DEFINITIONS, PROPOSITIONS AND THEOREMS

Real Analysis

Contents

1 Algebras of Sets

1.1 Definitions

(Boolean) algebra Omitted.

σ-algebra Omitted.

Borel algebra Defined only for topological spaces. The **Borel algebra** is the σ-algebra generated by all open sets, or equivalently, the intersection of all σ-algebra containing all open sets, or also, the smallest algebra containing all open sets. Note that for a topological space (X, τ) and its Borel sets $\mathscr{B}$, $\tau \subset \mathscr{B}$.

Note 1.1. This definition relies on the fact that any intersection of σ-algebras is still a σ-algebra.

Tensor product of σ-algebras Let $(X_1, \mathcal{A}_1)$ and $(X_2, \mathcal{A}_2)$ be measurable spaces. Define $\mathcal{R} = \{\text{All rectangles in } X_1 \times X_2\} = \{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}$ and $\widetilde{\mathcal{R}} = \{\text{Finite union of disjoint rectangles in } \mathcal{R}\}$. *Then $\widetilde{\mathcal{R}}$ is an algebra on $X_1 \times X_2$.* Then let $\mathcal{A}_1 \otimes \mathcal{A}_2$ denote the σ-algebra generated by $\mathcal{R}$ (or $\widetilde{\mathcal{R}}$), and call it the **tensor product** of $\mathcal{A}_1$ and $\mathcal{A}_2$.

Note 1.2. Need to verify that $\widetilde{\mathcal{R}}$ is an algebra. Let $A = \bigcup_{i=1}^{n}(A_i \times A_i')$, $B = \bigcup_{j=1}^{m}(B_j \times B_j') \in \widetilde{\mathcal{R}}$. Then

- Union:

$$A \cup B = \bigcup_{i=1}^{n}\bigcup_{j=1}^{m}((A_i \times A_i') \cup (B_j \times B_j'))$$

  As long as you deal with the union of two rectangles you can verify this. This is done by breaking it down to 7 disjoint parts.

- Intersection: this is much easier.

$$\begin{aligned} A \cap B &= (\bigcup_{i=1}^{n}(A_i \times A_i')) \cap (\bigcup_{i=1}^{m}(B_j \times B_j')) \\ &= \bigcup_{i=1}^{n}\bigcup_{j=1}^{m}((A_i \times A_i') \cap (B_j \times B_j')) \\ &= \bigcup_{i=1}^{n}\bigcup_{j=1}^{m}((A_i \cap B_j) \times (A_i' \cap B_j')) \end{aligned}$$

- Complement:

$$A^c = \left(\bigcup_{i=1}^{n} A_i \times A_i' \right)^c = \bigcap_{i=1}^{n} (A_i \times A_i')^c$$

It remains to break $A_i \times A_i'$ down into 8 disjoint parts.

Q 1.1. Why do I have to verify this? Why not just generate it?

2 Measures on $\mathbb{R}^n$ and Measurability of Sets

2.1 Definitions

Content of elementary sets Elementary sets in $\mathbb{R}$ are of the form

$$R = [a_1, b_1] \times \cdots \times [a_n, b_n]$$

Its **content** or **volume** is defined to be

$$|R| = (b_1 - a_1) \cdots (b_n - a_n)$$

Exterior measure For any subset $E \subset \mathbb{R}^n$, the **exterior measure** or **outer measure** of E is

$$m^*(E) = \inf_{\{Q_j\}} \sum_{j=1}^{\infty} |Q_j|$$

where the infimum is taken over all countable coverings $\{Q_j\}$ of E.

Lebesgue measure There are two equivalent conditions for Lebesgue measurability.

(i) (**Approximation by open sets**) A set E is measurable if

$$\forall \varepsilon > 0, \ \exists \text{ open set } \mathcal{O} \supset E, \ m^*(\mathcal{O} - E) \leqslant \varepsilon$$

(ii) (**The Carathéodory condition**) A set E is measurable if

$$m^*(T) = m^*(T \cap E) + m^*(T \cap E^c), \ \forall T \subset \mathbb{R}^n$$

The exterior measure restricted to Lebesgue measurable sets is call the **Lebesgue measure**.

Note 2.1. The notion of exterior measure and measurability can be generalized to any set. See Sect.3.1. But for now we restrict ourselves to the $\mathbb{R}^n$ case.

2.2 Propositions

Properties of the exterior measure

(i) (**Monotonicity**)

$$E_1 \subset E_2 \implies M^*(E_1) \leqslant m^*(E_2)$$

(ii) (**Subadditivity**)

$$m^*(\bigcup_{j=1}^{\infty} E_j) \leqslant \sum_{j=1}^{\infty} m^*(E_j)$$

(iii) (**Isolated additivity**)

$$d(E_1, E_2) > 0 \implies m^*(E_1 \cup E_2) = m^*(E_1) + m^*(E_2)$$

(iv) (**Countable additivity on almost disjoint cubes**)

$$\{Q_j\}_{j=1}^{\infty} \text{ are almost disjoint cubes} \implies m^*(\bigcup_{j=1}^{\infty} Q_j) = \sum_{j=1}^{\infty} |Q_j|$$

(v) (**Approximation by open sets**)

$$m^*(E) = \inf_{\substack{E \subset \mathcal{O} \\ \mathcal{O} \text{ is open}}} m^*(\mathcal{O})$$

Basic properties of measurable sets

(0) Null sets are measurable.

(i) Open sets are measurable.

(ii′) Compact sets are measurable.

(ii) Closed sets are measurable.

(iii) Measurable sets is closed under the taking of complement.

(iv) Measurable sets is closed under countable union.

(v) Measurable sets is closed under countable intersection.

(vi) The Lebesgue measure enjoys countable additivity.

(vii) Continuity on monotonic sequences.

The relations are illustrated in this diagram.

$$
\begin{tikzcd}
& & \text{null sets} & \text{C's lemma}^3 \arrow[ddl, Rightarrow] \\
& \text{Cond. } O \arrow[ur, Rightarrow] \arrow[r, Rightarrow] \arrow[d, Rightarrow] \arrow[ddr, Rightarrow] & \text{open sets} & \text{Cond. C} \arrow[ul, Rightarrow] \arrow[u, Rightarrow] \arrow[d, Rightarrow] \arrow[ddl, Rightarrow] \\
& \text{compact sets} \arrow[r, Rightarrow] & \text{closed sets} \arrow[u, Rightarrow] \arrow[r, Rightarrow] & \text{complement} \arrow[ul, Rightarrow] \arrow[d, Rightarrow] \\
& & \sigma\text{-union} \arrow[u, Rightarrow] \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \sigma\text{-intersection} \\
& & \sigma\text{-additivity} \arrow[d, Rightarrow] & \\
& & \text{continuity} &
\end{tikzcd}
$$

More properties of measurable sets I wanted to show the regularity, but got lazy.

3 Abstract Measure

3.1 Definitions

Exterior measure and measurability Let X be a set. An **exterior measure** is a function $\mu^* : 2^X \to [0, \infty]$ that satisfies

(i) $\mu^*(\varnothing) = 0$

(ii) $E_1 \subset E_2 \implies \mu^*(E_1) \leqslant \mu^*(E_1) \leqslant \mu^*(E_2)$

(iii) $\mu^*(\bigcup_{j=1}^{\infty} E_j) \leqslant \sum_{j=1}^{\infty} \mu^*(E_j)$

A set E is called **Carathéodory measurable** if

$$\mu^*(E) = \mu^*(E \cap T) + \mu^*(E^c \cap T), \ \forall T \in 2^X$$

Simple function A function $f : X \to \mathbb{R}$ is **simple** if $f(X)$ is a finite set.

Measurable space A **measurable space** is a set X equipped with a σ-algebra $\mathcal{F}$. More precisely, we call the tuple $(X, \mathcal{F})$ a **measurable space**. The sets in $\mathcal{F}$ are called **measurable sets**.

