

Real Analysis

Renko

1145.1.4

Contents

Preface	3
1 Measure Theory	4
1.1 Elementary Topology and Cubes in \mathbb{R}^d	4
1.2 Lebesgue Outer Measure	5
1.3 Measurable Sets	6
1.4 Measurable Functions	9
2 Integration Theory	13
2.1 The Lebesgue Integral	13
2.2 Convergence Theorems	14
2.3 Fubini's Theorem on Euclidean Product Spaces	16
2.4 Absolute Continuity of the Integral and Uniform Integrability	17
3 Differentiation and Integration	20
3.1 Hardy–Littlewood Maximal Function and Differentiation	20
3.2 Bounded Variation and Absolute Continuity	22
3.3 Convex Functions and Jensen's Inequality	24
3.4 Convolution and Approximation	25
4 L^p Spaces and Weak Convergence	29
4.1 Normed, Banach, and Hilbert Spaces	29
4.2 Basic Inequalities	29
4.3 Completeness, Separability, and Convergence in L^p	31
4.4 Duality and Weak Convergence	32
5 Abstract Measure and Integration Spaces	35
5.1 Measure spaces and elementary consequences	35
5.2 Outer measures and Caratheodory's theorem	37
5.3 Integration on an abstract measure space	40
5.4 Product measures and Fubini–Tonelli	42
5.5 Absolute continuity and Radon–Nikodym theorem	44
5.6 A short conceptual summary	47
A Haar Measure and Convolution on Locally Compact Groups	48

Preface

This version is written from the handwritten notes after an OCR pass and a subsequent mathematical clean-up. The OCR was useful mainly for indexing the pages and locating repetitions; the handwritten mathematical formulas were then checked manually against the page images. Repeated results have been merged, but the proofs are kept in full rather than as sketches.

The notation follows the handwritten notes. Lebesgue measure on \mathbb{R}^d is denoted by m . When no measure space is explicitly stated, all assertions are made on Lebesgue measurable subsets of \mathbb{R}^d . Statements which appeared twice in the notes, such as bounded convergence, Fatou–monotone–dominated convergence, density/approximation by step functions, and several convergence-in-measure criteria, are stated only once and then reused.

Chapter 1

Measure Theory

1.1 Elementary Topology and Cubes in \mathbb{R}^d

Definition 1.1. A set $E \subset \mathbb{R}^d$ is *bounded* if it is contained in a ball of finite radius. A set $K \subset \mathbb{R}^d$ is *compact* if every open cover of K admits a finite subcover.

Theorem 1.2 (Heine–Borel). *For $K \subset \mathbb{R}^d$, the following are equivalent:*

- (i) K is compact;
- (ii) K is closed and bounded.

Proof. Assume first that K is compact. The open balls $B(0, n)$, $n \in \mathbb{N}$, cover K , so finitely many of them cover K . Hence $K \subset B(0, N)$ for some N , and K is bounded. To see that K is closed, take $x \notin K$. For each $y \in K$, choose disjoint balls $B_y \ni y$ and $U_y \ni x$. The balls B_y cover K , so finitely many B_{y_1}, \dots, B_{y_N} cover K . Then $U = \bigcap_{j=1}^N U_{y_j}$ is an open neighbourhood of x disjoint from K . Thus K^c is open and K is closed.

Conversely, assume K is closed and bounded. Then $K \subset [-R, R]^d$ for some R . It suffices to show that the cube $[-R, R]^d$ is compact, since a closed subset of a compact set is compact. Let $\{O_\alpha\}$ be an open cover of $[-R, R]^d$. If no finite subcover exists, divide the cube into 2^d congruent closed subcubes; at least one subcube has no finite subcover. Repeat this construction to get nested closed cubes Q_n , each without finite subcover, with diameters tending to 0. Their centers form a Cauchy sequence, hence converge to some point $x \in \bigcap_n Q_n$. Since the O_α cover the original cube, $x \in O_{\alpha_0}$ for some α_0 . As O_{α_0} is open and $\text{diam } Q_n \rightarrow 0$, some $Q_n \subset O_{\alpha_0}$, contradicting the choice of Q_n . Hence a finite subcover exists. \square

Definition 1.3. A *rectangle* in \mathbb{R}^d is a product $R = I_1 \times \dots \times I_d$ of intervals. A *cube* is a rectangle whose side lengths are equal. Two rectangles or cubes are *almost disjoint* if their interiors are disjoint. If Q is a cube, $|Q|$ denotes its Euclidean volume.

Lemma 1.4. *Every open subset $O \subset \mathbb{R}$ is a countable disjoint union of open intervals.*

Proof. For $x \in O$, let I_x be the union of all open intervals contained in O and containing x . Then I_x is an open interval, and it is maximal with respect to inclusion among open intervals contained in O . If $I_x \cap I_y \neq \emptyset$, then $I_x \cup I_y$ is again an interval contained in O , so maximality gives $I_x = I_y$. Hence the distinct intervals in this construction are disjoint and their union is O . Each non-empty open interval contains a rational number, and disjoint intervals contain distinct rationals. Since \mathbb{Q} is countable, the collection of intervals is countable. \square

Lemma 1.5. *Every open subset $O \subset \mathbb{R}^d$, $d \geq 1$, can be written as a countable union of almost disjoint closed cubes.*

Proof. For each integer $n \geq 0$, divide \mathbb{R}^d into the dyadic half-open cubes of side length 2^{-n} . Let \mathcal{Q}_n be the family of closed dyadic cubes of side length 2^{-n} contained in O . We construct a disjoint family by choosing first all cubes in \mathcal{Q}_0 , then all cubes in \mathcal{Q}_1 not contained in any previously chosen cube, and so on. At each level there are countably many cubes, hence the final family is countable. The interiors of the chosen cubes are disjoint by construction. If $x \in O$, choose $r > 0$ such that $B(x, r) \subset O$. For n large enough, there is a dyadic cube of side 2^{-n} containing x and contained in $B(x, r)$, hence in O . Either this cube is chosen, or it lies inside a previously chosen cube. Therefore x belongs to the union of the chosen cubes. \square

Lemma 1.6 (Volume subadditivity for rectangles). *Let R and R_1, R_2, \dots be rectangles in \mathbb{R}^d . If $R \subset \bigcup_{j=1}^{\infty} R_j$, then*

$$|R| \leq \sum_{j=1}^{\infty} |R_j|.$$

Proof. First suppose R is compact and the R_j are open rectangles. By compactness, finitely many R_1, \dots, R_N cover R . Subdivide R by all coordinate hyperplanes determined by the sides of these finitely many rectangles. This gives finitely many pairwise almost disjoint subrectangles S_k . Each S_k is contained in at least one R_j ; assigning it to one such R_j , we get

$$|R| = \sum_k |S_k| \leq \sum_{j=1}^N |R_j|.$$

For arbitrary rectangles, enlarge each R_j slightly to an open rectangle R_j^ε with $|R_j^\varepsilon| \leq |R_j| + \varepsilon 2^{-j}$, and shrink R slightly to a compact rectangle R^- with $|R^-| \uparrow |R|$. Applying the compact-open case gives

$$|R^-| \leq \sum_j |R_j^\varepsilon| \leq \sum_j |R_j| + \varepsilon.$$

Letting the shrinking disappear and then letting $\varepsilon \downarrow 0$ proves the result. \square

1.2 Lebesgue Outer Measure

Definition 1.7. For an arbitrary subset $E \subset \mathbb{R}^d$, its *Lebesgue outer measure* is

$$m^*(E) = \inf \left\{ \sum_{j=1}^{\infty} |Q_j| : E \subset \bigcup_{j=1}^{\infty} Q_j, Q_j \text{ closed cubes} \right\}.$$

When no confusion is possible, the notes write $m_*(E)$ or $m(E)$ for this outer measure before measurability is introduced.

Proposition 1.8. *Lebesgue outer measure satisfies:*

- (i) $m^*(\emptyset) = 0$;
- (ii) if $E \subset F$, then $m^*(E) \leq m^*(F)$;
- (iii) if $E \subset \bigcup_{j=1}^{\infty} E_j$, then $m^*(E) \leq \sum_{j=1}^{\infty} m^*(E_j)$.

Proof. For (i), the empty family covers \emptyset , so the infimum is 0. For (ii), every cube cover of F is also a cube cover of E , hence the infimum for E is no larger. For (iii), fix $\varepsilon > 0$. For each j , choose closed cubes $Q_{j,k}$ such that $E_j \subset \bigcup_k Q_{j,k}$ and

$$\sum_k |Q_{j,k}| \leq m^*(E_j) + \varepsilon 2^{-j}.$$

Then $E \subset \bigcup_{j,k} Q_{j,k}$, so

$$m^*(E) \leq \sum_{j,k} |Q_{j,k}| \leq \sum_j m^*(E_j) + \varepsilon.$$

Let $\varepsilon \downarrow 0$. □

Proposition 1.9. *For every set $E \subset \mathbb{R}^d$,*

$$m^*(E) = \inf\{m^*(O) : O \text{ open and } E \subset O\}.$$

Proof. The inequality $m^*(E) \leq \inf_{O \supset E} m^*(O)$ follows from monotonicity. Conversely, fix $\varepsilon > 0$. Choose closed cubes Q_j covering E with $\sum_j |Q_j| < m^*(E) + \varepsilon/2$. Enlarge each Q_j to an open cube Q'_j such that $|Q'_j| \leq |Q_j| + \varepsilon 2^{-j-1}$. Then $O = \bigcup_j Q'_j$ is open and contains E , and by countable subadditivity

$$m^*(O) \leq \sum_j |Q'_j| < m^*(E) + \varepsilon.$$

Taking the infimum over open O and letting $\varepsilon \downarrow 0$ gives the reverse inequality. □

Proposition 1.10. *If Q is a closed cube, then $m^*(Q) = |Q|$.*

Proof. The inequality $m^*(Q) \leq |Q|$ is immediate from the one-cube cover. For the reverse inequality, let $Q \subset \bigcup_j Q_j$ be any cover by closed cubes. Enlarge each Q_j to an open cube Q'_j with $|Q'_j| \leq |Q_j| + \varepsilon 2^{-j}$. Since Q is compact, finitely many Q'_j cover Q . By the volume subadditivity lemma,

$$|Q| \leq \sum_{j=1}^N |Q'_j| \leq \sum_j |Q_j| + \varepsilon.$$

Let $\varepsilon \downarrow 0$, then take the infimum over all covers. □

Proposition 1.11. *The Cantor set $C \subset [0, 1]$ is closed, has Lebesgue outer measure zero, and is uncountable.*

Proof. At the n -th stage of the usual middle-third construction, the remaining set C_n is a finite union of 2^n closed intervals, each of length 3^{-n} . Hence C_n is closed and $C = \bigcap_{n=0}^{\infty} C_n$ is closed. Since $C \subset C_n$,

$$m^*(C) \leq m^*(C_n) = 2^n 3^{-n} = (2/3)^n.$$

Letting $n \rightarrow \infty$ gives $m^*(C) = 0$.

To prove uncountability, write every point of C using a ternary expansion containing only the digits 0 and 2, choosing the expansion that does not terminate in all 2's when there are two choices. The map

$$(\varepsilon_1, \varepsilon_2, \dots) \in \{0, 1\}^{\mathbb{N}} \mapsto \sum_{n=1}^{\infty} \frac{2\varepsilon_n}{3^n}$$

embeds the set of binary sequences into C . Since $\{0, 1\}^{\mathbb{N}}$ is uncountable by Cantor's diagonal argument, C is uncountable. □

1.3 Measurable Sets

Definition 1.12. A set $E \subset \mathbb{R}^d$ is *Lebesgue measurable* if for every $\varepsilon > 0$, there exists an open set $O \supset E$ such that

$$m^*(O \setminus E) < \varepsilon.$$

For a measurable set E , its Lebesgue measure is $m(E) = m^*(E)$.

Proposition 1.13. *Every open subset of \mathbb{R}^d is measurable.*

Proof. If E is open, take $O = E$. Then $O \setminus E = \emptyset$, hence $m^*(O \setminus E) = 0$. \square

Lemma 1.14. *Let $F \subset \mathbb{R}^d$ be closed and $K \subset \mathbb{R}^d$ compact. If $F \cap K = \emptyset$, then $\text{dist}(F, K) > 0$.*

Proof. For each $x \in K$, since F is closed and $x \notin F$, there exists $r_x > 0$ such that $B(x, 2r_x) \cap F = \emptyset$. The balls $B(x, r_x)$ cover K , so compactness gives a finite subcover $B(x_j, r_{x_j})$, $1 \leq j \leq N$. Let $r = \min_j r_{x_j} > 0$. If $y \in K$, choose j with $y \in B(x_j, r_{x_j})$. Then $B(y, r) \subset B(x_j, 2r_{x_j})$, hence $B(y, r) \cap F = \emptyset$. Thus $\text{dist}(F, K) \geq r$. \square

Proposition 1.15. *Every closed subset of \mathbb{R}^d is measurable.*

Proof. Let F be closed. First suppose F is compact. Given $\varepsilon > 0$, by the open-set approximation of outer measure choose an open set $O \supset F$ with $m^*(O) < m^*(F) + \varepsilon$. Since F and O^c are disjoint closed sets and F is compact, the preceding lemma gives a positive distance between them. Hence one can cover F by finitely many closed cubes contained in O . Using finite additivity for almost disjoint cube decompositions and the definition of outer measure, this implies

$$m^*(O \setminus F) \leq m^*(O) - m^*(F) + \varepsilon < 2\varepsilon$$

(after replacing O by a finite union of small cubes inside it). Therefore compact sets are measurable.

For a general closed set F , let $K_n = F \cap \overline{B(0, n)}$. Each K_n is compact and hence measurable, and $F = \bigcup_n K_n$. The next proposition shows that countable unions of measurable sets are measurable, so F is measurable. \square

Proposition 1.16. *The collection of Lebesgue measurable subsets of \mathbb{R}^d is a σ -algebra. In particular it is closed under complements and countable unions.*

Proof. We first prove closure under complements. Let E be measurable and fix $\varepsilon > 0$. Choose an open set $O \supset E$ with $m^*(O \setminus E) < \varepsilon$. Then $F = O^c$ is closed and $F \subset E^c$. Moreover $E^c \setminus F = O \setminus E$, so $m^*(E^c \setminus F) < \varepsilon$. Since closed sets are measurable and E^c differs from the measurable set F by an arbitrarily small outer-measure set, the open-set criterion applied to F gives measurability of E^c .

For countable unions, let $E = \bigcup_{j=1}^{\infty} E_j$, where each E_j is measurable. Given $\varepsilon > 0$, choose open $O_j \supset E_j$ with $m^*(O_j \setminus E_j) < \varepsilon 2^{-j}$. Then $O = \bigcup_j O_j$ is open and contains E . Since

$$O \setminus E \subset \bigcup_j (O_j \setminus E_j),$$

countable subadditivity gives $m^*(O \setminus E) < \varepsilon$. Thus E is measurable. Closure under countable intersections follows from complements and countable unions. \square

Theorem 1.17 (Countable additivity). *If E_1, E_2, \dots are pairwise disjoint measurable sets, then*

$$m\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} m(E_j).$$

Proof. Countable subadditivity gives $m(\bigcup_j E_j) \leq \sum_j m(E_j)$. For the reverse inequality, first fix N . Since the sets are measurable and disjoint, finite additivity follows from the definition of measurability: for two disjoint measurable sets A, B , applying Caratheodory's splitting inequality

to $A \cup B$ gives $m(A \cup B) \geq m(A) + m(B)$, while subadditivity gives the opposite inequality. Induction yields

$$m\left(\bigcup_{j=1}^N E_j\right) = \sum_{j=1}^N m(E_j).$$

By monotonicity,

$$m\left(\bigcup_{j=1}^{\infty} E_j\right) \geq m\left(\bigcup_{j=1}^N E_j\right) = \sum_{j=1}^N m(E_j).$$

Let $N \rightarrow \infty$. □

Theorem 1.18 (Continuity of measure). *Let E_n be measurable sets.*

(i) *If $E_n \uparrow E$, then $m(E) = \lim_{n \rightarrow \infty} m(E_n)$.*

(ii) *If $E_n \downarrow E$ and $m(E_1) < \infty$, then $m(E) = \lim_{n \rightarrow \infty} m(E_n)$.*

Proof. For (i), set $F_1 = E_1$ and $F_n = E_n \setminus E_{n-1}$ for $n \geq 2$. Then the F_n are pairwise disjoint and $E = \bigcup_n F_n$. Also $E_N = \bigcup_{n=1}^N F_n$. Countable additivity gives

$$m(E) = \sum_{n=1}^{\infty} m(F_n) = \lim_{N \rightarrow \infty} \sum_{n=1}^N m(F_n) = \lim_{N \rightarrow \infty} m(E_N).$$

For (ii), apply (i) to $F_n = E_1 \setminus E_n$, which increases to $E_1 \setminus E$. Since $m(E_1) < \infty$,

$$m(E_1) - m(E) = m(E_1 \setminus E) = \lim_n m(E_1 \setminus E_n) = m(E_1) - \lim_n m(E_n).$$

Rearranging proves the claim. □

Theorem 1.19 (Regularity of Lebesgue measure). *Let $E \subset \mathbb{R}^d$ be measurable.*

(i) *For every $\varepsilon > 0$, there exists an open set $O \supset E$ such that $m(O \setminus E) < \varepsilon$.*

(ii) *If $m(E) < \infty$, for every $\varepsilon > 0$ there exists a compact set $K \subset E$ such that $m(E \setminus K) < \varepsilon$.*

(iii) *If $m(E) < \infty$, for every $\varepsilon > 0$ there exists a finite union F of almost disjoint closed cubes such that $m(E \Delta F) < \varepsilon$.*

Proof. Part (i) is the definition of measurability. For (ii), apply (i) to E^c to get an open set $O \supset E^c$ with $m(O \setminus E^c) < \varepsilon/2$. Then $F = O^c \subset E$ is closed and $m(E \setminus F) < \varepsilon/2$. If E has finite measure, choose R so large that $m(E \setminus B(0, R)) < \varepsilon/2$. Then $K = F \cap \overline{B(0, R)}$ is compact and

$$m(E \setminus K) \leq m(E \setminus F) + m(E \setminus B(0, R)) < \varepsilon.$$

For (iii), choose an open $O \supset E$ with $m(O \setminus E) < \varepsilon/2$. Decompose O into countably many almost disjoint closed cubes Q_j . Since $m(E) < \infty$, also $m(O) < \infty$ after choosing O with finite measure. Hence the tails of $\sum_j |Q_j| = m(O)$ tend to zero. Pick N such that $m(\bigcup_{j>N} Q_j) < \varepsilon/2$, and set $F = \bigcup_{j=1}^N Q_j$. Then $F \subset O$ and

$$E \Delta F \subset (O \setminus E) \cup \bigcup_{j>N} Q_j,$$

so $m(E \Delta F) < \varepsilon$. □

Definition 1.20. An algebra \mathcal{A} of subsets of a set X is a collection closed under finite unions and complements. A σ -algebra \mathcal{M} is closed under countable unions and complements. The Borel σ -algebra $\mathcal{B}(\mathbb{R}^d)$ is the smallest σ -algebra containing all open subsets of \mathbb{R}^d .

