} \rightharpoonup f \quad \text{in } L^1(U, m).$$

Corollary 11.43 (Biting lemma). *Assume that the open subset U is bounded and let $(f_k)_{k=1}^{\infty}$ be a sequence of functions in $L^1(U, m)$ satisfying*

$$\sup_{k\geq 1} ||f_k||_{L^1(U,m)} < \infty.$$

Then there exists a subsequence $(f_{k_j})_{j=1}^{\infty}$ and a function $f \in L^1(U, m)$ such that for each $\delta > 0$ there exists a Lebesgue measurable set $E \subset U$ with

$$m(E) < \delta,$$

and

$$f_{k_j} \rightharpoonup f \quad \textit{in } L^p(U \setminus E, m).$$

For more properties of Radon measures, one can refer to [Evaa] and [For].

12. L^p SPACES

12.A. L^p space.

Definition 12.1 (L^p space). *Let $(X, \mathcal{X}, \mu)$ be a measure space. If f is a measurable function on X and $0 < p < \infty$, we define*

$$\|f\|_p = \left(\int |f|^p \, d\mu\right)^{1/p},$$

(allowing the possibility that $\|f\|_p = \infty$), and we define

$$L^p(X, \mathcal{X}, \mu) = \left\{f : X \to \mathbb{C} : f \text{ is measurable and } \|f\|_p < \infty\right\}.$$

Moreover, if f is a measurable function on X and $p = \infty$, we define

$$\|f\|_\infty = \inf\{a \geq 0 : \mu(\{x : |f(x)| > a\}) = 0\},$$

with the convention that $\inf \varnothing = \infty$, *and we define*

$$L^\infty(X, \mathcal{X}, \mu) = \left\{f : X \to \mathbb{C} : f \text{ is measurable and } \|f\|_\infty < \infty\right\}.$$

We abbreviate $L^p(X, \mathcal{X}, \mu)$ by $L^p(\mu)$, $L^p(X)$, or simply L^p when this will cause no confusion.

If A is any nonempty set, we define $l^p(A)$ to be $L^p(\mu)$ where μ is counting measure on $(A, \mathcal{P}(A))$, and we denote $l^p(\mathbb{N})$ simply by ℓ^p.

Remark 12.2. As like in remark 9.24, we usually **redefine $L^p(\mu)$ to be the set of equivalence classes of a.e.-defined L_p-integrable functions on** X, where f and g are considered equivalent iff $f = g$ a.e.

This new $L^p(\mu)$ is still a complex vector space, and we employ the notation "$f \in L^p(\mu)$" to mean that f is an a.e.-defined integrable function.

Remark 12.3. In particular, $\|f^r\|_p = \|f\|_{pr}^r$ for all $0 < p, r < \infty$.

12.B. Hölder inequality, Minkowski inequality.

Proposition 12.4. *Let $(X, \mathcal{X}, \mu)$ be a measure space. Let f and g be measurable functions on X.*

(1) (Hölder inequality) *If $1 < p < \infty$ and $p^{-1} + q^{-1} = 1$, then*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q. \tag{12.1}$$

In particular, if $f \in L^p$ and $g \in L^q$, then $fg \in L^1$. Besides,

$$f \in L^p \quad \text{and} \quad g \in L^q : \|fg\|_1 = \|f\|_p \|g\|_q \iff$$
$$\alpha|f|^p = \beta|g|^q \ \mu\text{-a.e. for some constants } \alpha, \beta \text{ with } \alpha\beta \neq 0.$$

Moreover, if $p = 1$ and $q = \infty$ we also have $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$. In particular, if $f \in L^1$ and $g \in L^\infty$, then $fg \in L^1$. Besides,

$$f \in L^1 \quad \text{and} \quad g \in L^\infty : \|fg\|_1 = \|f\|_1 \|g\|_\infty \iff$$
$$|g(x)| = \|g\|_\infty \ \mu\text{-a.e. on } \{x \in X : f(x) \neq 0\}.$$