Measure Let $(X, \mathcal{F})$ be a set X equipped with an algebra $\mathcal{F}$. An **additive function** $\mu : \mathcal{F} \to [0, \infty]$ on $\mathcal{F}$ is a function that satisfies

(i) $\mu(\varnothing) = 0$;

(ii) finite additivity on disjoint sets in $\mathcal{A}$.

If furthermore, the algebra is a σ-algebra and the additive function is σ-additive (countably additive). Then we have a **measure**, formalized as follows.

(i) A **positive measure** is a function $\mu : \mathcal{F} \to [0, \infty]$ which is **countably additive**, and $\mu(A) < \infty$ for at least one set A.

(ii) A **measure space** is a measurable space equipped with a positive measure on its measurable sets. More precisely, the triple $(X, \mathcal{F}, \mu)$ is called the **measure space**.

Note 3.1. Note that requiring $\mu(A) < \infty$ for some A is equivalent to saying $\mu(\varnothing) = 0$.

If a measure space $(X, \mathcal{F}, \mu)$ is such that $\mu(X) < \infty$, we say μ is **finite**. If there exists an ascending chain A_n of measurable sets such that $A_n \nearrow E$, $\mu(A_n) < \infty$, we say μ is **σ-finite**.

The following are some important examples of measures.

(i) (**Counting measure**) Let X be any set. The **counting measure** μ on $(X, 2^X)$ is defined by $\mu(E) = |E|$, $E \subset X$.

Note 3.2. The counting measure, when appearing in an integral, means "sum". For example, let μ be the counting measure on $(\mathbb{N}, 2^{\mathbb{N}})$, and we may consider a sequence $\{a_n\}$ as a measurable function $a : n \mapsto a_n$, the integral over a subset is then the sum over those indexes

$$\int_{E \subset \mathbb{N}} a \, \mathrm{d}\mu = \sum_{i \in E} a_i$$

(ii) (**probability measure**) Let μ be a measure on X. μ is a **probability measure** if we require in extra $\mu(X) = 1$.

3.2 Propositions

Simple function Let $f : X \to \mathbb{R}$ be a simple function on a measurable space X. f is measurable $\iff$ for any $y \in f(X)$, $f^{-1}(y)$ is measurable.

Positive measure Let μ be a positive measure on $(X, \mathcal{F})$. Then satisfies the following without any other restrictions than its definition.

(i) **On the empty set**: $\mu(\varnothing) = 0$;

(ii) **Finite additivity**: $A_1, \cdots, A_n \in \mathcal{F}$ are disjoint $\implies$ $\mu(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)$;

(iii) **Monotonicity**: $A \subset B \implies \mu(A) \leqslant \mu(B)$;

(iv) **Continuity on monotone sequences**:

$$A_n \nearrow A \implies \mu(A_n) \to \mu(A)$$

$$A_n \searrow A,\ \mu(A_1) < \infty \implies \mu(A_n) \to \mu(A)$$

3.3 Theorems

Carathéodory's Extension Theorem Let $(X, \mathcal{A}, \mu_0)$ be a set equipped with an algebra and an σ-finite additive function on that algebra.

(i) **Uniqueness, should an extension exist.** If there exists an extension μ of μ_0 that is a measure on $(X, \sigma(\mathcal{A}))$, then μ is unique.

(ii) **Existence.** First we define two conditions.

- **Condition** C: for all chains $\{A_n\} \subset \mathcal{A}$ such that $\mu_0(A_1) < \infty$, $A_n \searrow \varnothing$, we have $\mu_0(A_n) \to 0$;
- **Condition** C_∞: there exists an ascending chain $\{A_n\} \subset \mathcal{A}$ such that $A_n \nearrow X$, and that

$$A \in \mathcal{A},\ \mu_0(A) = \infty \implies \mu_0(A \cap A_n) \to \infty$$

The existence of an extension are considered in two cases.

- If μ_0 is finite, then C is sufficient.
- If $\mu_0(X) = \infty$, then both C and C_∞ are required.

Note 3.3. The idea of the proof is to show that the set on which two extensions agree is a monotone class.

Proof. Let μ, ν be two extensions of μ_0. If μ_0 is finite, let

$$\mathcal{M} = \{M \in \sigma(\mathcal{A}) : \mu(M) = \nu(M)\}$$

Then $\mathcal{M}$ is a monotone class containing $\mathcal{A}$ (The "finite" condition is used to verify the descending condition of a monotone class, or one can also show that $\mathcal{M}$ is closed under the taking of complement). Hence $\mathcal{M} \subset \sigma(\mathcal{A}) \implies \mathcal{M} = \sigma(\mathcal{A}) \implies \mu = \nu$ on $\sigma(\mathcal{A})$.
If μ_0 is not finite, the measure of the complement of a set may be infinity. Thus we need intermediate sets. Let

$$\mathcal{M} = \{M \in \sigma(\mathcal{A}) : \forall A \in \mathcal{A},\ \nu_0(A) < \infty,\ \mu(A \cap M) = \nu(A \cap M)\}$$

It is still a monotone class containing $\mathcal{A}$. One can then choose a chain $\{A_n\}$ ascending to X in place of A and reach the conclusion.

Note 3.4. Note that any probability measure is finite.

Q 3.1. What about condition C. Wu never mentioned it.

Littlewood's Three Principles In $\mathbb{R}^n$, we have

(i) *Every set of finite measure is nearly a finite union of intervals.* If $E \subset \mathbb{R}^n$ is such that $m(E) < \infty$, there exists a finite union $F = \bigcup_{i=1}^{N} Q_i$ of closed cubes such that $m(E \Delta F) \leqslant \varepsilon$.

(ii) (**Egorov**) *Every convergent sequence of measurable functions is nearly uniform convergent.* Let $\{f_k : E \to \overline{\mathbb{R}}\}_{k=1}^{\infty}$ is a sequence of measurable functions defined on a measurable set E with $m(E) < \infty$, and $f : E \to \mathbb{R}$ also a measurable function. Assume that $f_k \to f$ a.e. on E. Given $\varepsilon > 0$, we can find a closed set $A_\varepsilon \subset E$ such that $m(E - A_\varepsilon) \leqslant \varepsilon$ and $f_k \rightrightarrows f$ on A_ε.

Note 3.5. One proof is to consider the sets

$$E_s(p) = \bigcup_{k \geqslant s} \{|f_k - f| \geqslant 1/p\}$$

and prove that $\forall p \geqslant 1$, $\lim_s m(E_s(p)) = 0$. In this way for each p there is a s_p such that $m(E_{s_p}(p)) \leqslant \varepsilon/2^p$. Then f_k converges uniformly on its complement in E. Note that the uniformity lies in the fact that n_p is independent of x.

(iii) (**Lusin**) *Every measurable function is nearly continuous.* If f is measurable and $|f| < \infty$ on a measurable set E. Then $\forall \varepsilon > 0$, $\exists F_\varepsilon$ that is closed such that $F_\varepsilon \subset E$, $m(E - F_\varepsilon) \leqslant \varepsilon$ and that $f|_{F_\varepsilon}$ is continuous.

Note 3.6. If $m(E) < \infty$, you may use Egorov's theorem.

The Borel-Cantelli Lemma Let $(X, \mathcal{F}, \mu)$ be a measure space and $\{A_n \in \mathcal{F}\}$ a sequence of measure sets. Then

$$\sum_{n=1}^{\infty} \mu(A_n) < \infty \implies \mu\left(\limsup_n A_n\right) = 0$$

4 Modes of Convergence

4.1 Definitions

We may assume the following measure spaces are complete in order to use null sets.

L^p space Let $(X, \mathcal{F}, \mu)$ be a measure space.

(i) If $0 < p < \infty$ and $f : X \to \mathbb{C}$ is measurable, then we define

$$\|f\|_p = \left(\int_X |f|^p \, \mathrm{d}\mu\right)^{1/p}$$

to be the **L^p-norm** of f.