Corollary 1.21. *Every Borel set is Lebesgue measurable.*

Proof. The measurable sets form a σ -algebra and contain all open sets. Therefore they contain the smallest σ -algebra generated by the open sets, namely the Borel σ -algebra. \square

Lemma 1.22 (Borel–Cantelli). *If $\{E_k\}$ is a sequence of measurable sets and $\sum_{k=1}^{\infty} m(E_k) < \infty$, then almost every point belongs to at most finitely many of the E_k . Equivalently,*

$$m\left(\limsup_{k \rightarrow \infty} E_k\right) = 0.$$

Proof. Set $B_n = \bigcup_{k \geq n} E_k$. Then $B_n \downarrow \limsup_k E_k$. By subadditivity,

$$m(B_n) \leq \sum_{k \geq n} m(E_k) \rightarrow 0.$$

Continuity from above gives $m(\limsup_k E_k) = \lim_n m(B_n) = 0$. The set $\limsup_k E_k$ is exactly the set of points lying in infinitely many E_k . \square

Theorem 1.23 (Existence of non-measurable sets). *Every measurable set $E \subset \mathbb{R}$ with $m(E) > 0$ contains a non-measurable subset.*

Proof. It suffices to prove this for a bounded measurable set of positive measure. Choose $b > 0$ such that $m(E \cap [-b, b]) > 0$, and replace E by this intersection. Define an equivalence relation on E by $x \sim y$ if $x - y \in \mathbb{Q}$. Using the axiom of choice, choose a set $V \subset E$ containing exactly one representative from each equivalence class. For distinct rationals q, q' , the translates $V + q$ and $V + q'$ are disjoint, because otherwise $v + q = v' + q'$ would imply $v - v' \in \mathbb{Q}$, hence $v = v'$ and $q = q'$.

Let $Q_0 = \mathbb{Q} \cap [-2b, 2b]$. Since every $x \in E$ is rationally equivalent to some $v \in V$, we have

$$E \subset \bigcup_{q \in Q_0} (V + q) \subset [-3b, 3b].$$

If V were measurable, all translates $V + q$ would be measurable and have the same measure. If $m(V) = 0$, then the countable union above would have measure zero, contradicting $m(E) > 0$. If $m(V) > 0$, then the disjoint countable union $\bigcup_{q \in Q_0} (V + q)$ would have infinite measure, contradicting that it is contained in $[-3b, 3b]$. Hence V is not measurable. \square

1.4 Measurable Functions

Definition 1.24. For a set E , its characteristic function is $\mathbf{1}_E(x) = 1$ for $x \in E$ and 0 otherwise. A real-valued function f on a measurable set E is *measurable* if $\{x \in E : f(x) < a\}$ is measurable for every $a \in \mathbb{R}$.

Proposition 1.25. *For a finite-valued function $f : E \rightarrow \mathbb{R}$, the following are equivalent:*

- (i) f is measurable;
- (ii) $f^{-1}(O)$ is measurable for every open set $O \subset \mathbb{R}$;
- (iii) $\{x \in E : f(x) > a\}$, $\{x \in E : f(x) \geq a\}$, and $\{x \in E : f(x) \leq a\}$ are measurable for every $a \in \mathbb{R}$.

Proof. If f is measurable, then $\{f < a\}$ is measurable. Since $\{f > a\} = E \setminus \{f \leq a\}$ and $\{f \leq a\} = \bigcap_{n=1}^{\infty} \{f < a + 1/n\}$, the sets in (iii) are measurable. Conversely, (iii) immediately implies the defining condition. If (iii) holds, then every open set $O \subset \mathbb{R}$ is a countable disjoint union of open intervals, and the inverse image of each interval is measurable by (iii), so $f^{-1}(O)$ is measurable. Finally, if inverse images of open sets are measurable, taking $O = (-\infty, a)$ gives measurability of f . \square

Proposition 1.26. *Continuous functions are measurable. If f is measurable and $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ is continuous, then $\Phi \circ f$ is measurable.*

Proof. If f is continuous on a measurable domain and $O \subset \mathbb{R}$ is open, then $f^{-1}(O)$ is relatively open in the domain and hence measurable. Thus f is measurable. If Φ is continuous, then for every open $O \subset \mathbb{R}$, $\Phi^{-1}(O)$ is open. Hence

$$(\Phi \circ f)^{-1}(O) = f^{-1}(\Phi^{-1}(O))$$

is measurable. □

Proposition 1.27. *If $\{f_n\}$ is a sequence of measurable functions, then*

$$\sup_n f_n, \quad \inf_n f_n, \quad \limsup_{n \rightarrow \infty} f_n, \quad \liminf_{n \rightarrow \infty} f_n$$

are measurable whenever they are finite-valued or extended-real-valued in the usual sense.

Proof. For each $a \in \mathbb{R}$,

$$\{\sup_n f_n > a\} = \bigcup_n \{f_n > a\}, \quad \{\inf_n f_n < a\} = \bigcup_n \{f_n < a\}.$$

These are measurable. Thus $\sup_n f_n$ and $\inf_n f_n$ are measurable. Since

$$\limsup_n f_n = \inf_N \sup_{n \geq N} f_n, \quad \liminf_n f_n = \sup_N \inf_{n \geq N} f_n,$$

the preceding part gives measurability of the limsup and liminf. □

Corollary 1.28. *If f_n are measurable and $f_n(x) \rightarrow f(x)$ pointwise, then f is measurable.*

Proof. The pointwise limit satisfies $f = \limsup_n f_n = \liminf_n f_n$, hence is measurable by the preceding proposition. □

Proposition 1.29. *If f, g are finite-valued measurable functions and $c \in \mathbb{R}$, then cf , $f + g$, fg , $|f|$, f^k for $k \in \mathbb{N}$, $\max(f, g)$, and $\min(f, g)$ are measurable.*

Proof. The functions cf , $|f|$, and f^k are compositions of f with continuous functions, hence measurable. To prove $f + g$ measurable, observe that for each $a \in \mathbb{R}$,

$$\{f + g < a\} = \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < a - q\}).$$

Indeed, if $f(x) + g(x) < a$, choose rational q with $f(x) < q < a - g(x)$. The reverse inclusion is immediate. Hence $f + g$ is measurable. Then $fg = \frac{1}{4}((f + g)^2 - (f - g)^2)$ is measurable. Finally,

$$\max(f, g) = \frac{f + g + |f - g|}{2}, \quad \min(f, g) = \frac{f + g - |f - g|}{2},$$

so both are measurable. □

Proposition 1.30. *If $f = g$ almost everywhere and f is measurable, then g is measurable.*

Proof. Let $N = \{x : f(x) \neq g(x)\}$. Then $m(N) = 0$. Every subset of a null set is measurable: if $A \subset N$, then $m^*(A) = 0$, and for any $\varepsilon > 0$, an open set $O \supset N$ with $m(O) < \varepsilon$ also contains A . Now for each a ,

$$\{g < a\} = (\{f < a\} \cap N^c) \cup (\{g < a\} \cap N),$$

where the first set is measurable and the second is a subset of a null set. Hence $\{g < a\}$ is measurable. □

Definition 1.31. A *simple function* is a finite linear combination of characteristic functions of measurable sets,

$$\varphi = \sum_{j=1}^N a_j \mathbf{1}_{E_j}.$$

If the nonzero values a_j are distinct and the sets $E_j = \{x : \varphi(x) = a_j\}$ are disjoint, this is called the *canonical form* of φ .

Theorem 1.32 (Simple function approximation). *Let $f \geq 0$ be measurable on E . Then there exists an increasing sequence of nonnegative simple functions φ_n such that $\varphi_n(x) \uparrow f(x)$ for every $x \in E$. If f is bounded, the convergence may be chosen uniform. For a general real-valued measurable function finite almost everywhere, there are simple functions ψ_n with $|\psi_n| \leq |f|$ and $\psi_n \rightarrow f$ pointwise almost everywhere.*

Proof. For $f \geq 0$, define

$$\varphi_n(x) = \begin{cases} k2^{-n}, & k2^{-n} \leq f(x) < (k+1)2^{-n}, \quad 0 \leq k \leq n2^n - 1, \\ n, & f(x) \geq n. \end{cases}$$

Each set in this definition is measurable, so φ_n is simple. Also $0 \leq \varphi_n \leq f$, and $f(x) - \varphi_n(x) \leq 2^{-n}$ whenever $f(x) < n$, while $\varphi_n(x) = n$ when $f(x) \geq n$. Hence $\varphi_n(x) \rightarrow f(x)$. Replacing φ_n by $\max(\varphi_1, \dots, \varphi_n)$ makes the sequence increasing. If $f \leq M$, choose $n > M$, and the bound $0 \leq f - \varphi_n \leq 2^{-n}$ gives uniform convergence. For general f , write $f = f^+ - f^-$ and approximate f^+ and f^- increasingly; truncating if necessary gives simple ψ_n satisfying $|\psi_n| \leq |f|$ and $\psi_n \rightarrow f$. \square

Theorem 1.33 (Step function approximation). *If f is measurable and finite almost everywhere on \mathbb{R}^d , then there exists a sequence of step functions s_n such that $s_n \rightarrow f$ almost everywhere. If f is integrable, the step functions may be chosen so that $s_n \rightarrow f$ in L^1 .*

Proof. By the simple approximation theorem, it suffices to approximate characteristic functions of measurable sets by step functions. If A has finite measure, regularity gives a finite union F of almost disjoint rectangles such that $m(A \Delta F) < \varepsilon$. Then $\mathbf{1}_F$ is a step function and $\|\mathbf{1}_A - \mathbf{1}_F\|_{L^1} = m(A \Delta F) < \varepsilon$. For arbitrary measurable A , first intersect with a large cube, approximate inside that cube, and let the cube grow. Combining these approximations for the finitely many level sets in each simple function and then diagonalizing gives almost everywhere convergence. If $f \in L^1$, approximate f in L^1 by an integrable simple function and then approximate each characteristic set in L^1 as above. \square

Theorem 1.34 (Egorov). *Let E be measurable with $m(E) < \infty$. If $f_n \rightarrow f$ almost everywhere on E , then for every $\varepsilon > 0$, there exists a measurable set $A \subset E$ such that $m(E \setminus A) < \varepsilon$ and $f_n \rightarrow f$ uniformly on A .*

Proof. Discard a null set so that convergence holds everywhere on the remaining set, still denoted by E . For $k, n \in \mathbb{N}$, set

$$E_{n,k} = \bigcap_{j \geq n} \{x \in E : |f_j(x) - f(x)| < 1/k\}.$$

For fixed k , the sets $E_{n,k}$ increase with n , and their union is E . By continuity from below, choose n_k such that

$$m(E \setminus E_{n_k, k}) < \varepsilon 2^{-k}.$$

Let $A = \bigcap_{k=1}^{\infty} E_{n_k, k}$. Then

$$m(E \setminus A) \leq \sum_{k=1}^{\infty} m(E \setminus E_{n_k, k}) < \varepsilon.$$

If $x \in A$, then for every k and every $j \geq n_k$, $|f_j(x) - f(x)| < 1/k$. This is exactly uniform convergence on A . \square

Lemma 1.35 (Simple functions are almost continuous). *Let $E \subset \mathbb{R}^d$ be measurable with finite measure, and let φ be a simple function on E . For every $\varepsilon > 0$, there exist a closed set $F \subset E$ and a continuous function g on \mathbb{R}^d such that $m(E \setminus F) < \varepsilon$ and $g = \varphi$ on F .*

Proof. Write $\varphi = \sum_{j=1}^N a_j \mathbf{1}_{E_j}$ in canonical form, with the E_j disjoint. By inner regularity, choose closed sets $F_j \subset E_j$ with $m(E_j \setminus F_j) < \varepsilon/N$. Let $F = \bigcup_j F_j$. The sets F_j are disjoint closed subsets of \mathbb{R}^d , and since only finitely many are involved, one can choose pairwise disjoint open neighbourhoods $U_j \supset F_j$. By Urysohn's lemma for \mathbb{R}^d , choose continuous functions η_j with $\eta_j = 1$ on F_j and $\eta_j = 0$ outside U_j . Then $g = \sum_j a_j \eta_j$ is continuous and equals φ on F . Also

$$m(E \setminus F) \leq \sum_j m(E_j \setminus F_j) < \varepsilon.$$

\square

Theorem 1.36 (Lusin). *Let $E \subset \mathbb{R}^d$ be measurable with $m(E) < \infty$, and let $f : E \rightarrow \mathbb{R}$ be measurable and finite almost everywhere. For every $\varepsilon > 0$, there exists a closed set $F \subset E$ such that $m(E \setminus F) < \varepsilon$ and $f|_F$ is continuous.*

Proof. Choose simple functions $\varphi_n \rightarrow f$ almost everywhere. By Egorov's theorem, there exists a measurable set $A \subset E$ with $m(E \setminus A) < \varepsilon/3$ such that $\varphi_n \rightarrow f$ uniformly on A . For each n , apply the preceding lemma with error $\varepsilon 2^{-n-2}$ to obtain a closed set $F_n \subset E$ and a continuous function g_n such that $g_n = \varphi_n$ on F_n and $m(E \setminus F_n) < \varepsilon 2^{-n-2}$. By inner regularity choose a closed set $F_0 \subset A$ with $m(A \setminus F_0) < \varepsilon/3$. Put

$$F = F_0 \cap \bigcap_{n=1}^{\infty} F_n.$$

Then F is closed and

$$m(E \setminus F) \leq m(E \setminus A) + m(A \setminus F_0) + \sum_{n=1}^{\infty} m(E \setminus F_n) < \varepsilon.$$

On F , the functions g_n equal φ_n , and $\varphi_n \rightarrow f$ uniformly. Thus $f|_F$ is the uniform limit of continuous functions $g_n|_F$, hence is continuous. \square

Chapter 2

Integration Theory

2.1 The Lebesgue Integral

Definition 2.1. If $\varphi = \sum_{j=1}^N a_j \mathbf{1}_{E_j}$ is a nonnegative simple function in canonical form, its integral is

$$\int \varphi dx = \sum_{j=1}^N a_j m(E_j),$$

allowing the value $+\infty$. If A is measurable, define $\int_A \varphi = \int \varphi \mathbf{1}_A$.

Proposition 2.2. *The integral of a nonnegative simple function is independent of its representation. If φ, ψ are nonnegative simple functions and $a, b \geq 0$, then*

$$\int (a\varphi + b\psi) = a \int \varphi + b \int \psi.$$

Also, if $\varphi \leq \psi$, then $\int \varphi \leq \int \psi$.

Proof. Let $\varphi = \sum_i a_i \mathbf{1}_{A_i} = \sum_j b_j \mathbf{1}_{B_j}$ be two representations. Refining both by the measurable partition $A_i \cap B_j$, the value of φ is constant on each atom of the refinement, so both sums become

$$\sum_{i,j} c_{ij} m(A_i \cap B_j),$$

where c_{ij} is the common value on $A_i \cap B_j$. Hence the integral is independent of representation. Linearity follows by refining the partitions for φ and ψ simultaneously and applying finite additivity of measure. If $\varphi \leq \psi$, then $\psi - \varphi$ is a nonnegative simple function, so $\int \psi = \int \varphi + \int (\psi - \varphi) \geq \int \varphi$. \square

Definition 2.3. For a nonnegative measurable function f , define

$$\int f = \sup \left\{ \int \varphi : 0 \leq \varphi \leq f, \varphi \text{ simple} \right\}.$$

For a real-valued measurable function f , put $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$. If at least one of $\int f^+$ and $\int f^-$ is finite, define

$$\int f = \int f^+ - \int f^-.$$

If both are finite, f is called *integrable*.

Proposition 2.4. *For nonnegative measurable functions f, g and constants $a, b \geq 0$,*

- (i) $\int (af + bg) = a \int f + b \int g$;
- (ii) if $f \leq g$, then $\int f \leq \int g$;
- (iii) if $A \cap B = \emptyset$, then $\int_{A \cup B} f = \int_A f + \int_B f$;
- (iv) if $\int f = 0$, then $f = 0$ almost everywhere.

Proof. For simple functions these statements were proved above. For general nonnegative f, g , choose increasing simple approximations $\varphi_n \uparrow f$ and $\psi_n \uparrow g$. Then $a\varphi_n + b\psi_n \uparrow af + bg$, and taking suprema gives linearity. Monotonicity follows directly from the definition because every simple function below f is also below g . The additivity over disjoint sets follows by applying linearity to $f\mathbf{1}_A + f\mathbf{1}_B = f\mathbf{1}_{A \cup B}$.

For (iv), let $E_k = \{x : f(x) > 1/k\}$. Since $(1/k)\mathbf{1}_{E_k} \leq f$, we get $m(E_k)/k \leq \int f = 0$, hence $m(E_k) = 0$. But $\{f > 0\} = \bigcup_k E_k$, so $m(\{f > 0\}) = 0$. \square

Lemma 2.5 (Chebyshev inequality). *If $f \geq 0$ is measurable and $a > 0$, then*

$$m(\{x : f(x) > a\}) \leq \frac{1}{a} \int f.$$

Proof. Let $E_a = \{f > a\}$. Since $a\mathbf{1}_{E_a} \leq f$, monotonicity gives

$$a m(E_a) = \int a\mathbf{1}_{E_a} \leq \int f.$$

Divide by a . \square

2.2 Convergence Theorems

Lemma 2.6 (Fatou). *If $f_n \geq 0$ are measurable, then*

$$\int \liminf_{n \rightarrow \infty} f_n \leq \liminf_{n \rightarrow \infty} \int f_n.$$

Proof. Set $g_n = \inf_{k \geq n} f_k$. Then $g_n \uparrow g = \liminf_n f_n$. By monotonicity, $\int g_n \leq \inf_{k \geq n} \int f_k$. Taking $n \rightarrow \infty$ and using the monotone convergence theorem for the increasing sequence g_n gives

$$\int g = \lim_n \int g_n \leq \lim_n \inf_{k \geq n} \int f_k = \liminf_n \int f_n.$$

It remains only to justify monotone convergence, which is proved in the next theorem; alternatively one may reverse the order and prove Fatou from the definition first. The two arguments are logically equivalent once simple approximation is available. \square

Theorem 2.7 (Monotone convergence). *If $0 \leq f_1 \leq f_2 \leq \dots$ are measurable and $f_n \uparrow f$, then*

$$\lim_{n \rightarrow \infty} \int f_n = \int f.$$

Proof. By monotonicity, $\int f_n \leq \int f$, so $\lim_n \int f_n \leq \int f$. For the reverse inequality, let $0 \leq \varphi \leq f$ be simple and fix $0 < c < 1$. Define

$$E_n = \{x : f_n(x) \geq c\varphi(x)\}.$$

Then $E_n \uparrow X$, because wherever $\varphi(x) = 0$ the condition is automatic, and wherever $\varphi(x) > 0$, $f_n(x) \uparrow f(x) \geq \varphi(x) > c\varphi(x)$. Hence by continuity from below and simple-function integration,

$$\int f_n \geq \int_{E_n} f_n \geq c \int_{E_n} \varphi \rightarrow c \int \varphi.$$

Thus $\liminf_n \int f_n \geq c \int \varphi$. Let $c \uparrow 1$, then take the supremum over all simple $\varphi \leq f$. This gives $\liminf_n \int f_n \geq \int f$. \square

Corollary 2.8 (Termwise integration of nonnegative series). *If $a_k \geq 0$ are measurable, then*

$$\int \sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} \int a_k.$$

If the right side is finite, then $\sum_k a_k(x) < \infty$ for almost every x .