(2) (Minkowski inequality) *If $1 \leq p \leq \infty$ and $f, g \in L^p$, then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p. \tag{12.2}$$

Proof. For Hölder inequality, the case that $p = 1$ and $q = \infty$ is trivial, in the next we assume that $1 < p < \infty$. Note that the result is trivial if $\|f\|_p = 0$ or $\|g\|_q = 0$ (use proposition 9.23), or if $\|f\|_p = \infty$ or $\|g\|_q = \infty$. Then, setting

$$a(x) = \frac{|f(x)|}{\|f\|_p} \quad \text{and} \quad b(x) = \frac{|g(x)|}{\|g\|_q},$$

it follows from the Jesen inequality with respect to e^x that

$$a(x)b(x) \leq \frac{a^p(x)}{p} + \frac{b^q(x)}{q},$$

and hence

$$\int a(x)b(x)\,d\mu(x) \leq \frac{1}{p}\int a^p(x)\,d\mu(x) + \frac{1}{q}\int b^q(x)\,d\mu(x) = \frac{1}{p} + \frac{1}{q} = 1.$$

Thus (12.1) follows, and if $f \in L^p$ and $g \in L^q$, then $fg \in L^1$; moreover, in this case equality holds in (12.1) iff $a(x) = b(x)$ μ-a.e. iff $\alpha|f|^p = \beta|g|^q$ μ-a.e. for some constants α, β with $\alpha\beta \neq 0$.

For Minkowski inequality, the result is trivial if $p = 1$ or ∞, or if $f + g = 0$ μ-a.e. Otherwise, setting $p^{-1} + q^{-1} = 1$, it follows from Hölder inequality that

$$\begin{aligned} \|f+g\|_p^p = \|(f+g)^p\|_1 &\leq \left\||f+g|^{p-1}|f|\right\|_1 + \left\||f+g|^{p-1}|g|\right\|_1 \\ &\leq \left\||f+g|^{p-1}\right\|_q \left(\|f\|_p + \|g\|_p\right) \\ &= \|f+g\|_{q(p-1)}^{p-1} \left(\|f\|_p + \|g\|_p\right) \end{aligned}$$

Note that $q(p-1) = p$; then Minkowski inequality follows. □

The Minkowski inequality has the following generalized edition.

Proposition 12.5 (Minkowski inequality for integration). *Let $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ be σ-finite measure spaces, and let f be a measurable function on $(X \times Y, \mathcal{X} \times \mathcal{Y}, \mu \times \nu)$.*

(1) If $f \in L^+(\mathcal{X} \times \mathcal{Y})$ and $1 \leq p < \infty$, then

$$\left(\int_X \left(\int_Y f(x,y)\,d\nu(y)\right)^p d\mu(x)\right)^{1/p} \leq \int_Y \left(\int_X f(x,y)^p\,d\mu(x)\right)^{1/p} d\nu(y).$$

(2) If $1 \leq p \leq \infty$, $f(\cdot, y) \in L^p(\mu)$ for ν-a.e. y, and the function $y \mapsto \|f(\cdot,y)\|_p$ is in $L^1(\nu)$, then $f(x,\cdot) \in L^1(\nu)$ for μ-a.e. x, the function $x \mapsto \int f(x,y)\,d\nu(y)$ is in $L^p(\mu)$, and

$$\left\|\int f(\cdot,y)\,d\nu(y)\right\|_p \leq \int \|f(\cdot,y)\|_p\,d\nu(y).$$

Proof. When $p = 1$, (1) is merely Fubini-Tonelli theorem 9.28. When $1 < p < \infty$, set $p^{-1} + q^{-1} = 1$ and set

$$H : X \to [0, \infty], \quad x \mapsto \|f(x,y)\|_{1;Y} = \int_Y f(x,y)\,d\nu(y).$$