(ii) Let

$$L^p(X, \mathcal{F}, \mu) = \left\{ f : X \to \mathbb{C} \;\middle|\; \int_X |f|^p \, \mathrm{d}\mu < \infty \right\}$$

and call it the L^p-space of $(X, \mathcal{F}, \mu)$.

Convergence: pointwise Let X be a set and (Y, τ) a topological space. Let $\{f_n : X \to Y\}$ and $f : X \to Y$ be functions from X to Y. Then we say f_n **converges pointwise** to f if

$$\forall x \in X, \; \lim_n f_n(x) = f(x)$$

Convergence: uniform Let X be a set and (Y, d) a metric space. Let $\{f_n : X \to Y\}$ and $f : X \to Y$ be functions from X to Y. Then we say f_n **converges uniformly** to f if the convergence is uniform with respect to x. In formal language:

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \; \forall x \in X, \; \forall n > N, \; d(f_n(x), f(x)) < \varepsilon$$

Convergence: pointwise, almost everywhere Let $(X, \mathcal{F}, \mu)$ be a measure space and (Y, τ) a topological space. Let $\{f_n : X \to Y\}$ and $f : X \to Y$ be functions from X to Y. Then we say f_n **converges (pointwise) to f almost everywhere** if

$$\exists \mathcal{N} \in \mathcal{F} : \mu(\mathcal{N}) = 0,\ \forall x \in X \backslash \mathcal{N},\ \lim_n f_n(x) = f(x)$$

Convergence: uniform, almost everywhere Let $(X, \mathcal{F}, \mu)$ be a measure space and (Y, d) a metric space. Let $\{f_n : X \to Y\}$ and $f : X \to Y$ be functions from X to Y. Then we say f_n **converges uniformly almost everywhere** if after removing a null set in X, f_n converges uniformly to f. In formal language:

$$\exists \mathcal{N} \in \mathcal{F} : \mu(\mathcal{N}) = 0,\ \forall \varepsilon > 0,\ \exists N \in \mathbb{N},$$
$$\forall n > N,\ \forall x \in X \backslash \mathcal{N},\ d(f_n(x), f(x)) < \varepsilon$$

Convergence: almost uniform Let $(X, \mathcal{F}, \mu)$ be a measure space and (Y, d) a metric space. Let $\{f_n : X \to Y\}$ and $f : X \to Y$ be functions from X to Y. Then we say f_n **converges almost uniformly** if one can remove a set of arbitrarily small measure such that f_n converges uniformly in the remaining part. In formal language:

$$\forall \varepsilon > 0,\ \exists F_\varepsilon \subset X,\ \mu(X - F_\varepsilon) < \varepsilon, f_n \rightrightarrows f \text{ on } F_\varepsilon$$

Convergence: in L^p norm That is, convergence in L^p as a metric space. Let $f_n, f \in L^p(X, \mathcal{F}, \mu)$. We say that f_n **converges to f in L^p** if

$$\lim_n \|f_n - f\|_p = 0$$

Convergence: in L^∞ norm That is, uniformly almost everywhere when $Y = \mathbb{C}$.

Convergence: in measure Let $(X, \mathcal{F}, \mu)$ be a measure space and (Y, d) a metric space. Let $\{f_n : X \to Y\}$ and $f : X \to Y$ be functions from X to Y. Then we say f_n **converges to f in measure (μ)** if

$$\forall \varepsilon > 0,\ \lim_n \mu(\{d(f_n, f) \geqslant \varepsilon\}) = 0$$

Note 4.1. Note that for a sequence of functions to converge, Y must be a topological space no matter what; for we to use "almost", "in measure", X must be a measure space; for we to use "uniform", "norm", Y must be a metric space.

5 Abstract Integration

5.1 Definitions

Lebesgue integral Let $(X, \mathcal{F}, \mu)$ be a measure space. The definition of the Lebesgue integral for functions $X \to [-\infty, \infty]$ are defined through the following stages.

(i) **Simple, positive, measurable**: Let $s : X \to [0, \infty)$ be a simple function that is measurable. We may write

$$s(x) = \sum_{i=1}^{n} a_i \chi_{A_i}(x)$$

with the a_i's distinct. The **Lebesgue integral of** s **over** $E \in \mathcal{F}$ **with respect to** μ is defined as

$$\int_E s \,\mathrm{d}\mu = \sum_{i=1}^{n} a_i \mu(A_i \cap E)$$

Note 5.1. We need the convention that $0 \cdot \infty = 0$ in case $\mu(A_i \cap E) = \infty$ and $a_i = 0$.

Note 5.2. The basic idea is

$$\text{integral} = \text{value} \times \text{measure of its preimage}$$

Note 5.3. Note that the integral is defined for all positive, measurable functions, we have not yet encountered situations in which functions are measurable but not integrable.

(ii) **Positive, measurable**: Let $f : X \to [0, \infty]$ be measurable. The **Lebesgue integral of** f **over** $E \in \mathcal{F}$ **with respect to** μ is defined as

$$\int_E f \,\mathrm{d}\mu = \sup_{\substack{0 \leqslant s \leqslant f \\ s \text{ is simple}}} \int_E s \,\mathrm{d}\mu$$

Note 5.4. One need to verify that the second definition is consistent with the first for simple functions.

(iii) **Measurable**: Let $f : X \to [-\infty, \infty]$ be measurable. The **Lebesgue integral of** f **over** $E \in \mathcal{F}$ **with respect to** μ is defined as

$$\int_E f \,\mathrm{d}\mu = \int_E f^+ \,\mathrm{d}\mu - \int_E f^- \,\mathrm{d}\mu$$

if at least one term on the right is finite. Otherwise the integral does not exist.

L^1**: Lebesgue integrable, summable** Let $(X, \mathcal{F}, \mu)$ be a measure space. $L^1(X, \mathcal{F}, \mu)$ is the set of functions $f : X \to \mathbb{R}$ (or more generally $X \to \mathbb{C}$) satisfying $\int_X |f| \, \mathrm{d}\mu < \infty$.

5.2 Propositions

Integration of positive functions The following are immediate consequences of the definitions. The functions and sets occurring below are assumed to be measurable.

(i) **Monotonicity on functions**: $0 \leqslant f \leqslant g \implies \int_E f \, \mathrm{d}\mu \leqslant \int_E g \, \mathrm{d}\mu$;

(ii) **Monotonicity on domains**: $0 \leqslant f,\ A \subset B \implies \int_A f \, \mathrm{d}\mu \leqslant \int_B f \, \mathrm{d}\mu$;

(iii) **Scalar multiplication**: $0 \leqslant f,\ 0 \leqslant c < \infty \implies \int_E cf \, \mathrm{d}\mu = c \int_E f \, \mathrm{d}\mu$;

(iv) **Integral of** 0: $\int_E 0 \, \mathrm{d}\mu = 0$;

(v) **integral on a null set**: $\mu(E) = 0 \implies \int_E f \, \mathrm{d}\mu = 0$;

(vi) **Zero extension**: $0 \leqslant f \implies \int_E f \, \mathrm{d}\mu = \int_X f \chi_E \, \mathrm{d}\mu$.

(vii) **Weighted new measure**: Let s be nonnegative, measurable, simple. Then $\varphi : \mathcal{F} \to [0, \infty]$ defined by

$$\varphi(E) = \int_E s \, \mathrm{d}\mu$$

is a measure on $\mathcal{F}$.

(viii) σ**-additivity on sets**: Let s be nonnegative, measurable, simple. Let A_n be countably many disjoint measurable sets. Then

$$\int_{\cup_i A_i} s \, \mathrm{d}\mu = \sum_{i=1}^{n} \int_{A_i} s \, \mathrm{d}\mu$$

(ix) **Continuity on sets**: Let s be nonnegative, measurable, simple. Let A_n be an ascending or descending chain of measurable sets (in the second case we require $\mu(A_1) < \infty$) converging to A. Then

$$\lim_n \int_{A_n} s \, \mathrm{d}\mu = \int_A s \, \mathrm{d}\mu$$

(x) **Linearity on simple functions**: Let s and t be nonnegative, measurable, simple. Then

$$\int_X (s+t)\mathrm{d}\mu = \int_X s\,\mathrm{d}\mu + \int_X t\,\mathrm{d}\mu$$

(xi) **Continuity of functions, linearity on general positive measurable functions**: See convergence theorems and their corollaries.