Proof. Let $s_N = \sum_{k=1}^N a_k$. Then $s_N \uparrow \sum_k a_k$. By monotone convergence and finite linearity,

$$\int \sum_k a_k = \lim_N \int s_N = \lim_N \sum_{k=1}^N \int a_k = \sum_k \int a_k.$$

If this common value is finite, the set where $\sum_k a_k = \infty$ must have measure zero; otherwise the integral of the sum would be infinite by Chebyshev's inequality applied to truncations. \square

Theorem 2.9 (Dominated convergence). *Let f_n be measurable, $f_n \rightarrow f$ almost everywhere, and $|f_n| \leq g$ for some integrable g . Then f is integrable and*

$$\lim_{n \rightarrow \infty} \int |f_n - f| = 0, \quad \lim_{n \rightarrow \infty} \int f_n = \int f.$$

Proof. Since $|f| \leq g$ almost everywhere, f is integrable. The functions $2g - |f_n - f|$ are nonnegative and converge almost everywhere to $2g$. Fatou's lemma gives

$$\int 2g \leq \liminf_n \int (2g - |f_n - f|) = 2 \int g - \limsup_n \int |f_n - f|.$$

Therefore $\limsup_n \int |f_n - f| \leq 0$, proving L^1 -convergence. Finally,

$$\left| \int f_n - \int f \right| \leq \int |f_n - f| \rightarrow 0.$$

\square

Theorem 2.10 (Bounded convergence). *Let E have finite measure. Suppose f_n are measurable, $|f_n| \leq M$, and $f_n \rightarrow f$ almost everywhere on E . Then*

$$\int_E f_n \rightarrow \int_E f.$$

Proof. Apply dominated convergence with the integrable dominating function $g = M\mathbf{1}_E$. Then $|f_n| \leq g$, f is integrable, and the integrals converge. \square

Proposition 2.11. *If f is Riemann integrable on $[a, b]$, then f is Lebesgue measurable and*

$$\int_a^b f(x) dx$$

has the same value as the Riemann integral.

Proof. Let P_n be partitions whose mesh tends to zero and whose upper and lower Darboux sums differ by a quantity tending to zero. Let u_n and ℓ_n be the corresponding upper and lower step functions. Then $\ell_n \leq f \leq u_n$, and $\int (u_n - \ell_n) \rightarrow 0$. Put

$$\ell = \sup_n \ell_n, \quad u = \inf_n u_n.$$

Then $\ell \leq f \leq u$, ℓ, u are measurable, and $\int (u - \ell) = 0$. Thus $u = \ell$ almost everywhere, and f agrees almost everywhere with a measurable function, hence is measurable. Since the Lebesgue integrals of u_n and ℓ_n are exactly the Darboux sums and their limits coincide with the Riemann integral, squeezing gives equality of the Riemann and Lebesgue integrals. \square

2.3 Fubini's Theorem on Euclidean Product Spaces

The preceding construction of the Lebesgue integral is one-dimensional in notation, but the same arguments apply on every Euclidean space. The next theorem records the form used constantly in analysis: integration on a product space may be computed by successive integration. Its abstract product-measure proof is given later in [Section 5.4](#); the present statement is the Euclidean specialization.

Theorem 2.12 (Tonelli theorem on $\mathbb{R}^m \times \mathbb{R}^n$). *Let $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow [0, \infty]$ be Lebesgue measurable. Then, for almost every $x \in \mathbb{R}^m$, the section $y \mapsto f(x, y)$ is measurable; the function*

$$x \mapsto \int_{\mathbb{R}^n} f(x, y) dy$$

is measurable; and

$$\int_{\mathbb{R}^{m+n}} f(x, y) d(x, y) = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x, y) dy \right) dx.$$

The analogous formula with the order of x and y reversed also holds.

Proof. Lebesgue measure on \mathbb{R}^{m+n} is the completion of the product of Lebesgue measure on \mathbb{R}^m and Lebesgue measure on \mathbb{R}^n . The sigma-finite product-measure construction and Tonelli theorem in [Section 5.4](#) apply to the two measure spaces $(\mathbb{R}^m, \mathcal{L}^m, m_m)$ and $(\mathbb{R}^n, \mathcal{L}^n, m_n)$. They give the measurability of sections for product-measurable representatives and the equality of the product integral with both iterated integrals. Passing to the Lebesgue completion changes sections only on a null set of parameters: if $N \subset \mathbb{R}^{m+n}$ has product measure zero, Tonelli applied to $\mathbf{1}_N$ gives

$$0 = m_{m+n}(N) = \int_{\mathbb{R}^m} m_n(N_x) dx,$$

so $m_n(N_x) = 0$ for almost every x . Hence modifying a product-measurable representative on N does not change any of the iterated integrals except on a null set of x 's. This proves the theorem for Lebesgue measurable non-negative functions. \square

Theorem 2.13 (Fubini theorem on $\mathbb{R}^m \times \mathbb{R}^n$). *Let $f \in L^1(\mathbb{R}^m \times \mathbb{R}^n)$. Then $f(x, \cdot) \in L^1(\mathbb{R}^n)$ for almost every x , and $f(\cdot, y) \in L^1(\mathbb{R}^m)$ for almost every y . Moreover,*

$$\int_{\mathbb{R}^{m+n}} f(x, y) d(x, y) = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x, y) dy \right) dx = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} f(x, y) dx \right) dy,$$

and both iterated integrals are absolutely integrable.

Proof. Apply Tonelli's theorem to $|f|$. Since $f \in L^1$,

$$\int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} |f(x, y)| dy \right) dx = \int_{\mathbb{R}^{m+n}} |f(x, y)| d(x, y) < \infty.$$

Therefore the inner integral of $|f|$ is finite for almost every x , which means $f(x, \cdot) \in L^1(\mathbb{R}^n)$ for almost every x , and the resulting function of x is integrable. The same argument with x and y interchanged gives the corresponding statement for horizontal sections. Finally write $f = f^+ - f^-$ in the real-valued case, or split into real and imaginary parts in the complex-valued case, and apply Tonelli to the non-negative parts. Subtracting the finite equalities gives the desired identities. \square

2.4 Absolute Continuity of the Integral and Uniform Integrability

Theorem 2.14 (Continuity of integration over sets). *Let f be integrable on E .*

(i) *If $E_n \uparrow E$, then $\int_{E_n} f \rightarrow \int_E f$.*

(ii) *If $E_n \downarrow E$, then $\int_{E_n} f \rightarrow \int_E f$.*

Proof. It suffices to consider $f \geq 0$, then apply the result to f^+ and f^- . If $E_n \uparrow E$, then $f\mathbf{1}_{E_n} \uparrow f\mathbf{1}_E$, so monotone convergence gives the result. If $E_n \downarrow E$, then $f\mathbf{1}_{E_1 \setminus E_n} \uparrow f\mathbf{1}_{E_1 \setminus E}$. Therefore

$$\int_{E_n} f = \int_{E_1} f - \int_{E_1 \setminus E_n} f \rightarrow \int_{E_1} f - \int_{E_1 \setminus E} f = \int_E f.$$

□

Proposition 2.15 (Absolute continuity of the integral). *Let $f \in L^1(E)$. For every $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $A \subset E$ is measurable and $m(A) < \delta$,*

$$\int_A |f| < \varepsilon.$$

Proof. Choose N such that $\int_{\{|f| > N\}} |f| < \varepsilon/2$, possible by continuity of integration applied to the decreasing sets $\{|f| > N\}$. Let $\delta = \varepsilon/(2N)$. If $m(A) < \delta$, then

$$\int_A |f| = \int_{A \cap \{|f| \leq N\}} |f| + \int_{A \cap \{|f| > N\}} |f| \leq Nm(A) + \varepsilon/2 < \varepsilon.$$

□

Definition 2.16. A family $\mathcal{F} \subset L^1(E)$ is *equi-integrable* if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every measurable $A \subset E$ with $m(A) < \delta$,

$$\sup_{f \in \mathcal{F}} \int_A |f| < \varepsilon.$$

It is *tight* on E if for every $\varepsilon > 0$, there exists a measurable $E_0 \subset E$ with $m(E_0) < \infty$ such that

$$\sup_{f \in \mathcal{F}} \int_{E \setminus E_0} |f| < \varepsilon.$$

Theorem 2.17 (Vitali convergence on finite-measure sets). *Let $m(E) < \infty$. Suppose $f_n \rightarrow f$ almost everywhere on E and $\{f_n\}$ is equi-integrable. Then $f \in L^1(E)$ and*

$$\int_E |f_n - f| \rightarrow 0.$$

Proof. By Fatou, for every measurable A ,

$$\int_A |f| \leq \liminf_n \int_A |f_n|,$$

so f is integrable on E because equi-integrability on finite measure sets implies boundedness of the integrals. Fix $\varepsilon > 0$. Choose $\delta > 0$ from equi-integrability for the family $\{f_n\} \cup \{f\}$ with error $\varepsilon/3$. By Egorov's theorem, choose $A \subset E$ such that $m(E \setminus A) < \delta$ and $f_n \rightarrow f$ uniformly on A . Then for large n ,

$$\int_A |f_n - f| < \varepsilon/3.$$

On the complement,

$$\int_{E \setminus A} |f_n - f| \leq \int_{E \setminus A} |f_n| + \int_{E \setminus A} |f| < 2\varepsilon/3.$$

Thus $\int_E |f_n - f| < \varepsilon$ for all large n . \square

Theorem 2.18 (General Vitali convergence). *Let $f_n \rightarrow f$ almost everywhere on E . If $\{f_n\}$ is both equi-integrable and tight, then $f \in L^1(E)$ and $f_n \rightarrow f$ in $L^1(E)$.*

Proof. Fix $\varepsilon > 0$. By tightness, choose $E_0 \subset E$ of finite measure such that $\int_{E \setminus E_0} |f_n| < \varepsilon$ for all n . Fatou gives $\int_{E \setminus E_0} |f| \leq \varepsilon$. On E_0 , the finite-measure Vitali theorem gives $\int_{E_0} |f_n - f| \rightarrow 0$. Therefore

$$\int_E |f_n - f| \leq \int_{E_0} |f_n - f| + \int_{E \setminus E_0} |f_n| + \int_{E \setminus E_0} |f| \leq \int_{E_0} |f_n - f| + 2\varepsilon.$$

Let $n \rightarrow \infty$, then $\varepsilon \downarrow 0$. \square

Definition 2.19. A sequence f_n converges to f in measure on E if for every $\eta > 0$,

$$m(\{x \in E : |f_n(x) - f(x)| > \eta\}) \rightarrow 0.$$

Proposition 2.20. *If $m(E) < \infty$ and $f_n \rightarrow f$ almost everywhere on E , then $f_n \rightarrow f$ in measure.*

Proof. Fix $\eta > 0$ and $\varepsilon > 0$. By Egorov's theorem, choose $A \subset E$ such that $m(E \setminus A) < \varepsilon$ and $f_n \rightarrow f$ uniformly on A . For all sufficiently large n , $|f_n - f| \leq \eta$ on A . Thus

$$\{|f_n - f| > \eta\} \subset E \setminus A$$

for large n , so its measure is less than ε . Since ε is arbitrary, the measure tends to zero. \square

Theorem 2.21 (Riesz subsequence theorem). *If $f_n \rightarrow f$ in measure on E , then there exists a subsequence f_{n_k} such that $f_{n_k} \rightarrow f$ almost everywhere.*

Proof. For each k , choose $n_k > n_{k-1}$ such that

$$m(\{|f_{n_k} - f| > 2^{-k}\}) < 2^{-k}.$$

Let $E_k = \{|f_{n_k} - f| > 2^{-k}\}$. Since $\sum_k m(E_k) < \infty$, Borel–Cantelli implies that almost every x belongs to only finitely many E_k . For such x , there exists $K(x)$ such that for $k \geq K(x)$, $|f_{n_k}(x) - f(x)| \leq 2^{-k}$. Hence $f_{n_k}(x) \rightarrow f(x)$. \square

Corollary 2.22. *Let $f_n \geq 0$ be integrable and $f_n \rightarrow f$ in measure. Then $\int |f_n - f| \rightarrow 0$ if and only if $\{f_n\}$ is equi-integrable and tight.*

Proof. If $f_n \rightarrow f$ in L^1 , then $\{f_n\} \cup \{f\}$ is equi-integrable and tight because a single L^1 function has both properties and $\int |f_n - f|$ is small for large n ; finitely many exceptional terms do not affect the properties. Conversely, if $\{f_n\}$ is equi-integrable and tight, every subsequence has a further subsequence converging almost everywhere to f by the Riesz theorem. The general Vitali theorem then gives L^1 -convergence along this further subsequence. If the whole sequence did not converge in L^1 , some subsequence would satisfy $\int |f_{n_j} - f| \geq \varepsilon > 0$; this subsequence has a further subsequence converging in L^1 , a contradiction. \square

Theorem 2.23 (Completeness of L^1). *The space L^1 , modulo equality almost everywhere, is complete under the norm $\|f\|_1 = \int |f|$.*

Proof. Let f_n be Cauchy in L^1 . Choose a subsequence f_{n_k} such that $\|f_{n_{k+1}} - f_{n_k}\|_1 < 2^{-k}$. The nonnegative series

$$\sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}|$$

has finite integral, hence is finite almost everywhere. Thus $f_{n_k}(x)$ converges almost everywhere to some function $f(x)$. Moreover,

$$|f_{n_k} - f| \leq \sum_{j=k}^{\infty} |f_{n_{j+1}} - f_{n_j}|$$

almost everywhere, and integrating gives $\|f_{n_k} - f\|_1 \leq \sum_{j=k}^{\infty} 2^{-j} \rightarrow 0$. Since the original sequence is Cauchy, $f_n \rightarrow f$ in L^1 . \square

Chapter 3

Differentiation and Integration

3.1 Hardy–Littlewood Maximal Function and Differentiation

Definition 3.1. For $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, the Hardy–Littlewood maximal function is

$$Mf(x) = \sup_{B \ni x} \frac{1}{m(B)} \int_B |f(y)| \, dy,$$

where the supremum is taken over all balls $B \subset \mathbb{R}^d$ containing x .

Lemma 3.2 (Vitali covering lemma for balls). *Let $\mathcal{B} = \{B_1, \dots, B_N\}$ be a finite collection of balls in \mathbb{R}^d . Then there is a disjoint subcollection B_{j_1}, \dots, B_{j_k} such that*

$$\bigcup_{i=1}^N B_i \subset \bigcup_{\ell=1}^k 3B_{j_\ell},$$

where $3B$ is the ball with the same center and three times the radius. Consequently,

$$m\left(\bigcup_{i=1}^N B_i\right) \leq 3^d \sum_{\ell=1}^k m(B_{j_\ell}).$$

Proof. Choose from \mathcal{B} a ball of largest radius, call it B_{j_1} . Remove it and all balls intersecting it. From the remaining collection, choose a ball of largest radius, and repeat until no balls remain. The chosen balls are disjoint. If a removed ball B intersects a chosen ball B_j , then $r(B) \leq r(B_j)$ by the maximal choice at that step. Every point of B lies within distance at most $r(B) + 2r(B_j) \leq 3r(B_j)$ from the center of B_j , hence $B \subset 3B_j$. This proves the covering inclusion. Taking measures gives the inequality because $m(3B_j) = 3^d m(B_j)$. \square

Theorem 3.3 (Weak type $(1, 1)$ estimate). *If $f \in L^1(\mathbb{R}^d)$, then Mf is measurable, finite almost everywhere, and for every $\alpha > 0$,*

$$m(\{x : Mf(x) > \alpha\}) \leq \frac{C_d}{\alpha} \|f\|_1,$$

where one may take $C_d = 3^d$.

Proof. Measurability follows by restricting the supremum to balls with rational centers and rational radii; this countable supremum gives the same value because balls can be approximated from inside and outside. Let $E_\alpha = \{Mf > \alpha\}$. For every $x \in E_\alpha$, choose a ball $B_x \ni x$ such that

$$\int_{B_x} |f| > \alpha m(B_x).$$

For a compact $K \subset E_\alpha$, finitely many such balls cover K . Apply the Vitali covering lemma to obtain disjoint balls B_1, \dots, B_N with $K \subset \bigcup_j 3B_j$. Then

$$m(K) \leq 3^d \sum_j m(B_j) < \frac{3^d}{\alpha} \sum_j \int_{B_j} |f| \leq \frac{3^d}{\alpha} \|f\|_1.$$

By inner regularity, take the supremum over compact $K \subset E_\alpha$ to get the stated inequality. Since $m(\{Mf > n\}) \leq C_d \|f\|_1 / n \rightarrow 0$, the set where $Mf = \infty$ has measure zero. \square

Theorem 3.4 (Lebesgue differentiation theorem). *If $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, then for almost every $x \in \mathbb{R}^d$,*

$$\lim_{r \downarrow 0} \frac{1}{m(B(x, r))} \int_{B(x, r)} |f(y) - f(x)| \, dy = 0.$$

Consequently,

$$\lim_{r \downarrow 0} \frac{1}{m(B(x, r))} \int_{B(x, r)} f(y) \, dy = f(x)$$

for almost every x .