Then for all $g \in L^q(\mu)$, by Fubini-Tonelli theorem 9.28 and Hölder inequality 12.4 we have

$$\begin{aligned} \|H(x)\|_{p;X}^p = \|H^p(x)\|_{1;X} &= \|H^{p-1}(x)H(x)\|_{1;X} \\ &= \left\|H^{p-1}(x)\|f(x,y)\|_{1;Y}\right\|_{1;X} \\ &= \left\|\|H^{p-1}(x)f(x,y)\|_{1;Y}\right\|_{1;X} \\ &= \left\|\|H^{p-1}(x)f(x,y)\|_{1;X}\right\|_{1;Y} \\ &\le \left\|\|H^{p-1}(x)\|_{q;X}\|f(x,y)\|_{p;X}\right\|_{1;Y} \\ &= \|H(x)\|_{p;X}^{p-1}\left\|\|f(x,y)\|_{p;X}\right\|_{1;Y}. \end{aligned}$$

Then (1) follows.

When $p < \infty$, (2) follows from (1) with f replaced by $|f|$ and Fubini-Tonelli theorem 9.28. When $p = \infty$, it is a simple consequence of the monotonicity of the integration. $\square$

Remark 12.6. Setting $(Y, \mathcal{Y}, \nu) = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu)$ where ν is the counting measure, by Minkowski inequality for integration 12.5 we know that if $(X, \mathcal{X}, \mu)$ is a σ-finite measure space and $(f_j)_{j=1}^{\infty}$ are a sequence functions in $L^+(X)$, then for $1 \le p < \infty$ we have

$$\left\|\sum_{j=1}^{\infty} f_j\right\|_p \le \sum_{j=1}^{\infty} \|f_j\|_p, \tag{12.3}$$

which is a generalization of Minkowski inequality 12.4. Since we have corollary 9.14, we can remove the condition that X is σ-finite, and the above proof applies.

12.C. **Banach, separability, dual, weak compact, uniformly convex.**

Proposition 12.7. *Let $(X, \mathcal{X}, \mu)$ be a measure space. Then*

(1) $(L^p, \|\cdot\|_p)$ is a Banach space for $1 \le p \le \infty$.
(2) For $1 \le p < \infty$, the set Λ of simple functions $f = \sum_{j=1}^n a_j \chi_{E_j}$, where $\mu(E_j) < \infty$ for all j, is dense in L^p.
(3) Let $1 < p < \infty$ and $p^{-1} + q^{-1} = 1$, for each $\varphi \in (L^p)^*$ there exists $g \in L^q$ such that $\varphi(f) = \int fg\,d\mu$ for all $f \in L^p$, and hence L^q is isometrically isometric to $(L^p)^*$. The same conclusion holds for $p = 1$ (and $q = \infty$) provided μ is σ-finite.

Proof. It easily follows from Minkowski inequality 12.4 that $(L^p, \|\cdot\|_p)$ is an N.V.S. for $1 \le p \le \infty$. It's obvious that $(L^\infty, \|\cdot\|_\infty)$ is Banach. In the next assume that $1 \le p < \infty$. Let (f_n) be a Cauchy sequence in L^p, then

$$\forall \varepsilon > 0,\ \exists N = N(\varepsilon) \in \mathbb{N},\ \forall n, m \ge N : \|f_n - f_m\|_p < \varepsilon.$$

Hence there exists a sequence of strictly increasing numbers $(n_j)_{j=1}^{\infty}$ such that

$$\left\|f_{n_{j+1}} - f_{n_j}\right\|_p < \frac{1}{2^j}, \quad \forall j \in \mathbb{N}.$$

By (12.3) we know

$$\left\|\sum_{j=1}^{\infty}\left|f_{n_{j+1}}-f_{n_j}\right|\right\|_p \le \sum_{j=1}^{\infty}\left\|f_{n_{j+1}}-f_{n_j}\right\|_p \le 1,$$

and hence it follows from properties 9.9 that

$$\sum_{j=1}^{\infty}\left|f_{n_{j+1}}-f_{n_j}\right| < \infty, \quad \mu\text{-a.e.}$$