5.3 Theorems

Lebesgue's Monotone Convergence Theorem Let $(X, \mathcal{F}, \mu)$ be a measure space. Let $\{f_n\}$ be a sequence of measurable functions on X such that $\forall x \in X,\ 0 \leqslant f_1(x) \leqslant f_2(x) \leqslant \cdots \leqslant \infty$. Then $f = \lim_n f_n$ exists and is measurable, and

$$\lim_n \int_X f_n\,\mathrm{d}\mu = \int_X f\,\mathrm{d}\mu$$

In brief, **nonnegative, monotone, pointwise convergence implies the convergence of integral**.

Note 5.5. The fact that f is measurable (and hence integrable for being positive) and one direction of the equality are easy to prove. The other direction is by exhibiting that the integral of f_n can be arbitrarily close to that of f.

Proof. $\int_X f_n\,\mathrm{d}\mu$ converges to some α because it is monotone. Meanwhile $\int_X f_n \mathrm{d}\mu \leqslant \int_X f\,\mathrm{d}\mu \implies \alpha \leqslant \int_X f\,\mathrm{d}\mu$. For the converse, let $E_n(c) = \{f_n \geqslant cf\}$ for $0 < c < 1$. Then we have $E_n \nearrow X$ and $\int_X f_n\,\mathrm{d}\mu \geqslant \int_{E_n} f_n\,\mathrm{d}\mu \geqslant c\int_{E_n} f\,\mathrm{d}\mu$. Pass to the limit $n \to \infty$ and the $c \to 1$ and we obtain $\alpha \geqslant \int_X f\,\mathrm{d}\mu$, whence $\alpha = \lim_n \int_X f_n\,\mathrm{d}\mu = \int_X f\,\mathrm{d}\mu$.

Corollary 1: additivity on functions If f_1, f_2 are nonnegative, measurable, then

$$\int_X (f_1+f_2)\,\mathrm{d}\mu = \int_X f_1\,\mathrm{d}\mu + \int_X f_2\,\mathrm{d}\mu$$

Proof. Take two sequences of simple functions $\{s_i\}, \{s'_j\}$ that converge monotonically and pointwise to f_1, f_2 respectively. Notice that $s_i + s'_i \to f_1 + f_2$ and use linearity for simple functions.

Corollary 2 If $a_{ij} \geqslant 0$, then $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}$.

Fatou's Lemma Let $(X, \mathcal{F}, \mu)$ be a measure space. Let $f_n : X \to [0, \infty]$ be measurable and nonnegative. Then

$$\int_X \left(\liminf_n f_n\right) \mathrm{d}\mu \leqslant \liminf_n \int_X f_n \, \mathrm{d}\mu$$

Proof. Use the inequality $(\inf_{k \geqslant n} f_k) \leqslant f_m$ for $m \geqslant n$. Integrate it.

Lebesgue's Dominated Convergence Theorem Let $(X, \mathcal{F}, \mu)$ be a measure space. Let $\{f_n\}$ be a sequence of measurable functions on X such that $\lim_n f_n = f$ for some f. If there exists $g \in L^1(X, \mathcal{F}, \mu)$, such that

$$|f_n| \leqslant g,$$

then $f \in L^1$, and

$$\lim_n \int_X f_n \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mu$$

Fubini's Theorem Let $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \lambda)$ be σ-finite measure spaces, and let f be an $(\mathcal{S} \times \mathcal{T})$-measurable function on $X \times Y$.

(i) (**Real, positive case**) If $0 \leqslant f \leqslant \infty$, and if

$$\varphi(x) = \int_Y f_x \, \mathrm{d}\lambda, \quad \psi(y) = \int_X f^y \, \mathrm{d}\mu \quad (x \in X, y \in Y)$$

then they are respectively $\mathcal{S}$-measurable and ψ-measurable, and

$$\int_X \varphi \, \mathrm{d}\mu = \int_{X \times Y} f \, \mathrm{d}(\mu \times \lambda) = \int_Y \psi \, \mathrm{d}\lambda$$

(ii) (**Complex case**)

(iii) (**L^1 case**)

6 Topics

6.1 The Cantor Set

In this section we talk about various kinds of Cantor sets and functions on them.

Cantor set of constant dissection Starting from $[0, 1]$, at each step remove the central part (open) of *relative length* ξ of each remaining interval. What is left in the end is the Cantor set $\mathcal{C}_\xi$. The case in which $\xi = 1/3$ is the most common one. The Cantor set has the following properties.

(i) Its complement is a countable union of open intervals with total length 1.

(ii) $\mathcal{C}_\xi$ is compact.

(iii) $\mathcal{C}_\xi$ is totally disconnected.

(iv) $\mathcal{C}_\xi$ is perfect.

(v) $m^*(\mathcal{C}_\xi) = 0$

(vi) $\mathcal{C}_\xi$ is uncountable.

Cantor-like sets Construct a closed set $\hat{\mathcal{C}}$ by starting from $[0,1]$ and removing at the kth stage 2^{k-1} centrally located open intervals of length l_k of each remaining interval (assuming l_k are small enough). The Cantor-like set has the following properties.

(i) If $\sum_{k=1}^{\infty} 2^{k-1} l_k < 1$, then $m(\hat{\mathcal{C}}) = 1 - \sum_{k=1}^{\infty} 2^{k-1} l_k > 0$.

(ii) $\hat{\mathcal{C}}$ is totally disconnected.

(iii) $\hat{\mathcal{C}}$ is perfect.

An important property that a Cantor set has is that each point of corresponds to a binary sequence (the kth digit in the sequence denotes "left" or "right" of the choice at the kth stage). This allows us to define functions on Cantor sets. The standard **Cantor-Lebesgue function** on $\mathcal{C} = \mathcal{C}_{1/3}$ is defined by expanding each $x \in \mathcal{C}$ into a ternary form with coefficients $= 0, 2$:

$$x = \sum_{k=1} a_k 3^{-k}, \quad a_k = 0, 2$$

and then setting

$$F(x) = \sum_{k=1}^{\infty} \frac{a_k}{2} 2^{-k}$$

The function $F : \mathcal{C} \to [0,1]$ is surjective, increasing. at each point $x \in [0,1] \backslash \mathcal{C}$, let (a,b) denote the maximal open interval not intersecting $\mathcal{C}$ but containing x, then we have $F(a) = F(b)$, and by defining F to be constant on that interval we can extend F to $F : [0,1] \to [0,1]$ that is monotonically increasing, surjective.

In the above example, the reason why $F(a) = F(b)$ is that a, b corresponds to binary sequences of the form

$$a \sim \cdots 01111111111111 \cdots$$

$$b \sim \cdots 10000000000000 \cdots$$

by interpreting them as the fraction part of a number they are equal. However, if we preserve this information we may obtain a bijective function from $\mathcal{C}$ to $2^{\mathbb{N}}$, and hence to any other Cantor set $\mathcal{C}'$. The latter function $\mathcal{C} \to \mathcal{C}'$ is strictly increasing, hence for any maximal interval (a, b) in the complement of $\mathcal{C}$ in $[0, 1]$, we have $F(a) < F(b)$. Then by setting F to be linear on that interval, we have a strictly increasing, continuous, bijective function $F : [0, 1] \to [0, 1]$ such that $F(\mathcal{C}) = \mathcal{C}'$.

6.2 Non-measurable Sets

The set $\mathcal{N}$ constructed by choosing exactly one[4] representative from each equivalence class in $[0, 1]/\mathbb{Q}$ is non-measurable. To see this, first note that each translate of $\mathcal{N}$ by a rational number is unique. Then we may take the union (countable, because rational numbers are numerable) of some translations close enough to get a set of reasonable size. This union is then countable and disjoint, to which we can apply the σ-additivity such and obtain a contradiction.