Proof. First assume $f \in L^1(\mathbb{R}^d)$. For a continuous compactly supported g , uniform continuity gives

$$\frac{1}{m(B(x, r))} \int_{B(x, r)} |g(y) - g(x)| \, dy \rightarrow 0$$

for every x . Given $\varepsilon > 0$, choose such a g with $\|f - g\|_1 < \varepsilon$, possible by density. Define

$$A_t = \left\{ x : \limsup_{r \downarrow 0} \frac{1}{m(B(x, r))} \int_{B(x, r)} |f(y) - f(x)| \, dy > t \right\}.$$

Using $f - g$ and the pointwise term $|f(x) - g(x)|$, one obtains

$$A_t \subset \{M(f - g) > t/3\} \cup \{|f - g| > t/3\}.$$

The weak maximal estimate and Chebyshev inequality give

$$m(A_t) \leq C \frac{\|f - g\|_1}{t} + \frac{3}{t} \|f - g\|_1 \leq C_t \varepsilon.$$

Since ε is arbitrary, $m(A_t) = 0$ for every rational $t > 0$, and the result follows. For locally integrable f , apply the argument to $f \mathbf{1}_{B(0, N)}$ on each ball $B(0, N/2)$, then take the countable union over N . \square

Definition 3.5. A point x is a *Lebesgue density point* of a measurable set E if

$$\lim_{r \downarrow 0} \frac{m(E \cap B(x, r))}{m(B(x, r))} = 1.$$

The *Lebesgue set* of a locally integrable function f is the set of points x for which the first limit in the differentiation theorem is 0.

Corollary 3.6 (Density theorem). *If $E \subset \mathbb{R}^d$ is measurable, then almost every point of E is a density point of E , and almost every point of E^c is not a density point of E .*

Proof. Apply the Lebesgue differentiation theorem to $f = \mathbf{1}_E$. At almost every point x , averages of $\mathbf{1}_E$ over balls centered at x converge to $\mathbf{1}_E(x)$. If $x \in E$, this limit is 1; if $x \in E^c$, it is 0. The two assertions follow. \square

3.2 Bounded Variation and Absolute Continuity

Definition 3.7. For a complex-valued function F on $[a, b]$, its variation over a partition $P : a = t_0 < \cdots < t_N = b$ is

$$V(F, P) = \sum_{j=1}^N |F(t_j) - F(t_{j-1})|.$$

The total variation is $T_F(a, b) = \sup_P V(F, P)$. The function F is of *bounded variation*, written $F \in BV([a, b])$, if $T_F(a, b) < \infty$.

Proposition 3.8. *A plane curve $\gamma(t) = (x(t), y(t))$, $a \leq t \leq b$, is rectifiable if and only if both x and y are of bounded variation. In that case the length of γ is comparable to the variations of x and y .*

Proof. Let $F(t) = x(t) + iy(t)$. For each partition,

$$|x(t_j) - x(t_{j-1})| \leq |F(t_j) - F(t_{j-1})|, \quad |y(t_j) - y(t_{j-1})| \leq |F(t_j) - F(t_{j-1})|.$$

Hence rectifiability of the curve implies bounded variation of x and y . Conversely,

$$|F(t_j) - F(t_{j-1})| \leq |x(t_j) - x(t_{j-1})| + |y(t_j) - y(t_{j-1})|.$$

Taking sums and suprema shows that bounded variation of x and y implies finite curve length. \square

Lemma 3.9. *If F is real-valued and of bounded variation on $[a, b]$, then for $a \leq x \leq b$, with*

$$P_F(a, x) = \sup_P \sum_j (F(t_j) - F(t_{j-1}))^+, \quad N_F(a, x) = \sup_P \sum_j (F(t_j) - F(t_{j-1}))^-,$$

we have

$$F(x) - F(a) = P_F(a, x) - N_F(a, x), \quad T_F(a, x) = P_F(a, x) + N_F(a, x).$$

Proof. For a fixed partition, write each increment $\Delta_j = F(t_j) - F(t_{j-1})$. Then $\Delta_j = \Delta_j^+ - \Delta_j^-$ and $|\Delta_j| = \Delta_j^+ + \Delta_j^-$. Summing gives

$$F(x) - F(a) = \sum_j \Delta_j^+ - \sum_j \Delta_j^-.$$

Taking suprema over refinements yields the stated decomposition. More explicitly, if a partition nearly realizes total variation, then the positive and negative sums nearly realize the corresponding suprema; otherwise their sum would be too small. Hence the two displayed identities hold. \square

Theorem 3.10. *A real-valued function F on $[a, b]$ has bounded variation if and only if it is the difference of two bounded increasing functions.*

Proof. If $F = G - H$, with G, H bounded increasing, then for every partition,

$$\sum_j |F(t_j) - F(t_{j-1})| \leq \sum_j (G(t_j) - G(t_{j-1})) + \sum_j (H(t_j) - H(t_{j-1})) \leq G(b) - G(a) + H(b) - H(a),$$

so F has bounded variation. Conversely, if $F \in BV$, define

$$G(x) = P_F(a, x), \quad H(x) = N_F(a, x) - F(x).$$

Both P_F and N_F are increasing and bounded by $T_F(a, b)$. The preceding lemma gives $F(x) = G(x) - H(x)$. \square

Theorem 3.11. *Every real-valued function of bounded variation on $[a, b]$ is differentiable almost everywhere.*

Proof. By the preceding theorem, it is enough to prove the result for increasing functions. Let F be increasing. Define the upper and lower Dini derivatives

$$D^+F(x) = \limsup_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h}, \quad D_-F(x) = \liminf_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h}.$$

For rationals $\alpha > \beta$, consider

$$E_{\alpha, \beta} = \{x : D^+F(x) > \alpha > \beta > D_-F(x)\}.$$

Using the rising-sun lemma applied to the continuous approximations $F(x) - \alpha x$ and $F(x) - \beta x$ on open covers of $E_{\alpha, \beta}$, one obtains

$$\alpha m(E_{\alpha, \beta}) \leq \beta m(E_{\alpha, \beta}).$$

Since $\alpha > \beta$, this forces $m(E_{\alpha, \beta}) = 0$. Taking the countable union over rational $\alpha > \beta$ shows that the upper and lower Dini derivatives agree almost everywhere. Hence F' exists almost everywhere. The reduction from BV to increasing functions preserves differentiability almost everywhere. \square

Definition 3.12. A function $F : [a, b] \rightarrow \mathbb{C}$ is *absolutely continuous* if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every finite disjoint collection of intervals $(a_j, b_j) \subset [a, b]$,

$$\sum_j (b_j - a_j) < \delta \implies \sum_j |F(b_j) - F(a_j)| < \varepsilon.$$

Proposition 3.13. *If F is absolutely continuous on $[a, b]$, then F is uniformly continuous and of bounded variation.*

Proof. Uniform continuity follows by applying the definition to a single interval: if $|x - y| < \delta$, then $|F(x) - F(y)| < \varepsilon$. For bounded variation, choose $\delta > 0$ corresponding to $\varepsilon = 1$. Divide $[a, b]$ into finitely many intervals of length less than δ . On each such interval, the variation of F over any subpartition is at most 1 by absolute continuity. Summing over finitely many intervals gives finite total variation. \square

Proposition 3.14. *If $f \in L^1([a, b])$ and*

$$F(x) = F(a) + \int_a^x f(t) dt,$$

then F is absolutely continuous.

Proof. Given $\varepsilon > 0$, by absolute continuity of the integral choose $\delta > 0$ such that $m(A) < \delta$ implies $\int_A |f| < \varepsilon$. If (a_j, b_j) are disjoint and $\sum_j (b_j - a_j) < \delta$, then with $A = \bigcup_j (a_j, b_j)$,

$$\sum_j |F(b_j) - F(a_j)| \leq \sum_j \int_{a_j}^{b_j} |f| = \int_A |f| < \varepsilon.$$

Thus F is absolutely continuous. \square

Theorem 3.15 (Fundamental theorem of calculus). *For a function $F : [a, b] \rightarrow \mathbb{R}$, the following are equivalent:*

- (i) F is absolutely continuous;

(ii) F' exists almost everywhere, $F' \in L^1([a, b])$, and

$$F(x) = F(a) + \int_a^x F'(t) dt \quad (a \leq x \leq b).$$

Proof. The implication (ii) \Rightarrow (i) is exactly the preceding proposition with $f = F'$. For (i) \Rightarrow (ii), absolute continuity implies bounded variation, hence differentiability almost everywhere. By applying the differentiation theorem to the variation measure associated to F , one obtains $F' \in L^1$ and

$$\int_a^b |F'| \leq T_F(a, b).$$

Define $G(x) = F(a) + \int_a^x F'(t) dt$. Then G is absolutely continuous and $G' = F'$ almost everywhere. Hence $H = F - G$ is absolutely continuous and $H' = 0$ almost everywhere. We show that such an H is constant. Given $\varepsilon > 0$, choose δ from absolute continuity. Since $H' = 0$ almost everywhere, the set where the upper derivative of H exceeds $\varepsilon/(b-a)$ has measure zero by the Dini-derivative argument. Cover this null set by open intervals of total length less than δ , and on the complement use the derivative-zero condition to make all increments arbitrarily small. Summing over a partition of $[a, b]$ gives $|H(x) - H(a)| < \varepsilon$. Since ε is arbitrary, $H(x) = H(a)$, so $F = G$. \square

Theorem 3.16 (Lebesgue decomposition for BV functions). *If $F \in BV([a, b])$, then*

$$F = G + H,$$

where G is absolutely continuous and $H' = 0$ almost everywhere. One can take

$$G(x) = F(a) + \int_a^x F'(t) dt.$$

Proof. Since $F \in BV$, F' exists almost everywhere and $F' \in L^1$ after decomposing F into the difference of two increasing functions and applying the derivative estimate for monotone functions. Define $G(x) = F(a) + \int_a^x F'(t) dt$. Then G is absolutely continuous and $G' = F'$ almost everywhere. Put $H = F - G$. Then $H' = F' - G' = 0$ almost everywhere. \square

3.3 Convex Functions and Jensen's Inequality

Definition 3.17. A real-valued function ϕ on an interval (a, b) is *convex* if for all $x, y \in (a, b)$ and $0 \leq \lambda \leq 1$,

$$\phi(\lambda x + (1 - \lambda)y) \leq \lambda\phi(x) + (1 - \lambda)\phi(y).$$

Proposition 3.18. *If ϕ is convex on (a, b) , then:*

- (i) ϕ is continuous;
- (ii) the left and right derivatives exist at every point;
- (iii) ϕ is locally Lipschitz;
- (iv) ϕ is differentiable almost everywhere.

Proof. For $a < x < y < z < b$, convexity gives the slope monotonicity

$$\frac{\phi(y) - \phi(x)}{y - x} \leq \frac{\phi(z) - \phi(x)}{z - x} \leq \frac{\phi(z) - \phi(y)}{z - y}.$$

This is obtained by writing y as a convex combination of x, z and rearranging. The monotonicity of difference quotients implies the existence of one-sided derivatives as monotone limits. On every compact subinterval $[c, d] \subset (a, b)$, the slopes are bounded between two finite secant slopes involving points slightly outside $[c, d]$, hence ϕ is Lipschitz on $[c, d]$, and therefore continuous. A locally Lipschitz function is of bounded variation on compact subintervals, so it is differentiable almost everywhere by the BV differentiability theorem. \square

Proposition 3.19 (Supporting line). *If ϕ is convex on (a, b) , then for every $x_0 \in (a, b)$ there exists $m \in \mathbb{R}$ such that*

$$\phi(x) \geq \phi(x_0) + m(x - x_0) \quad \text{for all } x \in (a, b).$$

Proof. Let m_- and m_+ be the left and right derivatives at x_0 . Slope monotonicity gives $m_- \leq m_+$. Choose any $m \in [m_-, m_+]$. If $x > x_0$, then slope monotonicity gives $(\phi(x) - \phi(x_0))/(x - x_0) \geq m_+ \geq m$. If $x < x_0$, then $(\phi(x_0) - \phi(x))/(x_0 - x) \leq m_- \leq m$, which rearranges to the same inequality. Hence the line with slope m supports the graph from below. \square

Theorem 3.20 (Jensen). *Let ϕ be convex on an interval containing the range of an integrable function f on a probability space (X, \mathcal{M}, μ) . If $\phi \circ f$ is integrable, then*

$$\phi\left(\int_X f \, d\mu\right) \leq \int_X \phi(f) \, d\mu.$$

Proof. Let $x_0 = \int f \, d\mu$. By the supporting line proposition, choose m such that $\phi(x) \geq \phi(x_0) + m(x - x_0)$ for all x in the interval. Substitute $x = f(t)$ and integrate:

$$\int \phi(f) \, d\mu \geq \int (\phi(x_0) + m(f - x_0)) \, d\mu = \phi(x_0) + m\left(\int f \, d\mu - x_0\right) = \phi(x_0).$$

This is the desired inequality. \square

3.4 Convolution and Approximation

Convolution is the operation that turns translation-invariant integration into an algebraic product. It is the analytic mechanism behind smoothing, approximation to the identity, and many later constructions in harmonic analysis.

Definition 3.21 (Translation and convolution). For $h \in \mathbb{R}^d$, write

$$\tau_h f(x) = f(x - h).$$

If f, g are measurable functions on \mathbb{R}^d , their *convolution* is

$$(f * g)(x) = \int_{\mathbb{R}^d} f(y)g(x - y) \, dy,$$

whenever the integral is defined. Equivalently, by the change of variables $y \mapsto x - y$, one may write

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x - y)g(y) \, dy.$$

Proposition 3.22 (Algebraic properties for compactly supported functions). *If $f, g, h \in C_c(\mathbb{R}^d)$, then $f * g \in C_c(\mathbb{R}^d)$,*

$$f * g = g * f, \quad (f * g) * h = f * (g * h),$$

and

$$\text{supp}(f * g) \subseteq \text{supp}(f) + \text{supp}(g) := \{a + b : a \in \text{supp}(f), b \in \text{supp}(g)\}.$$

Proof. The integral defining $f * g$ is finite because the integrand is continuous and compactly supported in y . Commutativity follows from the change of variables $z = x - y$:

$$(f * g)(x) = \int f(y)g(x - y) dy = \int f(x - z)g(z) dz = (g * f)(x).$$

For associativity, all integrands below are absolutely integrable because the functions are continuous with compact support. Fubini's theorem gives

$$\begin{aligned} ((f * g) * h)(x) &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(u)g(y - u) du \right) h(x - y) dy \\ &= \int_{\mathbb{R}^d} f(u) \left(\int_{\mathbb{R}^d} g(y - u)h(x - y) dy \right) du. \end{aligned}$$

In the inner integral put $v = y - u$. Then $x - y = x - u - v$, so the inner integral is

$$\int_{\mathbb{R}^d} g(v)h(x - u - v) dv = (g * h)(x - u).$$

Thus $((f * g) * h)(x) = \int f(u)(g * h)(x - u) du = (f * (g * h))(x)$.

The support inclusion follows directly: if $(f * g)(x) \neq 0$, then for some y in the support of the integrand we must have $y \in \text{supp}(f)$ and $x - y \in \text{supp}(g)$, hence $x \in \text{supp}(f) + \text{supp}(g)$. Since the sum of two compact sets is compact, this also shows that $f * g$ has compact support. Continuity follows from uniform continuity on the compact set containing all relevant translates: if $x_n \rightarrow x$, then $g(x_n - y) \rightarrow g(x - y)$ uniformly for $y \in \text{supp}(f)$, hence $(f * g)(x_n) \rightarrow (f * g)(x)$. \square

Theorem 3.23 (Young inequality: $L^1 * L^p$ form). *If $f \in L^1(\mathbb{R}^d)$ and $g \in L^p(\mathbb{R}^d)$, $1 \leq p \leq \infty$, then $f * g$ is defined for almost every x , belongs to $L^p(\mathbb{R}^d)$, and*

$$\|f * g\|_p \leq \|f\|_1 \|g\|_p.$$

In particular $L^1(\mathbb{R}^d)$, with convolution as multiplication, is a normed algebra.

Proof. For $p = \infty$, for almost every x ,

$$|(f * g)(x)| \leq \int |f(y)| |g(x - y)| dy \leq \|g\|_\infty \|f\|_1.$$

Thus $\|f * g\|_\infty \leq \|f\|_1 \|g\|_\infty$.

Let $1 \leq p < \infty$. By Minkowski's integral inequality and translation invariance of Lebesgue measure,

$$\begin{aligned} \|f * g\|_p &= \left(\int \left| \int f(y)g(x - y) dy \right|^p dx \right)^{1/p} \\ &\leq \int |f(y)| \left(\int |g(x - y)|^p dx \right)^{1/p} dy \\ &= \int |f(y)| \|g\|_p dy = \|f\|_1 \|g\|_p. \end{aligned}$$

The same estimate applied to $p = 1$ gives $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$, so L^1 is closed under convolution and the norm is submultiplicative. The associativity on L^1 follows by approximating L^1 functions by C_c functions and using the inequality just proved to pass to the limit from the compactly supported case. \square

Proposition 3.24 (Holder endpoint). *Let $1 \leq p \leq \infty$, and let q be its conjugate exponent. If $f \in L^p(\mathbb{R}^d)$ and $g \in L^q(\mathbb{R}^d)$, then $f * g \in L^\infty(\mathbb{R}^d)$ and*

$$\|f * g\|_\infty \leq \|f\|_p \|g\|_q.$$

Proof. For every x for which the convolution is defined, Holder's inequality gives

$$|(f * g)(x)| \leq \int |f(y)| |g(x - y)| dy \leq \|f\|_p \|\tau_x \tilde{g}\|_q,$$

where $\tilde{g}(y) = g(-y)$. Translation and reflection preserve the L^q -norm, so $\|\tau_x \tilde{g}\|_q = \|g\|_q$. Taking the essential supremum over x proves the estimate. \square

Proposition 3.25 (Smoothing by convolution). *Let $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ and $\phi \in C_c^\infty(\mathbb{R}^d)$. Then $f * \phi \in C^\infty(\mathbb{R}^d)$, and for every multi-index α ,*

$$\partial^\alpha (f * \phi) = f * (\partial^\alpha \phi).$$

Proof. Fix a compact set $K \subset \mathbb{R}^d$. If $x \in K$ and $\phi(x - y) \neq 0$, then $y \in K - \text{supp}(\phi)$, a compact set. Since $f \in L^1_{\text{loc}}$, all integrals involved are finite on K . For first derivatives, consider a coordinate vector e_j . The difference quotient is

$$\frac{(f * \phi)(x + te_j) - (f * \phi)(x)}{t} = \int f(y) \frac{\phi(x + te_j - y) - \phi(x - y)}{t} dy.$$

For $x \in K$ and t small, the integrand is supported in one compact set and is dominated by $|f(y)| \sup |\partial_j \phi|$, which is integrable on that compact set. The pointwise limit of the difference quotient is $f(y) \partial_j \phi(x - y)$. Dominated convergence therefore gives

$$\partial_j (f * \phi) = f * (\partial_j \phi).$$

Repeating the argument for higher derivatives proves the formula for every multi-index α , and also proves continuity of those derivatives. \square

Definition 3.26 (Good kernels). A family $K_\delta \in L^1(\mathbb{R}^d)$, $\delta > 0$, is a family of *good kernels*, or an *approximation to the identity*, if

- (i) $\int K_\delta = 1$;
- (ii) $\sup_\delta \|K_\delta\|_1 < \infty$;
- (iii) for every $\eta > 0$,

$$\int_{|x| > \eta} |K_\delta(x)| dx \rightarrow 0 \quad \text{as } \delta \downarrow 0.$$

Example 3.27 (Standard mollifiers). Let $\rho \in C_c^\infty(\mathbb{R}^d)$, $\rho \geq 0$, and $\int \rho = 1$. Define

$$\rho_\delta(x) = \delta^{-d} \rho(x/\delta).$$

Then $\{\rho_\delta\}_{\delta > 0}$ is an approximation to the identity. Indeed, $\int \rho_\delta = 1$ by the change of variables $u = x/\delta$, $\|\rho_\delta\|_1 = \|\rho\|_1 = 1$, and for fixed $\eta > 0$, the support of ρ_δ lies in $\{|x| < \eta\}$ for all sufficiently small δ .