Hence f_{n_j} converges to some measurable f μ-a.e pointwisely via proposition 6.13 and proposition 6.14. It follows from Fatou's lemma 9.19 that

$$\forall n \ge N(\varepsilon) : \|f_n - f\|_p = \left\|\liminf_{j\to\infty}\left|f_n - f_{n_j}\right|\right\|_p \le \liminf_{j\to\infty}\|f_n - f_{n_j}\|_p \le \varepsilon$$

Hence f_n converges to f in L^p. Then (1) follows.

For (2), clearly $\Lambda \subset L^p$. If $f \in L^p$, by theorem 6.22 choose a sequence (f_n) of simple functions such that $0 \le |f_1| \le |f_2| \le \cdots \le |f|$, $f_n \to f$ pointwise. Then $f_n \in L^p$ and $|f_n - f|^p \le 2^p|f|^p \in L^1$, so by the dominated convergence theorem 9.25, $\|f_n - f\|_p \to 0$. Moreover, the simple functions $f_n = \sum_{j=1}^{N_n}\chi_{E_j}$, where the E_j's are disjoint and the a_j's are nonzero, must satisfy that $\mu(E_j) < \infty$ since $\sum_{j=1}^{N_n}|a_j|^p\mu(E_j) = \int |f_n|^p\, d\mu < \infty$. Then (2) follows.

For (3), the existence of g follows from Radon-Nikodym-Lebesgue theorem 10.24. One can refer to [For] or [Tao] for details. □

Remark 12.8. When $0 < p < 1$, $(L^p, \|\cdot\|_p)$ is not an N.V.S.

Remark 12.9. (2) is not true for $p = \infty$ in general. Consider $f \equiv 1$ on $\mathbb{R}$.

Proposition 12.10. *There are some basic conclusions about separability.*

(1) If Ω is Lebesgue measurable ($1 \le p \le \infty$), then $L^\infty(\Omega, \mathcal{L}(\Omega), m)$ is separable.
(2) If Ω is Lebesgue measurable with $m(\Omega) > 0$, then $L^\infty(\Omega, \mathcal{L}(\Omega), m)$ is not separable.
(3) ℓ^p is separable ($1 \le p \le \infty$), but ℓ^∞ is not separable.

Proof. Easy. □

Now we prove a general conclusion about separability.

Lemma 12.11. *Let $(X, \mathcal{X}, \mu)$ be a σ-finite measure space. If $\mathcal{X}$ is countably generated, then the metric spaces $(\mathcal{X}, d)$ is separable, where*

$$d(A, B) = \mu(A\Delta B).$$

sketch of the proof. (1) WLOG we assume that μ is finite.
(2) Suppose $\mathcal{C} = \{C_n : n \in \mathbb{N}\}$ generates $\mathcal{X}$. Define $\mathcal{A}_n$ to be the algebra generated by $\{C_1, \dots, C_n\}$. Sicne $\mathcal{A}_n$ is finite, $\mathcal{A} = \bigcup_n \mathcal{A}_n$ is a countable algebra that generates $\mathcal{X}$.
(3) We construct the outer measure

$$\mu^*(E) = \inf\left\{\sum_n \mu(A_n) : A_n \in \widehat{\mathcal{A}}, , E \subset \bigcup_n A_n\right\}.$$

Carathéodory theorem 7.19 yields a complete measure that extends μ to a larger σ-algebra $\mathcal{M} \supset \mathcal{X}$.

(4) From this construction, we know that any measurable set can be approximate by sets in $\mathcal{A}$, and hence $(X, \mathcal{M})$ is separable.

(5) The conclusion follows by taking rational linear combinations of sets (indicator functions rather) in $\mathcal{A}$.