For any measurable subset E of $\mathcal{N}$ we use the same technique to show that $m(E) = 0$. First we enumerate all rationals in $[0, 1]$ by r_k and let $E_k = \mathcal{N} + r_k$, then we have

$$\bigcup_{k=1}^{\infty} E_k \subset [0, 2]$$

as a countable, disjoint union of measurable sets E_k. Then

$$0 \leqslant m(\bigcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} (E_k) = \sum_{k=1}^{\infty} (E) \leqslant 2 \implies m(E) = 0$$

The existence of non-measurable sets is more common than one might think. Actually, in any set of positive (exterior) measure there is a non-measurable subset. The construction is exactly like before. For details see the lecture notes by Xuguang Lu, or read the text by Minqiang Zhou.

6.3 Construction of the Integration and Convergence Theorems

6.3.1 Stein's Way

The construction requires approximation by simple functions. We begin by giving some propositions on these approximations.

- If f is nonnegative and measurable, then one can choose an increasing sequence of nonnegative, simple functions converging to f.

[4] Acknowledging the axiom of choice

- If f is measurable, then one can choose a sequence of simple functions converging to f while increasing in its absolute value.
- If f is measurable, then there exists a sequence of step functions converging to f a.e..

Stein constructs the integration for different functions in the following order:

- Simple functions (**bounded convergence for simple functions** is introduced here).
- Bounded functions supported on a set of finite measure (**bounded convergence for functions supported on a set of finite measure** and the **consistency with the proper Riemann integral** are introduced here).
- Nonnegative functions (**Fatou's lemma**, **Beppo-Levi's theorem** and the **monotone convergence theorem** are introduced here).
- General complex-valued functions (**absolute continuity** is introduced here).

At each step, one needs to verify the well-definedness and the following properties of the integral.

- Linearity on the integrand.
- Additivity on the domain of integration.
- Monotonicity.
- Triangle inequality.

Along the construction, the following convergence theorems arise.

Bounded convergence theorem for simple functions Let $\{\varphi_n\}$ be simple functions, bounded by M, supported on E with $m(E) < \infty$. Then

$$\varphi_n \to f \text{ a.e.} \implies f \text{ is measurable, bounded by } M\text{, supported on } E$$

$$\varphi_n \to f \text{ a.e.} \implies \lim_{n\to\infty} \int \varphi_n \text{ exists.}$$

$$\varphi_n \to 0 \text{ a.e.} \implies \lim_{n\to\infty} \int \varphi_n = 0$$

Bounded convergence theorem Let $\{f_n\}$ be measurable functions, bounded by M, supported on E with $m(E) < \infty$. Then

$$f_n \to f \text{ a.e.} \implies f \text{ is measurable, bounded by } M, \text{ supported on } E$$

$$f_n \to f \text{ a.e.} \implies \|f_n - f\|_{L^1} = \int |f_n - f| \to 0 \implies \lim_{n\to\infty} \int f_n \to \int f$$

Fatou's lemma For nonnegative measurable functions $\{f_n\}$, $f_n \geqslant 0$,

$$\int \liminf_{n\to\infty} f_n \leqslant \liminf_{n\to\infty} \int f_n$$

Dominated convergence theorem for nonnegative functions For nonnegative measurable functions $\{f_n\}$ and f, if $f_n \leqslant f$ and $f_n \to f$ a.e., then

$$\lim_{n\to\infty} \int f_n = \int f$$

Monotone convergence theorem for nonnegative functions For nonnegative measurable functions $\{f_n\}$, if $f_n \nearrow f$ a.e., then

$$\lim_{n\to\infty} \int f_n = \int f$$

Dominated convergence theorem for general functions For measurable functions $\{f_n\}$, if $f_n \to f$ a.e. and $|f_n| \leqslant g$ with g integrable, then

$$\|f_n - f\| = \int |f_n - f| \to 0$$

and

$$\lim_{n\to\infty} \int f_n = \int f$$

Having known our objectives, we shall begin to construct the integration.

Simple functions The integral of a simple function φ is defined as the sum of its possible values weighted by the measures of the sets corresponding to those values. In other words, the integral of a simple function must be explicitly written in terms of its **canonical representation**. However, one can prove that we can use any representation of φ to calculate $\int \varphi$. This done by decomposing the sets in the arbitrary representation into disjoint ones. (Q: how?) As to the properties of the integral, note that the integral of a simple function is essentially a sum, and the equalities and inequalities can be easily obtained.

Bounded functions supported on a set of finite measure The integral of a bounded function f supported on a set E of finite measure is defined as the limit of any simple sequence $\{\varphi_n\}$ converging to it, while having the same bound and supported on E themselves. One needs to verify that such sequence exists, this limit exists, and is independent of the choice of the sequence.

Proof. $\{\varphi_n\}$ exists as a consequence of f begin measurable. $\lim_{n\to\infty} \int \varphi_n$ exists by Egorov's theorem. The "almost uniform" convergence allows us to apply the Cauchy's criterion to the sequence $\{\int \varphi_n\}$. The limit being independent of the sequence follows from $\varphi_n \to 0 \implies \int \varphi_n \to 0$, which also can be obtained by using Egorov's theorem.

Nonnegative functions The integral of a nonnegative measurable function f is defined as the supremum of $\int g$ taken over all measurable $0 \leqslant g \leqslant f$ with g bounded and $m(\mathrm{supp}(g)) < \infty$. TBA.

7 Integration over $\mathbb{R}^n$

7.1 Properties

Arithmetic properties Linearity on the integrand, additivity on the domain, monotonicity, triangle inequality.

Integrable and finiteness $f \geqslant 0,\ \int f < \infty \implies f < \infty$ a.e.

Having norm zero $f \geqslant 0,\ \int f = 0 \implies f = 0$ a.e.

Absolute continuity Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be measurable.

(i) (**At infinity**) If f is integrable, then $\forall \varepsilon > 0,\ \exists R > 0$ such that

$$\int_{B(0,R)^c} |f| < \varepsilon$$

(i′) If f is integrable, then

$$\lim_{\alpha\to\infty} \int_{|f|\geqslant\alpha} |f| = 0$$

(i″) If f is integrable, then

$$\lim_{\alpha\to\infty} \int_{|x|\geqslant\alpha} |f| = 0$$

(ii) (**Within** $\mathbb{R}^n$) If f is integrable, then $\forall \varepsilon > 0$, $\exists \delta < 0$ such that

$$m(E) < \delta \implies \int_E |f| < \varepsilon$$

Proof. Truncate f in the proper way and apply MCT. For the second, the trouble is that $|f|$ may not be bounded, so we use a bounded function to approximate it (truncating the range).

Invariance properties If $f \in L^1(\mathbb{R}^n)$, $\delta = (\delta_1, \cdots, \delta_n)$ with $d_i > 0$, then

(i) (Translation)

$$\int f(x-h)\mathrm{d}x = \int f(x)\mathrm{d}x$$

(ii) (Dilation)

$$\int f(\delta x)\mathrm{d}x = \frac{1}{\delta_1 \cdots \delta_n} \int f(x)\mathrm{d}x$$

(iii) (Reflection)

$$\int f(-x)\mathrm{d}x = \int f(x)\mathrm{d}x$$

Proof. First prove for simple functions and then generalize to all functions.

Continuity under translation Let $f \in L^1$ and $f_h(x) = f(x-h)$. Then

$$\lim_{h \to 0} \|f_h - f\| = 0$$

Proof. Prove this for C_K first and approximate any integrable function f using some $g \in C_K$.

7.2 Convergence Theorems

Bounded convergence theorem for simple functions Let $\{\varphi_n\}$ be simple functions, bounded by M, supported on E with $m(E) < \infty$. Then

$$\varphi_n \to f \text{ a.e.} \implies f \text{ is measurable, bounded by } M \text{, supported on } E$$

$$\varphi_n \to f \text{ a.e.} \implies \lim_{n\to\infty} \int \varphi_n \text{ exists.}$$

$$\varphi_n \to 0 \text{ a.e.} \implies \lim_{n\to\infty} \int \varphi_n = 0$$

Proof. Egorov's theorem and Cauchy's criterion.