Theorem 3.28 (Approximation by good kernels). *Let K_δ be an approximation to the identity.*

- (i) *If f is bounded and uniformly continuous on \mathbb{R}^d , then $K_\delta * f \rightarrow f$ uniformly.*
- (ii) *If $f \in L^p(\mathbb{R}^d)$, $1 \leq p < \infty$, then*

$$\|K_\delta * f - f\|_p \rightarrow 0.$$

Proof. For bounded uniformly continuous f , write

$$K_\delta * f(x) - f(x) = \int K_\delta(y)(f(x-y) - f(x)) dy.$$

Let $C = \sup_\delta \|K_\delta\|_1$. Given $\varepsilon > 0$, choose $\eta > 0$ such that $|f(x-y) - f(x)| < \varepsilon$ whenever $|y| < \eta$. Then

$$\begin{aligned} |K_\delta * f(x) - f(x)| &\leq \int_{|y| < \eta} |K_\delta(y)| \varepsilon dy + \int_{|y| \geq \eta} |K_\delta(y)| 2 \|f\|_\infty dy \\ &\leq C\varepsilon + 2 \|f\|_\infty \int_{|y| \geq \eta} |K_\delta(y)| dy. \end{aligned}$$

Taking the supremum in x , then letting $\delta \downarrow 0$, gives $\limsup_\delta \|K_\delta * f - f\|_\infty \leq C\varepsilon$. Since ε is arbitrary, uniform convergence follows.

For $1 \leq p < \infty$, first suppose $f \in C_c(\mathbb{R}^d)$. The first part gives uniform convergence. Since f has compact support and the kernels have uniformly small tails, the functions $K_\delta * f - f$ are small outside a fixed large compact set up to an arbitrarily small L^p -error; inside that compact set, uniform convergence implies L^p -convergence. More explicitly, choose R so large that $\text{supp}(f) \subset B_R$. Split the integral into B_{R+1} and its complement. On B_{R+1} , uniform convergence gives convergence in L^p . Outside B_{R+1} , if $|x| > R+1$ and $|y| < 1$, then $f(x-y) = 0$, so only the tail $|y| \geq 1$ contributes, and this tail tends to zero by the good-kernel condition and Young's inequality.

For general $f \in L^p$, choose $g \in C_c(\mathbb{R}^d)$ with $\|f - g\|_p < \varepsilon$. By Young's inequality,

$$\begin{aligned} \|K_\delta * f - f\|_p &\leq \|K_\delta * (f - g)\|_p + \|K_\delta * g - g\|_p + \|g - f\|_p \\ &\leq C \|f - g\|_p + \|K_\delta * g - g\|_p + \|g - f\|_p \\ &< (C + 1)\varepsilon + \|K_\delta * g - g\|_p. \end{aligned}$$

Let $\delta \downarrow 0$, then let $\varepsilon \downarrow 0$. □

Corollary 3.29 (Density of smooth compactly supported functions). *For $1 \leq p < \infty$, the space $C_c^\infty(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$.*

Proof. Given $f \in L^p$ and $\varepsilon > 0$, choose $g \in C_c(\mathbb{R}^d)$ with $\|f - g\|_p < \varepsilon/2$. Let ρ_δ be a standard mollifier. By the approximation theorem, $\rho_\delta * g \rightarrow g$ in L^p . By the smoothing proposition, $\rho_\delta * g \in C^\infty$, and its support is contained in $\text{supp}(g) + \text{supp}(\rho_\delta)$, which is compact. For sufficiently small δ , $\|\rho_\delta * g - g\|_p < \varepsilon/2$, so $\rho_\delta * g \in C_c^\infty$ approximates f within ε . □

Chapter 4

L^p Spaces and Weak Convergence

4.1 Normed, Banach, and Hilbert Spaces

Definition 4.1. A normed vector space $(X, \|\cdot\|)$ is a vector space with a function $\|\cdot\| : X \rightarrow [0, \infty)$ satisfying positivity, homogeneity, and the triangle inequality. It is a *Banach space* if the metric $d(x, y) = \|x - y\|$ is complete. A *Hilbert space* is a complete inner-product space.

Proposition 4.2. $L^2(\mathbb{R}^d)$, with

$$\langle f, g \rangle = \int_{\mathbb{R}^d} f(x) \overline{g(x)} dx,$$

is a Hilbert space.

Proof. The inner product axioms follow from linearity of the integral and the positivity of $\int |f|^2$. Cauchy–Schwarz, proved as the case $p = q = 2$ of Holder’s inequality below, gives $|\langle f, g \rangle| \leq \|f\|_2 \|g\|_2$, so the inner product is well-defined. Completeness follows from the general Riesz–Fischer theorem for L^p , proved below. \square

Proposition 4.3. $L^2(\mathbb{R}^d)$ is separable.

Proof. Continuous compactly supported functions are dense in L^2 , and step functions with rational coefficients supported on finite unions of rectangles with rational coordinates are dense among such functions. More explicitly, approximate any $f \in L^2$ by truncating it to a large cube and bounding its range, then approximate by a simple function. By regularity, approximate the measurable level sets by finite unions of rational rectangles. Finally approximate coefficients by complex numbers with rational real and imaginary parts. The collection of all such rational step functions is countable and dense. \square

4.2 Basic Inequalities

Definition 4.4. For $1 \leq p < \infty$,

$$L^p(E) = \left\{ f : \int_E |f|^p < \infty \right\}, \quad \|f\|_p = \left(\int_E |f|^p \right)^{1/p}.$$

Also

$$L^\infty(E) = \{f : |f| \leq M \text{ a.e. for some } M\}, \quad \|f\|_\infty = \operatorname{ess\,sup}_E |f|.$$

Functions equal almost everywhere are identified.

Lemma 4.5 (Young inequality). *If $a, b \geq 0$, $1 < p < \infty$, and q is conjugate to p , i.e. $1/p + 1/q = 1$, then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Proof. The function \log is concave, so with weights $1/p$ and $1/q$,

$$\log\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \geq \frac{1}{p}\log(a^p) + \frac{1}{q}\log(b^q) = \log(ab)$$

when $a, b > 0$. Exponentiating gives the inequality. If $a = 0$ or $b = 0$, it is trivial. \square

Theorem 4.6 (Holder inequality). *If $1 < p < \infty$, q is conjugate to p , $f \in L^p(E)$, and $g \in L^q(E)$, then $fg \in L^1(E)$ and*

$$\int_E |fg| \leq \|f\|_p \|g\|_q.$$

The endpoint inequality $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$ also holds.

Proof. The endpoint case is immediate from $|g| \leq \|g\|_\infty$ almost everywhere. For $1 < p < \infty$, if either norm is zero, then one function is zero almost everywhere and the result is trivial. Otherwise normalize $F = |f|/\|f\|_p$ and $G = |g|/\|g\|_q$. Young's inequality gives $FG \leq F^p/p + G^q/q$. Integrating,

$$\int FG \leq \frac{1}{p} \int F^p + \frac{1}{q} \int G^q = \frac{1}{p} + \frac{1}{q} = 1.$$

Multiplying by $\|f\|_p \|g\|_q$ proves Holder. \square

Theorem 4.7 (Minkowski inequality). *For $1 \leq p \leq \infty$,*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof. The cases $p = 1$ and $p = \infty$ follow from the pointwise triangle inequality. For $1 < p < \infty$, let q be conjugate to p . If $f + g = 0$ almost everywhere, there is nothing to prove. Otherwise,

$$\|f + g\|_p^p = \int |f + g|^{p-1} |f + g| \leq \int |f| |f + g|^{p-1} + \int |g| |f + g|^{p-1}.$$

By Holder,

$$\int |f| |f + g|^{p-1} \leq \|f\|_p \left(\int |f + g|^{(p-1)q} \right)^{1/q} = \|f\|_p \|f + g\|_p^{p-1},$$

and similarly for g . Hence

$$\|f + g\|_p^p \leq (\|f\|_p + \|g\|_p) \|f + g\|_p^{p-1}.$$

Divide by $\|f + g\|_p^{p-1}$. \square

Proposition 4.8. *If $m(E) < \infty$ and $1 \leq p < q \leq \infty$, then $L^q(E) \subset L^p(E)$ and*

$$\|f\|_p \leq m(E)^{1/p-1/q} \|f\|_q,$$

with the usual interpretation when $q = \infty$.

Proof. For $q < \infty$, apply Holder to $|f|^p \cdot 1$ with exponents q/p and $q/(q-p)$:

$$\int_E |f|^p \leq \left(\int_E |f|^q \right)^{p/q} m(E)^{1-p/q}.$$

Taking the p -th root gives the inequality. If $q = \infty$, then $|f| \leq \|f\|_\infty$ almost everywhere, so $\|f\|_p \leq m(E)^{1/p} \|f\|_\infty$. \square

Proposition 4.9. *If $p > 1$, then every norm-bounded family in $L^p(E)$ is equi-integrable on finite-measure sets. More precisely, if $\|f\|_p \leq M$, then*

$$\int_A |f| \leq M m(A)^{1-1/p}.$$

Proof. Apply Holder to $|f| \mathbf{1}_A$:

$$\int_A |f| \leq \|f\|_p \|\mathbf{1}_A\|_q = M m(A)^{1/q} = M m(A)^{1-1/p}.$$

Given $\varepsilon > 0$, choose $\delta = (\varepsilon/M)^q$ if $M > 0$; the case $M = 0$ is trivial. \square

4.3 Completeness, Separability, and Convergence in L^p

Theorem 4.10 (Riesz–Fischer). *For $1 \leq p \leq \infty$, $L^p(E)$ is a Banach space.*

Proof. For $1 \leq p < \infty$, let f_n be Cauchy. Choose a subsequence f_{n_k} such that $\|f_{n_{k+1}} - f_{n_k}\|_p < 2^{-k}$. Let

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)|.$$

By Minkowski and monotone convergence,

$$\|g\|_p \leq \|f_{n_1}\|_p + \sum_{k=1}^{\infty} 2^{-k} < \infty.$$

Hence $g < \infty$ almost everywhere, so f_{n_k} converges pointwise almost everywhere to some f , and $|f_{n_k} - f|^p \leq (2g)^p$ for large enough dominating function locally obtained from the tail. More directly,

$$|f_{n_k} - f| \leq \sum_{j=k}^{\infty} |f_{n_{j+1}} - f_{n_j}|,$$

so by Minkowski,

$$\|f_{n_k} - f\|_p \leq \sum_{j=k}^{\infty} 2^{-j} \rightarrow 0.$$

Since f_n is Cauchy, the whole sequence converges to f in L^p .

For $p = \infty$, choose a representative of each f_n . Since f_n is Cauchy in essential sup norm, there is a null set outside which $f_n(x)$ is uniformly Cauchy in \mathbb{C} . Define $f(x) = \lim_n f_n(x)$ outside this null set and 0 on it. Then $\|f_n - f\|_{\infty} \rightarrow 0$. \square

Theorem 4.11. *For $1 \leq p < \infty$, $L^p(\mathbb{R}^d)$ is separable.*

Proof. Simple functions with finite-measure support are dense in L^p : approximate f by truncating its support and values, then use simple approximation and dominated convergence. Characteristic functions of finite-measure measurable sets can be approximated in L^p by characteristic functions of finite unions of rectangles, since $\|\mathbf{1}_A - \mathbf{1}_F\|_p = m(A \Delta F)^{1/p}$. Rectangles can be approximated by rectangles with rational endpoints, and coefficients can be approximated by rational numbers. Thus the countable set of step functions with rational coefficients and rational rectangles is dense. \square

Theorem 4.12. *Let $1 \leq p < \infty$, and suppose $f_n \rightarrow f$ almost everywhere with $f_n, f \in L^p(E)$. Then the following are equivalent:*

- (i) $\|f_n - f\|_p \rightarrow 0$;
- (ii) $\|f_n\|_p \rightarrow \|f\|_p$;
- (iii) $\{|f_n - f|^p\}$ is equi-integrable and tight.

Proof. (i) implies (ii) by the reverse triangle inequality:

$$\left| \|f_n\|_p - \|f\|_p \right| \leq \|f_n - f\|_p.$$

Assume (ii). Since $f_n \rightarrow f$ a.e., the Brezis–Lieb type identity in this elementary setting follows from the inequality

$$\left| |a|^p - |a - b|^p \right| \leq C_p(|b| |a|^{p-1} + |b|^p)$$

and dominated truncation; equivalently apply Fatou to $|f_n - f|^p$ and use convergence of norms to get $\|f_n - f\|_p \rightarrow 0$. Thus (ii) implies (i). Finally, (i) is equivalent to $|f_n - f|^p \rightarrow 0$ in L^1 . By the L^1 Vitali criterion, this is equivalent to equi-integrability and tightness of $\{|f_n - f|^p\}$, proving equivalence with (iii). \square

4.4 Duality and Weak Convergence

Definition 4.13. Let X be a normed space. A linear functional $T : X \rightarrow \mathbb{C}$ is *bounded* if there exists M such that $|T(x)| \leq M \|x\|$ for all $x \in X$. The dual space X^* is the space of bounded linear functionals with norm

$$\|T\| = \sup_{\|x\| \leq 1} |T(x)|.$$

Proposition 4.14. A linear functional is bounded if and only if it is continuous. Moreover,

$$\|T\| = \sup_{\|x\|=1} |T(x)|.$$

Proof. If T is bounded, then $|T(x) - T(y)| \leq \|T\| \|x - y\|$, so T is continuous. If T is continuous at 0, there is $\delta > 0$ such that $\|x\| < \delta$ implies $|T(x)| < 1$. For $x \neq 0$, applying this to $\delta x / (2 \|x\|)$ gives $|T(x)| \leq 2 \|x\| / \delta$, so T is bounded. The equality of the two suprema follows by scaling nonzero vectors to norm one. \square

Theorem 4.15 (Dual of L^p). Let $1 < p < \infty$ and let q be conjugate to p . For every $g \in L^q(E)$,

$$T_g(f) = \int_E fg$$

defines a bounded linear functional on $L^p(E)$, with $\|T_g\| = \|g\|_q$. Conversely, every bounded linear functional on $L^p(E)$ has this form for a unique $g \in L^q(E)$.

Proof. Holder's inequality gives $|T_g(f)| \leq \|f\|_p \|g\|_q$, hence $\|T_g\| \leq \|g\|_q$. If $g \neq 0$, take

$$f = \frac{\operatorname{sgn}(\bar{g}) |g|^{q-1}}{\|g\|_q^{q-1}},$$

so that $\|f\|_p = 1$ and $T_g(f) = \|g\|_q$. Thus $\|T_g\| = \|g\|_q$.

Conversely, first suppose E has finite measure. For a bounded functional T , define a set function $\nu(A) = T(\mathbf{1}_A)$. If A_j are disjoint, then $\mathbf{1}_{\cup_{j=1}^N A_j} = \sum_{j=1}^N \mathbf{1}_{A_j}$, and continuity of T plus $\|\mathbf{1}_{\cup_{j>N} A_j}\|_p = m(\cup_{j>N} A_j)^{1/p} \rightarrow 0$ gives countable additivity. Also $\nu \ll m$, since $m(A) = 0$

implies $\mathbf{1}_A = 0$ in L^p . By the Radon–Nikodym theorem, $\nu(A) = \int_A g$ for some $g \in L^1$. For simple functions s , $T(s) = \int sg$, and by density for all $f \in L^p$. The estimate

$$\sup_{\|f\|_p \leq 1} \left| \int fg \right| < \infty$$

implies $g \in L^q$ by testing against truncated versions of $\text{sgn}(\bar{g})|g|^{q-1}$. For σ -finite general E , apply the finite-measure argument on an increasing exhaustion E_n of finite measure and patch the resulting functions; uniqueness follows because $\int_A(g-h) = 0$ for all finite-measure A implies $g = h$ almost everywhere. \square

Definition 4.16. A sequence x_n in a normed space X converges *weakly* to x , written $x_n \rightharpoonup x$, if $T(x_n) \rightarrow T(x)$ for every $T \in X^*$. It converges *strongly* if $\|x_n - x\| \rightarrow 0$.

Proposition 4.17. *Strong convergence implies weak convergence. In $L^p(E)$, $1 < p < \infty$, one has $f_n \rightharpoonup f$ if and only if*

$$\int_E f_n g \rightarrow \int_E f g$$

for every $g \in L^q(E)$, where q is conjugate to p .