□

Proposition 12.12. *Let $(X, \mathcal{X}, \mu)$ be a measure space. Suppose that μ is σ-finite and $\mathcal{X}$ is separable (i.e. countably generated). Then L^p is separable for all $1 \le p < \infty$.*

Proof. By lemma 12.11 we know that $(\mathcal{X}, d)$ is separable. Let $(A_n)_{n\in\mathbb{N}}$ be a dense subset in $(\mathcal{X}, d)$. We claim that the countable set $\mathrm{span}_{\mathbb{Q}}\{A_n : n \in \mathbb{N}\}$ is dense in L^p.

Note that for $A, B \in \mathcal{X}$ we have

$$\|\chi_A - \chi_B\|_p^p = \int |\chi_A - \chi_B|^p = \int_{A\setminus B} 1 + \int_{B\setminus A} 1 = \mu(A\Delta B) = d(A, B)$$

It follows that $d(A_{n_k}, B) \to 0 \iff \|\chi_{A_{n_k}} - \chi_B\|_{L^p} \to 0$. By proposition 12.7 (2), simple functions are dense in L^p, and then the conclusion easily follows.

One can suitably modify the proof for complex valued functions. □

One can refer to corollary 11.42 for the properties of weak compactness of L^p spaces. One can refer to problem 3.43 (4) for the properties of uniform convexity of L^p spaces.

12.D. **Convolution, regularization.**

Definition 12.13 (Convolution). *Let f and g be measurable functions on $\mathbb{R}^n$. The **convolution** of f and g is the function $f * g$ defined by*

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x - y)g(y)dy$$

for all x such that the integration exists.

We introduce the following basic property first.

Proposition 12.14 (Basic properties of convolution). *Assuming that all integrals blow exist, we have*

(1) $f * g = g * f$;

(2) $(f * g) * h = f * (g * h)$.

(3) For $z \in \mathbb{R}^n$, $\tau_z(f * g) = (\tau_z f) * g = f * (\tau_z g)$.

(4) If A is the closure of $\{x + y : x \in \mathrm{supp}(f), : y \in \mathrm{supp}(g)\}$, then $\mathrm{supp}(f * g) \subset A$.

Here we use the notation $\tau_y f(x) = (x - y)$.

Proof. Well-known. □

In the next we introduce some basic propositions which imply the existence of convolution under some typical conditions.

Theorem 12.15 (Young). *Let $f \in L^1(\mathbb{R}^n)$ and let $g \in L^p(\mathbb{R}^n)$ with $1 \le p \le \infty$. Then $f * g(x)$ exists for a.e. x, $f * g \in L^p(\mathbb{R}^n)$, and $\|f * g\|_p \le \|f\|_1 \|g\|_p$.*

Proof. Setting $\Phi(x,y) = g(x-y)f(y)$,[6] then $\Phi(\cdot,y) \in L^p$ for each $y \in \mathbb{R}^n$ and we have $\int \|\Phi(\cdot,y)\|_p\, dy < \infty$. Now the conclusion follows directly from the Minkowski inequality for integration 12.5 (2), proposition 12.14 (1) and the translation invariance of Lebesgue measure. □

Lemma 12.16. *For $1 \le p < \infty$, $\tau_z : L^p \to L^p$ is an isometric isomorphism and for $f \in L^p$, $z \in \mathbb{R}^n \mapsto \tau_z f \in L^p$ is continuous.*

Proof. It easily follows from corollary 11.10. □

Proposition 12.17. *Suppose that $p, q \in [1,\infty]$ and p and q are conjugate exponents, $f \in L^p$ and $g \in L^q$, then $f * g \in BC(\mathbb{R}^n)$, $\|f * g\|_\infty \le \|f\|_p \|g\|_q$ and if $p, q \in (1,\infty)$ then $f * g \in C_0(\mathbb{R}^n)$.*

Proof. The existence of $f * g(x)$ and the estimate $|f * g|(x) \le \|f\|_p \|g\|_q$ for all $x \in \mathbb{R}^n$ is a simple consequence of Hölder inequality 12.4 and the translation invariance of Lebesgue measure. In particular this shows $\|f * g\|_\infty \le \|f\|_p \|g\|_q$. By relabeling p and q if necessary we may assume that $p \in [1,\infty)$. By Hölder 12.4 and lemma 12.16 we have

$$\|\tau_z(f * g) - f * g\|_\infty = \|\tau_z f * g - f * g\|_\infty \le \|\tau_z f - f\|_p \|g\|_q \to 0 \quad \text{as } z \to 0.$$

It follows that $f * g$ is uniformly continuous.