Bounded convergence theorem Let $\{f_n\}$ be measurable functions, bounded by M, supported on E with $m(E) < \infty$. Then

$$f_n \to f \text{ a.e.} \implies f \text{ is measurable, bounded by } M \text{, supported on } E$$

$$f_n \to f \text{ a.e.} \implies \|f_n - f\|_{L^1} = \int |f_n - f| \to 0 \implies \lim_{n\to\infty} \int f_n \to \int f$$

Proof. Same as above.

Fatou's lemma For nonnegative measurable functions $\{f_n\}$, $f_n \geqslant 0$,

$$\int \liminf_{n\to\infty} f_n \leqslant \liminf_{n\to\infty} \int f_n$$

Proof. You know.

Dominated convergence theorem for nonnegative functions For measurable functions $\{f_n\}$ and f, if $0 \leqslant f_n \leqslant f$ and $f_n \to f$ a.e., then

$$\lim_{n\to\infty} \int f_n = \int f$$

Proof. An obvious direction + Fatou's lemma.

Monotone convergence theorem for nonnegative functions For nonnegative measurable functions $\{f_n\}$, if $f_n \nearrow f$ a.e., then

$$\lim_{n\to\infty} \int f_n = \int f$$

Proof. This is either a special case of the above, or can be proved using Beppo-Levi's method.

Dominated convergence theorem for general functions For measurable functions $\{f_n\}$, if $f_n \to f$ a.e. and $|f_n| \leqslant g$ with g integrable, then

$$\|f_n - f\| = \int |f_n - f| \to 0$$

and

$$\lim_{n\to\infty} \int f_n = \int f$$

Proof. Note that $|f_n| \leqslant g$ a.e. $\implies |f| \leqslant g$ a.e. $\implies |f_n - f| \leqslant 2g$. Then use the absolute continuity of $\int g$ at infinity and BCT.

8 The space L^1

The space $\mathcal{L}^1$ is defined as follows.

$$\mathcal{L}^1 = \{f : \mathbb{R}^n \to \overline{\mathbb{C}}\}$$

Then $\mathcal{L}^1$ is a vector space over $\mathbb{C}$. If moreover, $\mathcal{L}^1$ is given the norm $\|f\| = \int |f|$, then $\mathcal{L}^1$ is a normed vector space. Consider its subspace

$$N = \{f \in \mathcal{L}^1 : f = 0 \text{ a.e.}\}$$

and the quotient

$$L^1 = \mathcal{L}^1/N$$

Then the latter is a Banach space with a linear structure and a well-defined norm inherited from $\mathcal{L}^1$. The well-definedness follows from $f = g$ a.e. $\iff \|f - g\| = 0$. Other properties that need to be verified are listed as follows.

- $\|\cdot\|$ is a norm.
- (**Riesz-Fischer**) L^1 is complete under that norm.

8.1 Properties

Banach space L^1 is a Banach space.

Dense families Simple functions, step functions and continuous functions of compact support are dense families in L^1.

Cvg in L^1 and a.e. Convergence in L^1 implies convergence a.e. along a subsequence.

9 Fubini's Theorem

9.1 The Theorem

The Fubini theorem and its extended form are as follows.

Fubini's theorem Let $f(x, y)$ be integrable on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then

(i) For a.e. y, f^y is[5] integrable on $\mathbb{E}^{d_1}$.

(ii) For a.e. y, $\int f^y(x)\mathrm{d}x$ is integrable on $\mathbb{R}^{d_2}$.

[5] The slice f^y means to fix some y and see it as a function of x.

(iii)
$$\int_{\mathbb{R}^{d_2}}\left(\int_{\mathbb{R}^{d_1}} f(x,y)\mathrm{d}x\right)\mathrm{d}y=\int_{\mathbb{R}^{d_1+d_2}} f(x,y)\mathrm{d}x\mathrm{d}y$$

(iv) Everything above holds if we swap x, y by symmetry.

Proof. Let $\mathcal{F}$ be the set of functions that satisfy all the results above. The proof consists of the following steps.

(0) $\mathcal{F}$ is nonempty for $0 \in \mathcal{F}$.

(C1) $\mathcal{F}$ is closed under linear combinations.

(C2) $\mathcal{F}$ is closed under a monotonic limit (which is integrable).

(P3) $E \subset \mathbb{R}^{d_1+d_2}$ is a G_δ-set and $m(E) < \infty \implies \chi_E \in \mathcal{F}$.

(P4) $E \subset \mathbb{R}^{d_1+d_2}$ has $m(E) = 0 \implies \chi_E \in \mathcal{F}$.

(G5) $E \subset \mathbb{R}^{d_1+d_2}$ is measurable $\implies \chi_E \in \mathcal{F}$.

(G6) $L^1(\mathbb{R}^{d_1+d_2}) \subset \mathcal{F}$.

Note 9.1. The process is to show two kinds of **c**losedness of $\mathcal{F}$ and that two **p**articular types of functions belongs to $\mathcal{F}$, and finally to use these functions and closedness to **g**eneralize to any integrable function. I call this process the **CPG** (closed, particular, generalize) process.

Fubini's theorem (extended form, Tonelli) Let $f(x,y)$ be nonnegative and measurable on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then

(i) For a.e. y, f^y is measurable on $\mathbb{E}^{d_1}$.

(ii) For a.e. y, $\int f^y(x)\mathrm{d}x$ is integrable on $\mathbb{R}^{d_2}$.

(iii) In the sense that both sides may be ∞,
$$\int_{\mathbb{R}^{d_2}}\left(\int_{\mathbb{R}^{d_1}} f(x,y)\mathrm{d}x\right)\mathrm{d}y=\int_{\mathbb{R}^{d_1+d_2}} f(x,y)\mathrm{d}x\mathrm{d}y$$

(iv) Everything above holds if we swap x, y by symmetry.

Proof. Consider the truncation
$$f_k=\begin{cases} f(x,y) & , \quad |(x,y)|<k,\ f(x,y)<k \\ k & , \quad |(x,y)|<k,\ f(x,y)\geqslant k \\ 0 & , \quad \text{elsewhere} \end{cases}$$

Then $f_k \nearrow f$ and each f_k is integrable. By Fubini's theorem and its monotonicity we arrive at the following.

(1) f_k^y is integrable for a.e. y

(2) $f_k^y \nearrow f^y$

(3) $\int f_k^y(x)\mathrm{d}x$ is integrable for a.e. y

(4) $\int f_k^y(x)\mathrm{d}x \nearrow \int f^y(x)\mathrm{d}x$

(5) $\int f \leftarrow \int f_k = \int \left(\int f_k^y(x)\mathrm{d}x\right)\mathrm{d}y \to \int \left(\int f \mathrm{d}x\right)\mathrm{d}y$

Note that by consecutively applying Fubini's theorem we obtain (2) $\implies$ (4) $\implies$ (5), and the rest comes from the monotonicity.

Note 9.2. I still need a quick way to memorize this.

- Integrable $\implies$ a.e. integrable, a.e. integrable and equal in $\mathbb{R}$.
- Nonnegative measurable $\implies$ a.e. measurable, a.e. measurable and equal in $\overline{\mathbb{R}}$.

9.2 Applications

Slice of a measurable set Let $E \subset \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ be measurable, then for a.e. y the slice

$$E^y = \{x \in \mathbb{R}^{d_1} : (x, y) \in E\}$$

is measurable in $\mathbb{R}^{d_1}$, and

$$m(E) = \int \chi_E = \int \left(\int \chi_E^y(x, y)\mathrm{d}x\right)\mathrm{d}y = \int m(E^y)\mathrm{d}y$$

To put it simply, almost every slice of a measurable set is measurable, and the measure is the integral of the measures of slices.