Proof. If $x_n \rightarrow x$ in norm and $T \in X^*$, then $|T(x_n) - T(x)| \leq \|T\| \|x_n - x\| \rightarrow 0$. The characterization in L^p is exactly the duality theorem: every functional is integration against some $g \in L^q$, and every $g \in L^q$ gives such a functional. \square

Proposition 4.18. *A weak limit in $L^p(E)$, $1 < p < \infty$, is unique.*

Proof. Suppose $f_n \rightharpoonup f$ and $f_n \rightharpoonup h$. Then $\int (f-h)g = 0$ for all $g \in L^q$. Taking $g = \text{sgn}(f-h)|f-h|^{p-1}$ on the finite-measure truncations where this is in L^q , and then letting the truncations increase, gives $\int |f-h|^p = 0$. Hence $f = h$ almost everywhere. \square

Theorem 4.19 (Weak convergence tests). *Let $1 < p < \infty$ and let q be conjugate to p .*

- (i) *If $\{f_n\}$ is bounded in $L^p(E)$, then $f_n \rightharpoonup f$ if and only if $\int_A f_n \rightarrow \int_A f$ for every measurable $A \subset E$ with finite measure.*
- (ii) *If $E = [a, b]$, then $f_n \rightharpoonup f$ in $L^p([a, b])$ if and only if $\int_a^b f_n(t)\phi(t) dt \rightarrow \int_a^b f(t)\phi(t) dt$ for every step function, equivalently for every continuous ϕ .*

Proof. For (i), weak convergence clearly implies convergence against $\mathbf{1}_A \in L^q$ whenever $m(A) < \infty$. Conversely, finite linear combinations of such characteristic functions are dense in $L^q(E)$. If $g \in L^q$, choose simple s with $\|g - s\|_q < \varepsilon$. Since $\{f_n\}$ is bounded in L^p , say $\|f_n\|_p \leq M$, Holder gives

$$\left| \int (f_n - f)(g - s) \right| \leq (M + \|f\|_p)\varepsilon.$$

The integral against s converges by assumption. Let $\varepsilon \downarrow 0$. Part (ii) is the same argument using the density of step functions, and the density of continuous functions, in $L^q([a, b])$. \square

Theorem 4.20 (Helly selection principle in a separable normed space). *Let X be separable and let $T_n \in X^*$ be a bounded sequence: $\sup_n \|T_n\| < \infty$. Then there is a subsequence T_{n_k} and a bounded linear functional $T \in X^*$ such that $T_{n_k}(x) \rightarrow T(x)$ for every $x \in X$.*

Proof. Let $\{x_j\}$ be a countable dense subset of X . Since $|T_n(x_1)| \leq M \|x_1\|$, the scalar sequence $T_n(x_1)$ has a convergent subsequence. From that subsequence choose a further subsequence on

which $T_n(x_2)$ converges, and continue. The diagonal subsequence T_{n_k} has the property that $T_{n_k}(x_j)$ converges for every j . For arbitrary $x \in X$, choose x_j close to x . Then

$$|T_{n_k}(x) - T_{n_\ell}(x)| \leq 2M \|x - x_j\| + |T_{n_k}(x_j) - T_{n_\ell}(x_j)|.$$

First choose j so the first term is small, then k, ℓ so the second term is small. Hence $T_{n_k}(x)$ converges for each x . Define $T(x) = \lim_k T_{n_k}(x)$. Linearity follows by passing to the limit, and $|T(x)| \leq M \|x\|$, so $T \in X^*$. \square

Corollary 4.21. *For $1 < p < \infty$, every bounded sequence in $L^p(E)$ has a weakly convergent subsequence, provided E is such that $L^q(E)$ is separable, for instance $E \subset \mathbb{R}^d$ with Lebesgue measure and σ -finite.*

Proof. A bounded sequence $f_n \in L^p$ defines bounded functionals T_n on L^q by $T_n(g) = \int f_n g$. Since L^q is separable, Helly's theorem gives a subsequence T_{n_k} converging pointwise to some bounded functional $T \in (L^q)^*$. By the duality theorem, $T(g) = \int f g$ for some $f \in L^p$. Therefore $\int f_{n_k} g \rightarrow \int f g$ for all $g \in L^q$, which is exactly weak convergence. \square

Chapter 5

Abstract Measure and Integration Spaces

5.1 Measure spaces and elementary consequences

Definition 5.1 (Sigma-algebra and measure). Let X be a set. A *sigma-algebra* on X is a non-empty collection $\mathcal{M} \subseteq \mathcal{P}(X)$ such that

- (i) if $E \in \mathcal{M}$, then $X \setminus E \in \mathcal{M}$;
- (ii) if $E_1, E_2, \dots \in \mathcal{M}$, then $\bigcup_{n=1}^{\infty} E_n \in \mathcal{M}$.

A *measure* on (X, \mathcal{M}) is a function $\mu : \mathcal{M} \rightarrow [0, \infty]$ such that $\mu(\emptyset) = 0$ and, for every disjoint sequence $E_1, E_2, \dots \in \mathcal{M}$,

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

The triple (X, \mathcal{M}, μ) is called a *measure space*. It is *finite* if $\mu(X) < \infty$, and *sigma-finite* if there are sets $X_n \in \mathcal{M}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mu(X_n) < \infty$ for all n .

Example 5.2 (Counting measure). If X is any set and $\mathcal{M} = \mathcal{P}(X)$, the counting measure is

$$\#(E) = \begin{cases} \text{the number of elements of } E, & E \text{ finite,} \\ \infty, & E \text{ infinite.} \end{cases}$$

If $X = \mathbb{N}$, integration with respect to counting measure is summation.

Example 5.3 (Weighted measure). Let (X, \mathcal{M}, μ) be a measure space and let $w : X \rightarrow [0, \infty]$ be measurable. Then

$$\nu(E) = \int_E w \, d\mu$$

defines another measure on (X, \mathcal{M}) . This construction will later be recognized as the model case of absolute continuity.

Proposition 5.4 (Elementary monotonicity and subadditivity). *Let (X, \mathcal{M}, μ) be a measure space.*

- (i) If $E, F \in \mathcal{M}$ and $E \subseteq F$, then $\mu(E) \leq \mu(F)$.

- (ii) If $E_1, E_2, \dots \in \mathcal{M}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mu(E_n).$$

(iii) If $E \subseteq F$ and $\mu(E) < \infty$, then $\mu(F \setminus E) = \mu(F) - \mu(E)$, with the convention that the right side is allowed to be ∞ .

Proof. For (i), write $F = E \sqcup (F \setminus E)$. Countable additivity gives

$$\mu(F) = \mu(E) + \mu(F \setminus E) \geq \mu(E).$$

For (ii), define disjoint sets

$$F_1 = E_1, \quad F_n = E_n \setminus \bigcup_{k=1}^{n-1} E_k \quad (n \geq 2).$$

Then $F_n \subseteq E_n$, the sets F_n are disjoint, and $\bigcup_n F_n = \bigcup_n E_n$. Hence by additivity and part (i),

$$\mu\left(\bigcup_n E_n\right) = \sum_n \mu(F_n) \leq \sum_n \mu(E_n).$$

For (iii), again use $F = E \sqcup (F \setminus E)$. Since $\mu(E) < \infty$, subtraction is legitimate in the extended real line and gives the desired identity. \square

Proposition 5.5 (Continuity of measure). *Let (X, \mathcal{M}, μ) be a measure space.*

(i) If $E_1 \subseteq E_2 \subseteq \dots$ and $E = \bigcup_{n=1}^{\infty} E_n$, then

$$\mu(E) = \lim_{n \rightarrow \infty} \mu(E_n).$$

(ii) If $E_1 \supseteq E_2 \supseteq \dots$, $E = \bigcap_{n=1}^{\infty} E_n$, and $\mu(E_1) < \infty$, then

$$\mu(E) = \lim_{n \rightarrow \infty} \mu(E_n).$$

Proof. For (i), set $F_1 = E_1$ and $F_n = E_n \setminus E_{n-1}$ for $n \geq 2$. Then $E = \bigsqcup_{n=1}^{\infty} F_n$, and for every N , $E_N = \bigsqcup_{n=1}^N F_n$. Therefore

$$\mu(E) = \sum_{n=1}^{\infty} \mu(F_n) = \lim_{N \rightarrow \infty} \sum_{n=1}^N \mu(F_n) = \lim_{N \rightarrow \infty} \mu(E_N).$$

For (ii), apply (i) to the increasing sequence $F_n = E_1 \setminus E_n$. Since $\bigcup_n F_n = E_1 \setminus E$, we get

$$\mu(E_1 \setminus E) = \lim_{n \rightarrow \infty} \mu(E_1 \setminus E_n).$$

Because $\mu(E_1) < \infty$, this is

$$\mu(E_1) - \mu(E) = \lim_{n \rightarrow \infty} (\mu(E_1) - \mu(E_n)),$$

and the conclusion follows. \square

Definition 5.6 (Null set and complete measure space). A set $N \in \mathcal{M}$ is *null* if $\mu(N) = 0$. The measure space (X, \mathcal{M}, μ) is *complete* if every subset of every null measurable set is measurable.

Proposition 5.7 (Completion of a measure space). *Let (X, \mathcal{M}, μ) be a measure space. Define*

$$\overline{\mathcal{M}} = \{A \subseteq X : \text{there exist } E, F \in \mathcal{M} \text{ with } E \subseteq A \subseteq F \text{ and } \mu(F \setminus E) = 0\}.$$

For $A \in \overline{\mathcal{M}}$, choose such an E and set $\overline{\mu}(A) = \mu(E)$. Then $\overline{\mathcal{M}}$ is a sigma-algebra, $\overline{\mu}$ is a complete measure on it, and $\overline{\mu}|_{\mathcal{M}} = \mu$.

Proof. First we prove that $\bar{\mu}$ is well-defined. Suppose

$$E_1 \subseteq A \subseteq F_1, \quad E_2 \subseteq A \subseteq F_2, \quad \mu(F_i \setminus E_i) = 0.$$

Since $E_1 \setminus E_2 \subseteq F_2 \setminus E_2$, we have $\mu(E_1 \setminus E_2) = 0$. Similarly $\mu(E_2 \setminus E_1) = 0$. Hence $E_1 \Delta E_2$ is null, and therefore $\mu(E_1) = \mu(E_2)$. Thus the definition of $\bar{\mu}(A)$ is independent of the choice of E .

We next show that $\bar{\mathcal{M}}$ is a sigma-algebra. Clearly $\emptyset \in \bar{\mathcal{M}}$. If $A \in \bar{\mathcal{M}}$, choose $E \subseteq A \subseteq F$ with $E, F \in \mathcal{M}$ and $\mu(F \setminus E) = 0$. Then

$$X \setminus F \subseteq X \setminus A \subseteq X \setminus E,$$

and $(X \setminus E) \setminus (X \setminus F) = F \setminus E$ is null. Hence $X \setminus A \in \bar{\mathcal{M}}$. If $A_n \in \bar{\mathcal{M}}$, choose $E_n, F_n \in \mathcal{M}$ with $E_n \subseteq A_n \subseteq F_n$ and $\mu(F_n \setminus E_n) = 0$. Then

$$\bigcup_n E_n \subseteq \bigcup_n A_n \subseteq \bigcup_n F_n,$$

and

$$\left(\bigcup_n F_n \right) \setminus \left(\bigcup_n E_n \right) \subseteq \bigcup_n (F_n \setminus E_n),$$

which has measure zero by countable subadditivity. Thus $\bigcup_n A_n \in \bar{\mathcal{M}}$.

To prove countable additivity, let $A_1, A_2, \dots \in \bar{\mathcal{M}}$ be disjoint. Choose $E_n, F_n \in \mathcal{M}$ as above. Then $\bar{\mu}(A_n) = \mu(E_n) = \mu(F_n)$, since $F_n \setminus E_n$ is null. The sets E_n are disjoint, and $\bigcup_n E_n \subseteq \bigcup_n A_n \subseteq \bigcup_n F_n$, with the two measurable bounds differing by a null set. Therefore

$$\bar{\mu} \left(\bigcup_n A_n \right) = \mu \left(\bigcup_n E_n \right) = \sum_n \mu(E_n) = \sum_n \bar{\mu}(A_n).$$

Finally, if $A \subseteq N \in \bar{\mathcal{M}}$ and $\bar{\mu}(N) = 0$, choose $E \subseteq N \subseteq F$ with $\mu(F \setminus E) = 0$. Then $\mu(F) = \mu(E) = \bar{\mu}(N) = 0$. Since $\emptyset \subseteq A \subseteq F$ and F is null, $A \in \bar{\mathcal{M}}$. Hence $\bar{\mu}$ is complete. \square

5.2 Outer measures and Caratheodory's theorem

Definition 5.8 (Outer measure). Let X be a set. An *outer measure* on X is a function $\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$ satisfying

- (i) $\mu^*(\emptyset) = 0$;
- (ii) if $A \subseteq B$, then $\mu^*(A) \leq \mu^*(B)$;
- (iii) for every sequence $A_1, A_2, \dots \subseteq X$,

$$\mu^* \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

Definition 5.9 (Caratheodory measurability). Let μ^* be an outer measure on X . A set $E \subseteq X$ is called μ^* -measurable if for every $A \subseteq X$,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E).$$

Since the inequality \leq always follows from subadditivity, the point is the reverse inequality.

Theorem 5.10 (Caratheodory theorem). *Let μ^* be an outer measure on X , and let \mathcal{M}_{μ^*} denote the collection of μ^* -measurable sets. Then \mathcal{M}_{μ^*} is a sigma-algebra, and $\mu = \mu^*|_{\mathcal{M}_{\mu^*}}$ is a complete measure.*

Proof. The empty set is measurable, and measurability is invariant under complements because the defining equality is symmetric in E and $X \setminus E$.

First we prove closure under finite unions. Let $E, F \in \mathcal{M}_{\mu^*}$. For any $A \subseteq X$, applying measurability of E and then of F gives

$$\begin{aligned}\mu^*(A) &= \mu^*(A \cap E) + \mu^*(A \setminus E) \\ &= \mu^*(A \cap E) + \mu^*((A \setminus E) \cap F) + \mu^*((A \setminus E) \setminus F).\end{aligned}$$

The first two terms cover $A \cap (E \cup F)$, so by subadditivity

$$\mu^*(A \cap (E \cup F)) \leq \mu^*(A \cap E) + \mu^*((A \setminus E) \cap F).$$

Therefore

$$\mu^*(A) \geq \mu^*(A \cap (E \cup F)) + \mu^*(A \setminus (E \cup F)).$$

The opposite inequality is automatic from subadditivity, hence $E \cup F$ is measurable. Thus \mathcal{M}_{μ^*} is an algebra.

Moreover, the same argument shows finite additivity in the following useful form: if E_1, \dots, E_N are disjoint measurable sets, then for every $A \subseteq X$,

$$\mu^*\left(A \cap \bigcup_{j=1}^N E_j\right) = \sum_{j=1}^N \mu^*(A \cap E_j).$$

This follows by induction from the defining equality.

Now let E_1, E_2, \dots be disjoint measurable sets and put $E = \bigcup_{n=1}^{\infty} E_n$, $S_N = \bigcup_{n=1}^N E_n$. Since S_N is measurable, for every $A \subseteq X$,

$$\mu^*(A) = \mu^*(A \cap S_N) + \mu^*(A \setminus S_N) \geq \sum_{n=1}^N \mu^*(A \cap E_n) + \mu^*(A \setminus E).$$

Letting $N \rightarrow \infty$,

$$\mu^*(A) \geq \sum_{n=1}^{\infty} \mu^*(A \cap E_n) + \mu^*(A \setminus E).$$

By subadditivity,

$$\mu^*(A \cap E) \leq \sum_{n=1}^{\infty} \mu^*(A \cap E_n),$$

so

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \setminus E).$$

Again the reverse inequality is automatic, hence E is measurable. Since an arbitrary countable union can be written as a disjoint countable union by replacing E_n with $E_n \setminus \bigcup_{k < n} E_k$, \mathcal{M}_{μ^*} is a sigma-algebra.

Taking $A = E$ in the previous finite and countable formulas gives countable additivity of μ^* on \mathcal{M}_{μ^*} . Finally, if $N \subseteq X$ and $\mu^*(N) = 0$, then for every $A \subseteq X$,

$$\mu^*(A \cap N) + \mu^*(A \setminus N) \leq 0 + \mu^*(A) = \mu^*(A),$$

while the opposite inequality is subadditivity. Hence every outer-null set is measurable, and the resulting measure is complete. \square

Definition 5.11 (Algebra and premeasure). An *algebra* on X is a non-empty collection $\mathcal{A} \subseteq \mathcal{P}(X)$ closed under complements and finite unions. A *premeasure* on \mathcal{A} is a function $\mu_0 : \mathcal{A} \rightarrow [0, \infty]$ such that $\mu_0(\emptyset) = 0$ and, whenever $A_1, A_2, \dots \in \mathcal{A}$ are disjoint and $\bigcup_n A_n \in \mathcal{A}$,

$$\mu_0 \left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} \mu_0(A_n).$$

Theorem 5.12 (Extension from a premeasure). *Let \mathcal{A} be an algebra on X , and let μ_0 be a premeasure on \mathcal{A} . For $E \subseteq X$, define*

$$\mu^*(E) = \inf \left\{ \sum_{n=1}^{\infty} \mu_0(A_n) : E \subseteq \bigcup_{n=1}^{\infty} A_n, A_n \in \mathcal{A} \right\}.$$

Then μ^ is an outer measure, every set in \mathcal{A} is μ^* -measurable, and $\mu^*|_{\mathcal{A}} = \mu_0$. Consequently μ^* restricts to a complete measure on the Caratheodory sigma-algebra containing $\sigma(\mathcal{A})$. If μ_0 is sigma-finite on \mathcal{A} , the extension to $\sigma(\mathcal{A})$ is unique.*

Proof. The empty set is covered by the empty set, so $\mu^*(\emptyset) = 0$. Monotonicity is immediate: a cover of B is also a cover of any $A \subseteq B$. For countable subadditivity, let $E \subseteq \bigcup_j E_j$. If one of $\mu^*(E_j)$ is infinite there is nothing to prove. Otherwise, for each j and each $\varepsilon > 0$, choose $A_{j,k} \in \mathcal{A}$ such that

$$E_j \subseteq \bigcup_{k=1}^{\infty} A_{j,k}, \quad \sum_{k=1}^{\infty} \mu_0(A_{j,k}) \leq \mu^*(E_j) + \frac{\varepsilon}{2^j}.$$

Then the countable family $A_{j,k}$ covers E , and hence

$$\mu^*(E) \leq \sum_{j,k} \mu_0(A_{j,k}) \leq \sum_j \mu^*(E_j) + \varepsilon.$$

Letting $\varepsilon \downarrow 0$ gives subadditivity.

We next show that μ^* extends μ_0 . For $A \in \mathcal{A}$, the cover A itself gives $\mu^*(A) \leq \mu_0(A)$. Conversely, suppose $A \subseteq \bigcup_n A_n$ with $A_n \in \mathcal{A}$. Define disjoint sets

$$B_1 = A \cap A_1, \quad B_n = A \cap A_n \setminus \bigcup_{k=1}^{n-1} A_k \quad (n \geq 2).$$

Then $B_n \in \mathcal{A}$, the B_n are disjoint, $\bigcup_n B_n = A$, and $B_n \subseteq A_n$. Since μ_0 is a premeasure and is monotone on \mathcal{A} ,

$$\mu_0(A) = \sum_n \mu_0(B_n) \leq \sum_n \mu_0(A_n).$$

Taking the infimum over all covers gives $\mu_0(A) \leq \mu^*(A)$.

Now fix $A \in \mathcal{A}$. We prove that A is Caratheodory measurable. Let $E \subseteq X$, and choose a cover $E \subseteq \bigcup_n B_n$ with $B_n \in \mathcal{A}$. Then $B_n \cap A$ covers $E \cap A$, and $B_n \setminus A$ covers $E \setminus A$. Hence

$$\begin{aligned} \mu^*(E \cap A) + \mu^*(E \setminus A) &\leq \sum_n \mu_0(B_n \cap A) + \sum_n \mu_0(B_n \setminus A) \\ &= \sum_n \mu_0(B_n). \end{aligned}$$

Taking the infimum over all such covers gives

$$\mu^*(E \cap A) + \mu^*(E \setminus A) \leq \mu^*(E).$$

The reverse inequality follows from subadditivity, so A is measurable.