Finally if $p, q \in (1,\infty)$, setting $f_m = f\chi_{|f|\le m}$ and $g_m = g\chi_{|g|\le m}$, then by proposition 12.14 (4) and what we just proved, $f_n * g_n \in C_c(\mathbb{R}^n)$. Moreover,

$$\begin{aligned}
\|f * g - f_m * g_m\|_\infty &\le \|f * g - f_m * g\|_\infty + \|f_m * g - f_m * g_m\|_\infty \\
&\le \|f - f_m\|_p \|g\|_q + \|f_m\|_p \|g - g_m\|_q \\
&\le \|f - f_m\|_p \|g\|_q + \|f\|_p \|g - g_m\|_q \to 0 \quad \text{as } m \to \infty.
\end{aligned}$$

Hence $f * g \in C_0(\mathbb{R}^n)$ by proposition 11.21. □

Proposition 12.18. *Let $f \in C_c(\mathbb{R}^n)$ and $g \in L^1_{loc}(\mathbb{R}^n)$. Then $(f * g)(x)$ is well defined for every $x \in \mathbb{R}^n$, and $(f * g) \in C(\mathbb{R}^n)$.*

Proof. Transform the conclusion into a local estimate, and then use Hölder inequality 12.4 and lemma 12.16. □

Now we introduce the property of differentiation.

Proposition 12.19. *Suppose f and g are real-valued functions on $\mathbb{R}^d$. If $f \in C_c^k$ (resp. $f \in C^k$) and $g \in L^1_{loc}$ (resp. $g \in L^1$), then $f * g \in C^k$ and*

$$D^\alpha(f * g) = (D^\alpha f) * g, \quad \text{for all } \alpha \text{ with } |\alpha| \le k.$$

In particular, if $f \in C_c^\infty$ (resp. $f \in C^\infty$) and $g \in L^1_{loc}$ (resp. $g \in L^1$), then $f * g \in C^\infty$.

Proof. It immediately follows from theorem 2.27 of [For]. □

Finally we introduce the regularizations.

Definition 12.20 (mollifier). *If φ is a smooth function on $\mathbb{R}^n$, $n \ge 1$, satisfying the following three requirements*

[6] It will fail if we set $\Phi(x,y) = f(x-y)g(y)$.

(1) It is compactly supported;
(2) $\int_{\mathbb{R}^n} \varphi(x)\,dx = 1$;
(3) $\lim_{\varepsilon\to 0} \varphi_\varepsilon(x) = \lim_{\varepsilon\to 0} \varepsilon^{-n}\varphi(x/\varepsilon) = \delta(x)$;

where $\delta(x)$ *is the Dirac delta function and the limit must be understood in the space of Schwartz distributions, then* φ *is a* ***mollifier****. For example, we can define* $\eta \in C^\infty(\mathbb{R}^n)$ *by*

$$\eta(x) := \begin{cases} C\exp\left(\frac{1}{|x|^2-1}\right) & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1 \end{cases}$$

the constant $C > 0$ *selected so that* $\int_{\mathbb{R}^n} \eta(x)\,dx = 1$*. It's easy to see that* η *is a mollifier. We call* η *the* ***standard mollifier****.*

For each $\varepsilon > 0$ we set

$$\eta_\varepsilon(x) := \frac{1}{\varepsilon^n}\eta\left(\frac{x}{\varepsilon}\right)$$

Now we make smooth approximations of a locally integrable function using the mollifier and convolution.