Measurable product set $E = E_1 \times E_2$ measurable in $\mathbb{R}^{d_1+d_2}$, $m_*(E_2) > 0$ in $\mathbb{R}^{d_2}$ $\implies$ E_1 measurable.

Proof. Notice that for a product set we have $\chi_{E_1 \times E_2}^y(x) = \chi_{E_1}(x)\chi_{E_2}(y)$. And it suffices to prove there exists $y \in E_2$ such that $\chi_{E_1 \times E_2}^y(x) = \chi_{E_1}(x)$ is measurable.

General product set For any product set $E_1 \times E_2$,

(i) $m_*(E_1 \times E_2) \leqslant m_*(E_1)m_*(E_2)$

(ii) If moreover, E_1, E_2 are measurable, then $E_1 \times E_2$ is measurable and

$$m(E_1 \times E_2) = m(E_1)m(E_2)$$

Transdimensional orthogonal extension Let $f : \mathbb{R}^{d_1} \to \overline{\mathbb{R}}$ be measurable and the extension $\widetilde{f}$ of f to $\mathbb{R}^{d_1+d_2}$ is measurable.

Volume under a function Let $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ be nonnegative. The part "under" f is

$$\mathcal{A} = \{(x, y) \in \mathbb{R}^d \times \mathbb{R} : 0 \leqslant y \leqslant f(x)\}$$

Then

(i) f is measurable on $\mathbb{R}^d \iff \mathcal{A}$ is measurable on $\mathbb{R}^{d+1}$.

(ii) If (i) holds, then $m(\mathcal{A}) = \int f$.

10 Functions Constructed Using the Integral

10.1 Integral with Variable Upper Limit

If f is integrable on $\mathbb{R}^d$ then $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$ is uniformly continuous. This follows from the absolutely continuity of the integral.

10.2 Convolution

Let $f, g : \mathbb{R}^d \to \overline{\mathbb{R}}$. Let $F(x, y) = f(x - y)g(y)$

(i) f, g are measurable $\implies F$ is measurable on $\mathbb{R}^{2d}$.

(ii) f, g are integrable $\implies F$ is integrable on $\mathbb{R}^{2d}$.

(iii) f, g are integrable $\implies F_x$ is integrable with respect to $y \in \mathbb{R}^d$ for a.e. x. And the convolution

$$(f * g)(x) = \int_{\mathbb{R}^d} F_x \mathrm{d}y$$

is defined for a.e. x.

(iv) f, g are integrable $\implies (f * g)(x)$ is integrable.

(iv′) f, g are integrable and nonnegative $\implies \|f * g\| \leqslant \|f\| \, \|g\|$.

(v) f is integrable and g is bounded $\implies f * g$ is uniformly continuous.

(vi) f, g are integrable and g is bounded $\implies \lim_{|x|\to\infty}(f * g)(x) = 0$.

11 Differentiation

11.1 The Maximal Function and the Lebesgue Differentiation Theorem

If without further specification, B denotes a ball in this section.

If $f \in L^1(\mathbb{R}^d)$, its **(Hardy-Littlewood) maximal function** f^* is defined as

$$f^*(x) = \sup_{x \in B} \int_B |f(y)| \, dy$$

The maximal function has the following properties.

Theorem 11.1. $f \in L^1(\mathbb{R}^d) \implies$

(i) f^* *is measurable.*

(ii) $f^* < \infty$ *a.e.*

(iii) f^* *satisfies the* ***weak-type*** *inequality*

$$m\{f^* > \alpha\} \leqslant \frac{A}{\alpha} \|f\|$$

where A *can be chosen as* $A = 3^d$.

Proof.

(i) $\{f^* > \alpha\}$ is open.

(ii) This follows from (iii).

(iii) One needs the Vitali covering argument.

The properties of the maximal functions lead to the Lebesgue differentiation theorem.

Theorem 11.2. *If* $f \in L^1(\mathbb{R}^d)$*, then*

$$\lim_{\substack{m(B) \to 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y) \mathrm{d}y = f(x) \text{ a.e.}$$

If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$*, then*

$$\lim_{\substack{m(B) \to 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y) \mathrm{d}y = f(x) \text{ a.e.}$$

The following concept is a useful substitute for pointwise continuity. If $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, then we define the **Lebesgue set** of f to be

$$L(f) = \left\{ x \in \mathbb{R}^d : \lim_{\substack{m(B)\to 0 \\ x\in B}} \frac{1}{m(B)} \int_B |f(y) - f(x)| \, \mathrm{d}y = 0 \right\}$$

We have the following properties of the Lebesgue set.

Theorem 11.3. *Let $f \in L^1_{\text{loc}}(\mathbb{R}^d)$.*

(i) If x is a continuity of f, then x is in the Lebesgue set.

(ii) If x is in the Lebesgue set, then the differentiation formula holds.

$$\lim_{\substack{m(B)\to 0 \\ x\in B}} \frac{1}{m(B)} \int_B f(y)\mathrm{d}y = f(x)$$

(iii) Almost every point is in the Lebesgue set.

11.2 Good Kernels and Approximation to the Identity

In this part we consider good kernels that "flattens" a function. A family of **good kernels** is a family $\{K_\delta\}_{\delta>0}$ of integrable functions such that

(i) $\int K_\delta = 1$ (normalization condition)

(ii) $\int |K_\delta| \leqslant A$ (uniformly bounded condition)

(iii) $\forall \eta > 0,\ \lim_{\delta\to 0} \int_{|x|\geqslant\eta} |K_\delta| \, \mathrm{d}x = 0$ (converging condition)

Even better kernels, called **approximations to the identity**, are defined to be a family $\{K_\delta\}_{\delta>0}$ of integrable functions that satisfies

(i) $\int K_\delta = 1$ (normalization condition)

(ii′) $|K_\delta| \leqslant A\delta^{-d}$ (peak bound condition)

(iii′) $|K_\delta| \leqslant A\delta / |x|^{d+1}$ (decreasing rate condition)

These kernels are approximations to the identity element in the convolution. Note that the convolution

$$(f * K_\delta)(x) = \int f(x-y)K_\delta(y)\mathrm{d}y$$

means to assigns a weight function $K_\delta(y)$ with its mass centered at 0. The smaller the δ, the most concentrated the mass at 0. By translating $f(x)$ by y, we are actually considering the value of f near x. The precise formulation is as follows.

Theorem 11.4. *If $\{K_\delta\}_{\delta>0}$ is family of good kernels and $f \in L^1(\mathbb{R}^d)$, then $(f * K_\delta)$ is integrable and $(f * K_\delta) \to f$ in L^1. Moreover, if $\{K_\delta\}_{\delta>0}$ is an approximation to the identity, then $(f * K_\delta) \to f$ a.e.*

11.3 Covering Arguments

A collection $\mathcal{B} = \{B\}$ of balls is called a **Vitali covering** of a set E if for any $x \in E$ and any $\eta > 0$ there exists $B \in \mathcal{B}$ such that $x \in B$, i.e., each point x of E is covered by a ball B in $\mathcal{B}$ of arbitrarily small measure. The covering arguments are as follows.

Theorem 11.5. *Let $\mathcal{A} = \{A_1, \cdots, A_N\}$ be a finite collection of open balls. Let $\mathcal{B} = \{B\}$ be a Vitali covering of a set E with $m(E) < \infty$. Then*

(i) For $\mathcal{A}$, there exists a disjoint sub-collection $A_{k_1}, \cdots, A_{k_l}$ such that

$$m\left(\bigcup_{i=1}^{N} A_i\right) \leqslant 3^d \sum_{j=1}^{l} m(A_{k_j})$$

(ii) For $\mathcal{B}$, we can find for any $\delta > 0$ a disjoint, finite sub-collection $B_1, \cdots, B_m$ such that

$$\sum_{i=1}^{l} m(B_i) \geqslant m(E) - \delta, \quad m(E - \bigcup_{i=1}^{l} B_i) < 2\delta$$

(iii) For $\mathcal{B}$, if $m(E) > 0$, then we can find for any $\eta > 0$ a disjoint, countable sub-collection $\{B_j\}_{j=1}^{\infty}$ such that

$$m_*\left(E \backslash \bigcup_{j=1}^{\infty} B_j\right) = 0, \quad \sum_{j=1}^{\infty} |B_j| \leqslant (1+\eta) m_*(E)$$

11.4 Rising Sun Lemma

The rising sun lemma deals with continuous functions.