It remains only to prove uniqueness under sigma-finiteness. Suppose ν is another measure on $\sigma(\mathcal{A})$ extending μ_0 . First assume $X \in \mathcal{A}$ and $\mu_0(X) < \infty$. Let

$$\mathcal{D} = \{E \in \sigma(\mathcal{A}) : \nu(E) = \mu^*(E)\}.$$

Then $\mathcal{A} \subseteq \mathcal{D}$. Also \mathcal{D} is a Dynkin class: it contains X ; if $E, F \in \mathcal{D}$ and $E \subseteq F$, then finite additivity inside the finite-measure space gives $F \setminus E \in \mathcal{D}$; and if $E_n \in \mathcal{D}$ are disjoint, then countable additivity gives $\bigcup_n E_n \in \mathcal{D}$. Since an algebra is a pi-system, Dynkin's pi-lambda theorem gives $\sigma(\mathcal{A}) \subseteq \mathcal{D}$.

In the sigma-finite case, choose $X_n \in \mathcal{A}$ such that $X_n \uparrow X$ and $\mu_0(X_n) < \infty$. For each n , the preceding paragraph applied to the finite algebra $\mathcal{A}_n = \{A \cap X_n : A \in \mathcal{A}\}$ shows that $\nu(E \cap X_n) = \mu^*(E \cap X_n)$ for every $E \in \sigma(\mathcal{A})$. Letting $n \rightarrow \infty$ and using continuity from below for both measures gives $\nu(E) = \mu^*(E)$. \square

5.3 Integration on an abstract measure space

Throughout this section (X, \mathcal{M}, μ) is a measure space.

Definition 5.13 (Measurable function). A function $f : X \rightarrow [-\infty, \infty]$ is *measurable* if $\{x : f(x) > a\} \in \mathcal{M}$ for every $a \in \mathbb{R}$. Complex-valued functions are measurable if their real and imaginary parts are measurable.

Proposition 5.14 (Stability of measurable functions). *Let $f, g : X \rightarrow \mathbb{C}$ be measurable and let $c \in \mathbb{C}$. Then $f + g$, cf , fg , $|f|$, $\max(f, g)$, and $\min(f, g)$ are measurable whenever the expressions are defined. If f_n is a sequence of measurable real-valued functions, then $\sup_n f_n$, $\inf_n f_n$, $\limsup_n f_n$, and $\liminf_n f_n$ are measurable.*

Proof. The inverse image of an open set under a measurable real-valued function is measurable; this follows because open subsets of \mathbb{R} are countable unions of intervals, and intervals can be described using inequalities $f > a$. Hence sums and products are measurable as compositions of measurable maps into \mathbb{R}^2 with continuous functions. The absolute value is treated similarly. For suprema,

$$\left\{x : \sup_n f_n(x) > a\right\} = \bigcup_{n=1}^{\infty} \{x : f_n(x) > a\},$$

which is measurable. Infima follow from $\inf_n f_n = -\sup_n(-f_n)$, and limsup/liminf are obtained from

$$\limsup_n f_n = \inf_N \sup_{n \geq N} f_n, \quad \liminf_n f_n = \sup_N \inf_{n \geq N} f_n.$$

\square

Definition 5.15 (Integral of a simple function). A non-negative simple function has the form

$$\varphi = \sum_{j=1}^N a_j \mathbf{1}_{E_j}, \quad a_j \geq 0, \quad E_j \in \mathcal{M}.$$

If the sets E_j are disjoint, define

$$\int_X \varphi d\mu = \sum_{j=1}^N a_j \mu(E_j).$$

This is independent of the chosen disjoint representation. For a non-negative measurable function f , define

$$\int_X f d\mu = \sup \left\{ \int_X \varphi d\mu : 0 \leq \varphi \leq f, \varphi \text{ simple} \right\}.$$

For a general measurable complex-valued f , write $f = f_1 - f_2 + i(f_3 - f_4)$, where each $f_j \geq 0$. We say $f \in L^1(X, \mu)$ if $\int |f| d\mu < \infty$, and then define $\int f d\mu$ by integrating the positive and negative parts.

Theorem 5.16 (Monotone convergence theorem). *Let $0 \leq f_1 \leq f_2 \leq \dots$ be measurable functions and let $f = \lim_{n \rightarrow \infty} f_n$. Then*

$$\int_X f d\mu = \lim_{n \rightarrow \infty} \int_X f_n d\mu.$$

Proof. Since $f_n \leq f$, monotonicity of the integral gives

$$\lim_n \int f_n d\mu \leq \int f d\mu.$$

For the reverse inequality, let $0 < a < 1$, and let φ be a simple function with $0 \leq \varphi \leq f$. Put

$$E_n = \{x : f_n(x) \geq a\varphi(x)\}.$$

Then $E_n \uparrow X$. Indeed, if $\varphi(x) = 0$, then $x \in E_n$ for all n ; if $\varphi(x) > 0$, then $f(x) \geq \varphi(x) > a\varphi(x)$, so for large n , $f_n(x) \geq a\varphi(x)$. On E_n we have $f_n \geq a\varphi$, hence

$$\int f_n d\mu \geq a \int_{E_n} \varphi d\mu.$$

Since φ is simple and $E_n \uparrow X$, continuity from below gives $\int_{E_n} \varphi d\mu \rightarrow \int \varphi d\mu$. Therefore

$$\lim_n \int f_n d\mu \geq a \int \varphi d\mu.$$

Letting $a \uparrow 1$, and then taking the supremum over all simple $\varphi \leq f$, gives the desired reverse inequality. \square

Corollary 5.17 (Fatou's lemma). *If $f_n \geq 0$ are measurable, then*

$$\int_X \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int_X f_n d\mu.$$

Proof. Set $g_N = \inf_{n \geq N} f_n$. Then $g_N \uparrow \liminf_n f_n$. By monotone convergence,

$$\int \liminf_n f_n d\mu = \lim_N \int g_N d\mu.$$

For every $n \geq N$, $g_N \leq f_n$, so $\int g_N d\mu \leq \inf_{n \geq N} \int f_n d\mu$. Letting $N \rightarrow \infty$ gives the result. \square

Corollary 5.18 (Dominated convergence theorem). *Suppose $f_n \rightarrow f$ almost everywhere, and $|f_n| \leq g$ almost everywhere for some $g \in L^1(X, \mu)$. Then $f \in L^1(X, \mu)$, $f_n \in L^1(X, \mu)$, and*

$$\lim_{n \rightarrow \infty} \int_X |f_n - f| d\mu = 0.$$

In particular, $\int f_n d\mu \rightarrow \int f d\mu$.

Proof. Since $|f| \leq g$ almost everywhere, $f \in L^1$. Apply Fatou's lemma to the non-negative functions $2g - |f_n - f|$. We get

$$\int 2g d\mu = \int \liminf_n (2g - |f_n - f|) d\mu \leq \liminf_n \int (2g - |f_n - f|) d\mu.$$

Thus

$$\limsup_n \int |f_n - f| d\mu \leq 0.$$

Since the integrals are non-negative, they converge to zero. The convergence of the integrals follows from

$$\left| \int f_n d\mu - \int f d\mu \right| \leq \int |f_n - f| d\mu.$$

□

Proposition 5.19 (Measures defined by densities). *Let $f : X \rightarrow [0, \infty]$ be measurable. Define*

$$\nu(E) = \int_E f d\mu, \quad E \in \mathcal{M}.$$

Then ν is a measure on (X, \mathcal{M}) . Moreover, if $g \geq 0$ is measurable, then

$$\int_X g d\nu = \int_X gf d\mu.$$

The same identity holds for all measurable g for which either side is absolutely integrable.

Proof. If E_n are disjoint measurable sets, then

$$\mathbf{1}_{\bigcup_n E_n} f = \sum_{n=1}^{\infty} \mathbf{1}_{E_n} f,$$

where the partial sums increase pointwise to the left-hand side. By monotone convergence,

$$\nu\left(\bigcup_n E_n\right) = \int \mathbf{1}_{\bigcup_n E_n} f d\mu = \sum_n \int \mathbf{1}_{E_n} f d\mu = \sum_n \nu(E_n).$$

Thus ν is a measure.

For $g = \mathbf{1}_E$, the identity is exactly the definition of ν . By linearity it holds for non-negative simple functions. If $g \geq 0$ is measurable, choose simple $g_n \uparrow g$. Applying monotone convergence to both measures gives

$$\int g d\nu = \lim_n \int g_n d\nu = \lim_n \int g_n f d\mu = \int gf d\mu.$$

For general absolutely integrable g , apply the non-negative case to positive and negative parts of real and imaginary parts. □

5.4 Product measures and Fubini–Tonelli

Definition 5.20 (Product sigma-algebra). Let (X, \mathcal{M}) and (Y, \mathcal{N}) be measurable spaces. The *product sigma-algebra* $\mathcal{M} \otimes \mathcal{N}$ is the sigma-algebra on $X \times Y$ generated by measurable rectangles $A \times B$, where $A \in \mathcal{M}$ and $B \in \mathcal{N}$.

Theorem 5.21 (Existence and uniqueness of product measure). *Let (X, \mathcal{M}, μ) and (Y, \mathcal{N}, ν) be sigma-finite measure spaces. There exists a unique measure $\mu \times \nu$ on $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ such that*

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$$

for all $A \in \mathcal{M}$, $B \in \mathcal{N}$.

Proof. Let \mathcal{A} be the algebra of finite disjoint unions of measurable rectangles. If

$$E = \bigsqcup_{j=1}^N (A_j \times B_j),$$

define

$$\rho_0(E) = \sum_{j=1}^N \mu(A_j)\nu(B_j).$$

This is well-defined: if a set in \mathcal{A} has two disjoint rectangle decompositions, refine both decompositions by all intersections

$$(A_i \cap A'_k) \times (B_i \cap B'_k).$$

Finite additivity of μ and ν shows that the two resulting sums are equal.

The standard monotone-class argument shows that ρ_0 is a premeasure on \mathcal{A} . The point is that countable additivity for a disjoint union whose union still lies in \mathcal{A} can be checked first on rectangles, then on finite unions of rectangles after refinement, and finally under sigma-finite exhaustion by finite-measure rectangles. Thus the extension theorem applies and gives a measure on $\sigma(\mathcal{A}) = \mathcal{M} \otimes \mathcal{N}$. This extension has the required value on rectangles. Uniqueness follows from the sigma-finite uniqueness part of the extension theorem, since finite-measure rectangles exhaust the product space and generate the product sigma-algebra. \square

Lemma 5.22 (Measurability of sections). *Let $E \in \mathcal{M} \otimes \mathcal{N}$. For $x \in X$ and $y \in Y$, define*

$$E_x = \{y \in Y : (x, y) \in E\}, \quad E^y = \{x \in X : (x, y) \in E\}.$$

Then $E_x \in \mathcal{N}$ for every x , and $E^y \in \mathcal{M}$ for every y .

Proof. Let \mathcal{C} be the collection of all $E \in \mathcal{M} \otimes \mathcal{N}$ whose vertical and horizontal sections are measurable. If $E = A \times B$, then $E_x = B$ if $x \in A$ and $E_x = \emptyset$ otherwise; similarly $E^y = A$ if $y \in B$ and \emptyset otherwise. Thus rectangles belong to \mathcal{C} . The class \mathcal{C} is closed under complements and countable unions because taking sections commutes with these operations. Hence \mathcal{C} is a sigma-algebra containing the rectangles, and therefore contains $\mathcal{M} \otimes \mathcal{N}$. \square

Theorem 5.23 (Tonelli theorem). *Let (X, \mathcal{M}, μ) and (Y, \mathcal{N}, ν) be sigma-finite measure spaces. If $f : X \times Y \rightarrow [0, \infty]$ is $\mathcal{M} \otimes \mathcal{N}$ -measurable, then the functions*

$$x \mapsto \int_Y f(x, y) d\nu(y), \quad y \mapsto \int_X f(x, y) d\mu(x)$$

are measurable, and

$$\int_{X \times Y} f d(\mu \times \nu) = \int_X \left(\int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y).$$

Proof. We prove the first iterated-integral identity; the second is symmetric. First assume $\mu(X) < \infty$ and $\nu(Y) < \infty$. Let \mathcal{C} be the collection of all sets $E \in \mathcal{M} \otimes \mathcal{N}$ such that $x \mapsto \nu(E_x)$ is measurable and

$$(\mu \times \nu)(E) = \int_X \nu(E_x) d\mu(x).$$

Rectangles belong to \mathcal{C} , because if $E = A \times B$, then $E_x = B$ for $x \in A$ and $E_x = \emptyset$ otherwise. Thus $\nu(E_x) = \nu(B)\mathbf{1}_A(x)$, and the identity is exactly $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$.

The class \mathcal{C} is a monotone class. If $E_n \uparrow E$ and each $E_n \in \mathcal{C}$, then $(E_n)_x \uparrow E_x$, so $\nu((E_n)_x) \uparrow \nu(E_x)$; monotone convergence gives both measurability of $x \mapsto \nu(E_x)$ and the integral

identity. If $E_n \downarrow E$, the finiteness of $\nu(Y)$ and $(\mu \times \nu)(X \times Y)$ allows us to use continuity from above and the same conclusion follows. By the monotone class theorem, $\mathcal{C} = \mathcal{M} \otimes \mathcal{N}$. Hence the theorem holds for indicators in the finite case, and therefore for non-negative simple functions by linearity. Passing from simple functions $f_n \uparrow f$ to f uses the monotone convergence theorem.

In the sigma-finite case, choose $X_k \uparrow X$ and $Y_k \uparrow Y$ with finite measure. Apply the finite result to $f \mathbf{1}_{X_k \times Y_k}$, and let $k \rightarrow \infty$. Monotone convergence gives the desired formula for f . \square

Corollary 5.24 (Fubini theorem). *Let $f \in L^1(X \times Y, \mu \times \nu)$. Then for almost every x , the function $y \mapsto f(x, y)$ is integrable on Y , and for almost every y , the function $x \mapsto f(x, y)$ is integrable on X . Moreover,*

$$\int_{X \times Y} f d(\mu \times \nu) = \int_X \left(\int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y),$$

where the iterated integrals are absolutely integrable.

Proof. Apply Tonelli's theorem to $|f|$. Since $\int |f| d(\mu \times \nu) < \infty$, the iterated integrals of $|f|$ are finite almost everywhere and integrable. Hence the sections of f are integrable almost everywhere. Applying Tonelli to f^+ , f^- , and the positive and negative parts of the real and imaginary parts proves the formula. \square

5.5 Absolute continuity and Radon–Nikodym theorem

Definition 5.25 (Absolute continuity and singularity). Let μ and ν be positive measures on (X, \mathcal{M}) . We say that ν is *absolutely continuous* with respect to μ , written $\nu \ll \mu$, if

$$\mu(E) = 0 \implies \nu(E) = 0.$$

We say that ν and μ are *mutually singular*, written $\nu \perp \mu$, if there exists $A \in \mathcal{M}$ such that $\mu(A) = 0$ and $\nu(X \setminus A) = 0$.

Proposition 5.26 (Epsilon-delta form in the finite case). *Let μ and ν be finite positive measures on (X, \mathcal{M}) . Then $\nu \ll \mu$ if and only if for every $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\mu(E) < \delta \implies \nu(E) < \varepsilon.$$

Proof. The epsilon-delta condition immediately implies absolute continuity by taking $\mu(E) = 0$. Conversely, suppose the condition fails. Then there exist $\varepsilon_0 > 0$ and sets $E_n \in \mathcal{M}$ such that $\mu(E_n) < 2^{-n}$ but $\nu(E_n) \geq \varepsilon_0$. Put

$$F_N = \bigcup_{n \geq N} E_n, \quad F = \bigcap_{N=1}^{\infty} F_N.$$

Then $\mu(F_N) \leq \sum_{n \geq N} 2^{-n} \rightarrow 0$, so $\mu(F) = 0$. Since $F_N \supseteq E_n$ for all $n \geq N$, $\nu(F_N) \geq \varepsilon_0$. The sets F_N decrease, and $\nu(F_1) < \infty$, so continuity from above gives

$$\nu(F) = \lim_{N \rightarrow \infty} \nu(F_N) \geq \varepsilon_0.$$

This contradicts $\nu \ll \mu$. Hence the epsilon-delta condition holds. \square

Fact 5.27 (Hilbert-space input). We shall use the Riesz representation theorem for Hilbert spaces: if H is a Hilbert space and $L : H \rightarrow \mathbb{C}$ is a bounded linear functional, then there exists a unique $h \in H$ such that $L(u) = \langle u, h \rangle$ for all $u \in H$.

Theorem 5.28 (Radon–Nikodym theorem, finite case). *Let μ and ν be finite positive measures on (X, \mathcal{M}) , and suppose $\nu \ll \mu$. Then there exists a non-negative measurable function f such that*

$$\nu(E) = \int_E f d\mu$$

for every $E \in \mathcal{M}$. The function f is unique up to equality μ -almost everywhere. It is denoted by

$$f = \frac{d\nu}{d\mu}.$$

Proof. Set $\rho = \mu + \nu$. Then ρ is a finite positive measure. Define a linear functional on $L^2(X, \rho)$ by

$$L(\varphi) = \int_X \varphi d\nu.$$

For $\varphi \in L^2(X, \rho)$, Cauchy's inequality gives

$$|L(\varphi)| \leq \int_X |\varphi| d\nu \leq \nu(X)^{1/2} \left(\int_X |\varphi|^2 d\nu \right)^{1/2} \leq \nu(X)^{1/2} \|\varphi\|_{L^2(\rho)}.$$

Thus L is bounded. By the Hilbert-space Riesz theorem, there is $h \in L^2(X, \rho)$ such that

$$\int_X \varphi d\nu = \int_X \varphi h d\rho$$

for every $\varphi \in L^2(X, \rho)$. Taking $\varphi = \mathbf{1}_E$, which belongs to $L^2(\rho)$ because ρ is finite, gives

$$\nu(E) = \int_E h d\rho.$$

We claim that $0 \leq h \leq 1$ ρ -almost everywhere. If $A = \{h < 0\}$, then

$$0 \leq \nu(A) = \int_A h d\rho \leq 0,$$

and the inequality is strict unless $\rho(A) = 0$. Hence $h \geq 0$ almost everywhere. Similarly, since

$$\mu(E) = \rho(E) - \nu(E) = \int_E (1 - h) d\rho,$$

we get $h \leq 1$ ρ -almost everywhere.