Theorem 12.21 (Smooth approximation)**.** *Let* $U \subset \mathbb{R}^n$ *be an open subset, let* $\varepsilon > 0$*, and let* $f : U \to \mathbb{R}$ *be locally integrable. Set*

$$U_\varepsilon = \{x \in U : \operatorname{dist}(x, \partial U) > \varepsilon\}.$$

and

$$f^\varepsilon(x) := \eta_\varepsilon * f(x) = \int_U \eta_\varepsilon(x-y)f(y)\,dy = \int_{B(0,\varepsilon)} \eta_\varepsilon(y)f(x-y)\,dy \quad \text{in } U_\varepsilon.$$

Then we have

(1) $f^\varepsilon \in C^\infty(U_\varepsilon)$;
(2) $f^\varepsilon \to f$ *a.e. as* $\varepsilon \to 0$;
(3) If $f \in C(U)$*, then* $f^\varepsilon \to f$ *uniformly on compact subsets of* U;
(4) If $1 \leq p \leq \infty$*, and* $f \in L^p_{loc}(U)$*, then* $f^\varepsilon \to f$ *in* $L^p_{loc}(U)$*.*

Proof. One can refer to [Evab]. □

Corollary 12.22. *There are some density theorems.*

(1) $C_c^\infty(\mathbb{R}^n)$ *(and hence also* $\mathcal{S}$*) is dense in* $L^p(\mathbb{R}^n)$ *for* $1 \leq p \leq \infty$ *and in* $C_0(\mathbb{R}^n)$*.*
(2) Let $\Omega \subset \mathbb{R}^n$ *be an open subset. Then* $C_c^\infty(\Omega)$ *is dense in* $L^p(\Omega)$ *for any* $1 \leq p < \infty$*.*

Corollary 12.23. *Let* $\Omega \subset \mathbb{R}^d$ *be an open set and let* $u \in L^1_{loc}(\Omega)$ *be such that*

$$\int uf = 0 \quad \forall f \in C_c^\infty(\Omega).$$

Then $u = 0$ *a.e. on* Ω*.*

Proof. One can show that $\int ug = 0$ for any $g \in L^\infty(\Omega)$ through considering g^ε. Then for any compact subset K of Ω, taking g to be $\mathrm{sign}(u)$ on K and 0 otherwise, we get $u = 0$ a.e. on K. The conclusion follows from the arbitrariness of K. □

For more basic properties of L^p spaces, one can refer to [For] and [Jon].

13. Appendix A — Semicontinuity

Theorem 13.1. *A T_1-space X is perfectly normal iff for every lower (resp. upper) semicontinuous function f defined on X, there exists a sequence (f_i) of continuous real-valued function on X such that $f_i(x) \le f_{i+1}(x)$ (resp. $f_i(x) \ge f_{i+1}(x)$) for $i = 1, 2, \cdots$ and $x \in X$, and that $f(x) = \lim f_i(x)$ for every $x \in X$.*

Proof. One can refer to [Eng]. □

References

[Bre] Haim Brezis. *Funcional Analysis, Sobolev Spaces and Partial Differential Equations.*
[Eng] Ryszard Engelking. *General Topology.*
[Evaa] Lawrence C. Evans. *Measure theory and fine properties of functions.*
[Evab] Lawrence C. Evans. *Partial Differential Equations.*
[For] Gerald B. Forlland. *Real analysis.*
[Jon] Frank Jones. *Lebesgue integration on Euclidean space.*
[Mei] Jiaqiang Mei. *Mathematical analysis.*
[Rai] Armin Rainer. *Real analysis.*
[Tao] Terence Tao. *An epsilon of room: pages from year three of a mathematical blog.*
[Xio] Zhiyao Xiong. *Solutions of problems in lecture notes.*

Zhiyao Xiong, Department of Mathematics, Nanjing University, Nanjing, 210093, China
Email address: 181830203@smail.nju.edu.cn