Theorem 11.6. *Let $G : \mathbb{R} \to \mathbb{R}$ be a continuous function. Let*

$$E = \{x \in \mathbb{R} : \exists h_x > 0,\ G(x + h_x) > G(x)\}$$

Then

(i) E is open.

(ii) $E = \bigcup_k (a_k, b_k)$ *if nonempty, where* (a_k, b_k) *are disjoint.*

(iii) *If* $-\infty < a_k < b_k < \infty$*, then* $G(a_k) = G(b_k)$*.*

In particular, if

$$\Omega = \{x \in (a,b) : \exists h_x : 0 < h_x < b - x,\ G(x + h_x) > G(x)\}$$

then

(i) Ω *is open.*

(ii) $\Omega = \bigcup_k (c_k, d_k)$ *if nonempty, where* (c_k, d_k) *are disjoint.*

(iii) *If* $c_k > a$*, then* $G(c_k) = G(d_k)$*; if* $c_k = a$*, then* $G(c_k) \leqslant G(d_k)$*.*

Note 11.1. How to tell if a point is in the set E or Ω? Draw a horizontal line across the point under consideration. And see if it gets the sunshine.

11.5 Operators on Functions and Weak-type Inequalities

Suppose we are given some operator $\widetilde{(\cdot)}$ such that it takes an integrable function f to a measurable function $\widetilde{f}$. Then a **weak-type inequality** is an inequality of the form

$$m\{\widetilde{f} > \alpha\} \leqslant \frac{A\,\|f\|}{\alpha}$$

It is called a **weak-type** inequality because it is weaker than the Chebyshev inequality. The following are some examples.

Absolute value For an integrable function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$,

$$m\{|f| > \alpha\} \leqslant \frac{\|f\|}{\alpha}$$

Maximal operator associated to balls For an integrable function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$,

$$m\{f^* > \alpha\} \leqslant \frac{3^d\,\|f\|}{\alpha}$$

Maximal operator associated to rectangles For an integrable function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$, $f^*_{\mathcal{R}}$ **does not** satisfy

$$m\{f^*_{\mathcal{R}} > \alpha\} \leqslant \frac{A\,\|f\|}{\alpha}$$

Now we talk about the two maximal operators. the (ball) maximal operator $f \mapsto f^*$ is defined by

$$f^*(x) = \sup_{x \in B} \frac{1}{m(B)} \int_B |f(y)| \, \mathrm{d}y$$

where B ranges over all balls containing x. And the (rectangle) maximal operator $f \mapsto f^*_{\mathcal{R}}$ (where $\mathcal{R}$ is the set of all rectangles containing 0) is defined by

$$\begin{aligned} f^*_{\mathcal{R}}(x) &= \sup_{R \in \mathcal{R}} \frac{1}{m(R)} \int_R |f(x-y)| \, \mathrm{d}y \\ &= \sup_{R \in \mathcal{R}} \frac{1}{m(R)} \int_{R+x} |f(y)| \, \mathrm{d}y \\ &= \sup_{x \in R} \frac{1}{m(R)} \int_R |f(y)| \, \mathrm{d}y \end{aligned}$$

Now we see that these two maximal operators differ in the families that they are associated to. I guess the reason why $f \mapsto f^*_{\mathcal{R}}$ does not satisfy the weak-type inequality is that the family $\mathcal{R}^*$ does not have a bounded eccentricity.

11.6 Variation of a Function

Let $\mathcal{P}$ be the set of all partitions of the interval $[a,b]$ and let $P = \{a = x_0 < x_1 < \cdots < x_n = b\}$ denote some specific partition. The **total variation** of f on $[a,b]$ is

$$T_f(a,b) = \sup_{P \subset \mathcal{P}} \sum_{j=1}^{n} |f(x_i) - f(x_{i-1})|$$

The **positive** and **negative variations** are defined as follows.

$$P_f(a,b) = \sup_{P \subset \mathcal{P}} \sum_{f(x_i) \geqslant f(x_{i-1})} (f(x_i) - f(x_{i-1}))$$

$$N_f(a,b) = \sup_{P \subset \mathcal{P}} \sum_{f(x_i) < f(x_{i-1})} (f(x_{i-1}) - f(x_i))$$

Without assuming that any of them is finite, we have the following properties. And for the simplicity of statement, we let $V(f,P)$ denote the variation of f with respect to the partition P and let $V^+(f,P), V^-(f,P)$ denote respectively the positive and negative variations with respect to P.

Theorem 11.7. *For a function $f : [a,b] \to \mathbb{R}$, the following are true.*

(i) $T_f(a,b), P_f(a,b), N_f(a,b) \geqslant 0$

(ii) $V(f,P) = V^+(f,P) + V^-(f,P)$

(iii) $V(f,P), V^+(f,P), V^-(f,P)$ *only go larger under refinement.*

(iv) $T_f(a,b) = P_f(a,b) + N_f(a,b)$

(v) $T_f(a,b) = T_f(a,c) = T_f(c,b)$ *for* $a < c < b$.

Proof.

(i) Obvious.

(ii) Obvious.

(iii) See what happens if the refinement has only one point.

(iv) We start from the very beginning.

$$V(f,P) = V^+(f,P) + V^-(f,P)$$

It follows from this simple fact that

$$\sup_{P\in\mathcal{P}} V(f,P) \geqslant V(f,P) = V^+(f,P) + V^-(f,P)$$

$$\sup_{P\in\mathcal{P}} V(f,P) = \sup_{P\in\mathcal{P}}(V^+(f,P) + V^-(f,P)) \leqslant \sup_{P\in\mathcal{P}} V^+(f,P) + \sup_{P\in\mathcal{P}} V^-(f,P)$$

Hence

$$P_f + N_f \geqslant T_f \geqslant V^+(f,P) + V^-(f,P)$$

Now suppose f has bounded variation, then $T_f < \infty \implies P_f, N_f \leqslant \infty$, hence we can find partitions to approximate P_f, N_f by V^+, V^-. By a refinement we have a common partition that approximates P_f, N_f. If $T_f = \infty$, then $P_f + N_f \geqslant T_f \implies$ at least one of P_f, N_f is ∞, hence $T_f = P_f + N_f = \infty$.

(v) Choose approximations.

The applications of variations are in the next section.

11.7 Differentiable Functions

The differentiation problem is simplified if we use the following decompositions.

Theorem 11.8. *The following are true.*

(i) *If F is of bounded variation on $[a,b]$, then $F(x) = F(a) + P_F(a,x) - N_F(a,x)$, where both $P_F(a,x), N_F(a,x)$ are increasing.*

(ii) *If F is of bounded variation and continuous, then $F(x) = F_1(x) - F_2(x)$ are increasing and continuous functions.*

(iii) *If F is increasing on $[a,b]$, then $F(x) = (F(x) - J(x)) + J(x)$ where $F(x) - J(x)$ is increasing and continuous, and $J(x)$ is the jump function.*

The following functions are differentiable.

Theorem 11.9. *(i)* *If F is increasing and continuous, then F' exists a.e., F' is measurable, nonnegative,*

$$\int_a^b F'(x)\mathrm{d}x \leqslant F(b) - F(a)$$

Moreover, F is bounded on $\mathbb{R} \implies F'$ is integrable on $\mathbb{R}$.

(ii) *If F is absolutely continuous on $[a,b]$, then F' exists a.e., F' is integrable on $[a,b]$,*

$$\int_a^x F'(y)\mathrm{d}y = F(x) - F(a)$$

(iii) *If J is a jump function, then J' exists a.e. and vanishes a.e..*