Let $B = \{h = 1\}$. Then

$$\mu(B) = \int_B (1 - h) d\rho = 0.$$

Since $\nu \ll \mu$, $\nu(B) = 0$, and therefore $\rho(B) = \mu(B) + \nu(B) = 0$. Thus $h < 1$ almost everywhere with respect to ρ . Define

$$f = \frac{h}{1 - h}$$

on the set where $h < 1$, and define $f = 0$ on the remaining ρ -null set. Then $f \geq 0$ is measurable. For every $E \in \mathcal{M}$, using $\mu(E) = \int_E (1 - h) d\rho$, the density rule gives

$$\int_E f d\mu = \int_E \frac{h}{1 - h} (1 - h) d\rho = \int_E h d\rho = \nu(E).$$

This proves existence.

For uniqueness, suppose $f, g \geq 0$ and $\int_E f d\mu = \int_E g d\mu$ for all $E \in \mathcal{M}$. For each $n \in \mathbb{N}$, let $A_n = \{f \geq g + 1/n\}$. Then

$$0 = \int_{A_n} (f - g) d\mu \geq \frac{1}{n} \mu(A_n),$$

so $\mu(A_n) = 0$. Since $\{f > g\} = \bigcup_n A_n$, we get $f \leq g$ almost everywhere. Reversing the roles of f and g gives $g \leq f$ almost everywhere. \square

Theorem 5.29 (Radon–Nikodym theorem, sigma-finite case). *Let μ and ν be sigma-finite positive measures on (X, \mathcal{M}) . If $\nu \ll \mu$, then there exists a non-negative measurable function f such that*

$$\nu(E) = \int_E f d\mu$$

for every $E \in \mathcal{M}$. The function f is unique up to μ -almost everywhere equality.

Proof. Since both measures are sigma-finite, choose measurable sets A_n and B_n such that $X = \bigcup_n A_n = \bigcup_n B_n$, $\mu(A_n) < \infty$, and $\nu(B_n) < \infty$. Replacing these by finite intersections and then disjointizing, we may assume that

$$X = \bigsqcup_{n=1}^{\infty} X_n, \quad \mu(X_n) < \infty, \quad \nu(X_n) < \infty.$$

Apply the finite Radon–Nikodym theorem to the restricted measures $\mu_n(E) = \mu(E \cap X_n)$ and $\nu_n(E) = \nu(E \cap X_n)$. We obtain $f_n \geq 0$ on X_n such that

$$\nu(E \cap X_n) = \int_{E \cap X_n} f_n d\mu.$$

Define $f(x) = f_n(x)$ for $x \in X_n$. Then

$$\nu(E) = \sum_n \nu(E \cap X_n) = \sum_n \int_{E \cap X_n} f_n d\mu = \int_E f d\mu.$$

Uniqueness follows by applying the finite uniqueness argument to each X_n . \square

Proposition 5.30 (Chain rule for Radon–Nikodym derivatives). *Suppose $\lambda \ll \nu \ll \mu$ are sigma-finite positive measures. Then*

$$\frac{d\lambda}{d\mu} = \frac{d\lambda}{d\nu} \frac{d\nu}{d\mu} \quad \mu\text{-almost everywhere.}$$

Proof. Let $f = d\lambda/d\nu$ and $g = d\nu/d\mu$. For every measurable E , the density rule gives

$$\lambda(E) = \int_E f d\nu = \int_E fg d\mu.$$

By uniqueness in the Radon–Nikodym theorem, $d\lambda/d\mu = fg$ μ -almost everywhere. \square

Theorem 5.31 (Lebesgue decomposition). *Let μ and ν be sigma-finite positive measures on (X, \mathcal{M}) . Then there exist unique positive measures ν_a and ν_s such that*

$$\nu = \nu_a + \nu_s, \quad \nu_a \ll \mu, \quad \nu_s \perp \mu.$$

Moreover, $\nu_a(E) = \int_E f d\mu$ for a uniquely determined $f \geq 0$, namely $f = d\nu_a/d\mu$.

Proof. First assume that μ and ν are finite. Let $\rho = \mu + \nu$. By the Radon–Nikodym theorem, there exists $0 \leq h \leq 1$ such that

$$\nu(E) = \int_E h d\rho, \quad \mu(E) = \int_E (1-h) d\rho.$$

Let $S = \{h = 1\}$ and $A = X \setminus S$. Then $\mu(S) = 0$, so $\nu|_S \perp \mu$. On A , define

$$f = \frac{h}{1-h}.$$

For every E ,

$$\nu(E \cap A) = \int_{E \cap A} h \, d\rho = \int_{E \cap A} \frac{h}{1-h} \, d\mu = \int_E f \mathbf{1}_A \, d\mu.$$

Thus set $\nu_a(E) = \nu(E \cap A)$ and $\nu_s(E) = \nu(E \cap S)$. This gives the desired decomposition in the finite case.

For sigma-finite measures, decompose $X = \bigsqcup_n X_n$ so that both $\mu(X_n)$ and $\nu(X_n)$ are finite. Apply the finite result on each X_n , and sum the absolutely continuous and singular parts over n . The countable union of the singular supports is still μ -null, so the singular part remains singular.

For uniqueness, suppose

$$\nu = \nu_a + \nu_s = \tilde{\nu}_a + \tilde{\nu}_s$$

with $\nu_a, \tilde{\nu}_a \ll \mu$ and $\nu_s, \tilde{\nu}_s \perp \mu$. Choose $S, T \in \mathcal{M}$ such that $\mu(S) = \mu(T) = 0$, $\nu_s(X \setminus S) = 0$, and $\tilde{\nu}_s(X \setminus T) = 0$. Put $N = S \cup T$. Then $\mu(N) = 0$, so $\nu_a(N) = \tilde{\nu}_a(N) = 0$, while both singular measures are supported on N . For every E , on $E \setminus N$ the singular parts vanish, so $\nu_a(E \setminus N) = \tilde{\nu}_a(E \setminus N)$. On $E \cap N$, the absolutely continuous parts vanish, so $\nu_s(E \cap N) = \tilde{\nu}_s(E \cap N)$. Combining these equalities gives $\nu_a = \tilde{\nu}_a$ and $\nu_s = \tilde{\nu}_s$. \square

5.6 A short conceptual summary

The key point of abstract measure theory is that the Euclidean construction of measure has two separable parts. First, Caratheodory's theorem turns a primitive outer measure into a complete measure on a sigma-algebra. Second, once a measure space is obtained, the integral and the convergence theorems no longer depend on Euclidean geometry. Product measures and Fubini's theorem express how integration behaves under products. The Radon–Nikodym theorem explains precisely when one measure can be integrated against another by means of a density. Haar measure then shows that translation-invariant integration persists on locally compact groups, which is the measure-theoretic foundation of abstract harmonic analysis.

Appendix A

Haar Measure and Convolution on Locally Compact Groups

This section is independent of the previous one in spirit, but it uses the same abstract language. Lebesgue measure is characterized by translation invariance on \mathbb{R}^n . Haar measure is the corresponding object on a locally compact group.

Definition A.1 (Radon measure). Let X be a locally compact Hausdorff space. A positive Borel measure μ on X is a *Radon measure* if

- (i) $\mu(K) < \infty$ for every compact set $K \subseteq X$;
- (ii) for every Borel set E ,

$$\mu(E) = \inf \{ \mu(U) : E \subseteq U, U \text{ open} \};$$

- (iii) for every Borel set E ,

$$\mu(E) = \sup \{ \mu(K) : K \subseteq E, K \text{ compact} \}.$$

Definition A.2 (Left Haar measure). Let G be a locally compact Hausdorff group. A non-zero Radon measure μ on G is a *left Haar measure* if

$$\mu(gE) = \mu(E)$$

for every $g \in G$ and every Borel set $E \subseteq G$. A *right Haar measure* is defined similarly by requiring $\mu(Eg) = \mu(E)$.

Fact A.3 (Haar existence and uniqueness). Every locally compact Hausdorff group admits a left Haar measure. It is unique up to multiplication by a positive constant. Similarly, every locally compact Hausdorff group admits a right Haar measure, also unique up to a positive constant.

Remark A.4. The existence proof is one of the deeper foundational results of abstract harmonic analysis. It uses local compactness in an essential way and is not merely an application of Caratheodory extension from a simple algebra of sets. In these notes we record the theorem as a structural fact and prove the main formal consequences used in analysis.

Proposition A.5 (Elementary consequences of left invariance). *Let μ be a left Haar measure on G . Then:*

- (i) every non-empty open set has positive measure;
- (ii) if $K \subseteq G$ is compact and $U \subseteq G$ is non-empty and open, then finitely many left translates of U cover K ;

(iii) if G is compact, then $\mu(G) < \infty$, so μ can be normalized by requiring $\mu(G) = 1$.

Proof. Since μ is non-zero, there exists a Borel set E with $\mu(E) > 0$. By inner regularity, there is a compact set $K_0 \subseteq E$ such that $\mu(K_0) > 0$. Let U be a non-empty open set and choose $u \in U$. For each $x \in K_0$, put $g_x = xu^{-1}$. Then $x \in g_x U$, so $\{g_x U\}_{x \in K_0}$ is an open cover of K_0 . Compactness gives a finite subcover $K_0 \subseteq \bigcup_{j=1}^N g_j U$. If $\mu(U) = 0$, then by left invariance each $\mu(g_j U) = 0$, hence $\mu(K_0) = 0$, a contradiction. Therefore $\mu(U) > 0$.

The same compactness argument proves (ii): choose $u \in U$, and for each $x \in K$ the translate $xu^{-1}U$ contains x . A finite subcollection covers K .

For (iii), compactness and the Radon property give $\mu(G) < \infty$. Since G is non-empty and open in itself, part (i) gives $\mu(G) > 0$. Thus $\mu/\mu(G)$ is again left invariant and has total mass one. \square

Definition A.6 (Modular function). Let μ be a left Haar measure on G . For $a \in G$, define a measure μ_a by

$$\mu_a(E) = \mu(Ea).$$

The measure μ_a is again left Haar, hence by uniqueness there is a positive constant $\Delta(a)$ such that

$$\mu(Ea) = \Delta(a)\mu(E).$$

The function $\Delta : G \rightarrow (0, \infty)$ is called the *modular function*. If $\Delta \equiv 1$, the group is called *unimodular*.

Proposition A.7 (Properties of the modular function). *Let G be a locally compact group with left Haar measure μ . Then:*

- (i) $\Delta : G \rightarrow (0, \infty)$ is a continuous group homomorphism;
- (ii) for every non-negative Borel function f ,

$$\int_G f(xa) d\mu(x) = \Delta(a)^{-1} \int_G f(x) d\mu(x);$$

- (iii) every compact group, every abelian locally compact group, and every discrete group is unimodular.

Proof. For the homomorphism property, use the defining identity twice:

$$\mu(Eab) = \Delta(b)\mu(Ea) = \Delta(b)\Delta(a)\mu(E).$$

Thus $\Delta(ab) = \Delta(a)\Delta(b)$.

For (ii), first take $f = \mathbf{1}_E$. Then

$$\int_G \mathbf{1}_E(xa) d\mu(x) = \mu(Ea^{-1}) = \Delta(a^{-1})\mu(E) = \Delta(a)^{-1}\mu(E).$$

The identity extends from indicators to simple functions by linearity and then to non-negative Borel functions by monotone convergence.

To prove continuity of Δ , choose $0 \leq f \in C_c(G)$ with $\int f d\mu > 0$. By (ii),

$$\Delta(a)^{-1} \int_G f d\mu = \int_G f(xa) d\mu(x).$$

The right-hand side is continuous in a : near a fixed a_0 , the supports of $x \mapsto f(xa)$ lie in one compact set, and uniform continuity of f on a compact neighborhood permits passage under the integral. Hence Δ^{-1} , and therefore Δ , is continuous.

For (iii), if G is compact, then $0 < \mu(G) < \infty$, and

$$\mu(Ga) = \mu(G) = \Delta(a)\mu(G),$$

so $\Delta(a) = 1$. If G is abelian, then left and right translations are the same; hence left invariance already gives $\mu(Ea) = \mu(aE) = \mu(E)$. Thus $\Delta(a) = 1$. If G is discrete, counting measure is a left and right Haar measure, so again $\Delta(a) = 1$. \square

Remark A.8 (Convention check). Some authors define the modular function by $\mu(Ea) = \Delta(a)^{-1}\mu(E)$. This replaces the present Δ by Δ^{-1} . In these notes the convention is

$$\mu(Ea) = \Delta(a)\mu(E), \quad \int f(xa) d\mu(x) = \Delta(a)^{-1} \int f d\mu.$$

Proposition A.9 (Convolution on a locally compact group). *Let G be a locally compact group with left Haar measure μ . If $f, g \in C_c(G)$, define*

$$(f * g)(x) = \int_G f(y)g(y^{-1}x) d\mu(y).$$

*Then $f * g \in C_c(G)$. Moreover, convolution is associative.*

Proof. For fixed x , the integrand is continuous in y and supported in

$$\text{supp}(f) \cap x \text{supp}(g)^{-1},$$

which is compact. Hence the integral is finite. Standard continuity under the integral sign applies because when x varies in a compact neighborhood, the supports remain inside one compact set and the integrands vary uniformly. Thus $f * g$ is continuous. If $(f * g)(x) \neq 0$, then there is $y \in \text{supp}(f)$ with $y^{-1}x \in \text{supp}(g)$, so $x \in \text{supp}(f) \text{supp}(g)$, a compact set. Thus $f * g \in C_c(G)$.

For associativity, compute using Fubini. Since all functions have compact support, the relevant integrands are absolutely integrable. We have

$$\begin{aligned} ((f * g) * h)(x) &= \int_G (f * g)(z)h(z^{-1}x) d\mu(z) \\ &= \int_G \int_G f(y)g(y^{-1}z)h(z^{-1}x) d\mu(y) d\mu(z). \end{aligned}$$

For fixed y , substitute $z = yu$. Left invariance gives $d\mu(z) = d\mu(u)$. Thus

$$\begin{aligned} ((f * g) * h)(x) &= \int_G f(y) \int_G g(u)h(u^{-1}y^{-1}x) d\mu(u) d\mu(y) \\ &= \int_G f(y)(g * h)(y^{-1}x) d\mu(y) \\ &= (f * (g * h))(x). \end{aligned}$$

\square

Theorem A.10 (Young inequality on a locally compact group). *Let G be a locally compact group with left Haar measure μ . For $f, g \in C_c(G)$, define*

$$(f * g)(x) = \int_G f(y)g(y^{-1}x) d\mu(y).$$

Then, for $1 \leq p \leq \infty$,

$$\|f * g\|_{L^p(G)} \leq \|f\|_{L^1(G)} \|g\|_{L^p(G)}.$$

Consequently convolution extends uniquely to a bounded bilinear map

$$L^1(G) \times L^p(G) \longrightarrow L^p(G).$$

In particular, $L^1(G)$ is a Banach algebra under convolution.

Proof. For $p = \infty$, left invariance gives

$$|(f * g)(x)| \leq \int_G |f(y)| |g(y^{-1}x)| d\mu(y) \leq \|g\|_\infty \|f\|_1.$$

For $1 \leq p < \infty$, Minkowski's integral inequality gives

$$\|f * g\|_p \leq \int_G |f(y)| \left(\int_G |g(y^{-1}x)|^p d\mu(x) \right)^{1/p} d\mu(y).$$

For fixed y , the change of variables $x = yu$ is a left translation, so $d\mu(x) = d\mu(u)$. Hence

$$\left(\int_G |g(y^{-1}x)|^p d\mu(x) \right)^{1/p} = \|g\|_p,$$

and the desired inequality follows. Since $C_c(G)$ is dense in $L^p(G)$ for Radon measures on locally compact spaces, the inequality permits a unique continuous extension from $C_c(G) \times C_c(G)$ to $L^1(G) \times L^p(G)$. The case $p = 1$ gives the Banach-algebra estimate. \square

Definition A.11 (Approximate identity on a locally compact group). A net $\{\phi_i\} \subset L^1(G)$ is called a *left approximate identity* if

- (i) $\phi_i \geq 0$ and $\int_G \phi_i d\mu = 1$;
- (ii) the supports of ϕ_i eventually lie in every neighborhood of the identity element e .

Proposition A.12 (Approximation by group kernels). *Let $\{\phi_i\}$ be a left approximate identity on G .*

- (i) *If $f \in C_c(G)$, then $\phi_i * f \rightarrow f$ uniformly on G .*
- (ii) *If $1 \leq p < \infty$ and $f \in L^p(G)$, then $\|\phi_i * f - f\|_p \rightarrow 0$.*

Proof. For $f \in C_c(G)$,

$$\phi_i * f(x) - f(x) = \int_G \phi_i(y)(f(y^{-1}x) - f(x)) d\mu(y).$$

A continuous compactly supported function on a locally compact group is uniformly continuous with respect to left translation on compact sets: for every $\varepsilon > 0$, there exists a neighborhood U of e such that

$$|f(y^{-1}x) - f(x)| < \varepsilon \quad \text{for all } x \in G, y \in U.$$

Once $\text{supp}(\phi_i) \subset U$, the displayed integral has absolute value at most $\varepsilon \int \phi_i d\mu = \varepsilon$, proving uniform convergence.

For L^p , first approximate f in L^p by a function $g \in C_c(G)$. Young's inequality gives

$$\|\phi_i * f - f\|_p \leq \|\phi_i * (f - g)\|_p + \|\phi_i * g - g\|_p + \|g - f\|_p \leq 2\|f - g\|_p + \|\phi_i * g - g\|_p.$$

The middle term tends to zero by the compactly supported case, and g may be chosen arbitrarily close to f . This proves the result. \square

Example A.13 (Basic Haar measures). (i) On $(\mathbb{R}^n, +)$, Lebesgue measure is Haar measure. The group is abelian, hence unimodular.

- (ii) On any discrete group G , counting measure is both left and right Haar measure.

- (iii) On a compact group G , Haar measure may be normalized to a probability measure. This is the setting in which averaging over the group is most directly analogous to averaging over a finite group.
- (iv) On $GL_n(\mathbb{R})$, a left and right Haar measure is

$$d\mu(A) = \frac{dA}{|\det A|^n},$$

where dA is Lebesgue measure on the vector space of real $n \times n$ matrices restricted to the open set $GL_n(\mathbb{R})$. Indeed, left multiplication $A \mapsto BA$ scales dA by $|\det B|^n$, while $|\det(BA)|^n = |\det B|^n |\det A|^n$; the factors cancel. The same argument applies to right multiplication.

Bibliography

- [1] Elias M. Stein and Rami Shakarchi. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton Lectures in Analysis III. Princeton University Press, 2005.
- [2] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Wiley, second edition, 1999.
- [3] Walter Rudin. *Real and Complex Analysis*. McGraw–Hill, third edition, 1987.