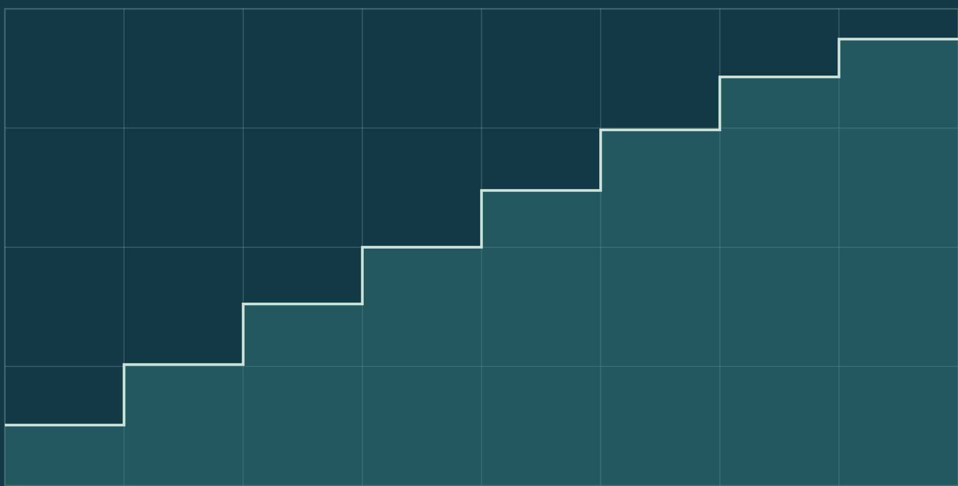
LECTURE NOTES 2026

Real Analysis

Measure, Integration and Function Spaces

Yi-Hsuan Lin



Department of Applied Mathematics
National Yang Ming Chiao Tung University

Yi-Hsuan Lin
Department of Applied Mathematics
National Yang Ming Chiao Tung University
Hsinchu, Taiwan

Real Analysis: Measure, Integration and Function Spaces
Lecture notes, 2026
Detailed revision R5.1

This revision develops the proofs in the author's handwritten course notes within the twelve-chapter organization of the earlier typeset version. The exterior measure is denoted by m_*. The principal references are listed at the end of the volume.

This is an independently prepared course manuscript, not a publication of any commercial publisher.

Preface

These notes develop Lebesgue measure and integration, differentiation of integrals, and the function spaces used in analysis. They grew out of a course on real analysis. The construction begins in Euclidean space, where cube coverings, compact approximation, and countable decompositions give explicit proofs of the basic properties of measure.

Convergence in measure and its relation to almost everywhere convergence are treated before integration. The integral is constructed in four stages: simple functions, bounded functions on finite-measure sets, nonnegative functions, and integrable functions. Repeated integration is proved through rectangles, open sets, null sets, and measurable sets, with the exceptional sections included in the argument.

The later chapters study distribution functions, rearrangements, L^p spaces, maximal functions, Fourier transforms, and weak derivatives. One-dimensional integration and integral estimates lead to the Sobolev inequalities. Hilbert space methods give Fourier series and boundary limits of holomorphic functions. The final chapter develops measure extension, Radon–Nikodym representation, and the characterization of absolutely continuous functions.

The presentation uses the author's handwritten notes as its main course reference. Intermediate constructions and limiting arguments are included in the proofs. Statements about L^p and Sobolev spaces concern equivalence classes unless a representative is specified.

The prerequisites are elementary real analysis, Euclidean topology, and linear algebra; the boundary-limit theorem also uses holomorphic power series. The main background texts are [1, 2], with complementary elementary material in [3, 4]. Exercises are separated from the main proofs.

Yi-Hsuan Lin

Contents

Preface ii

Notation vi

1 Lebesgue measure 1
1.1 Cubes and exterior measure . . . 1
1.2 Measurable sets and regularity . . . 5
1.3 Invariance and examples . . . 9
Exercises . . . 12

2 Measurable functions and convergence 14
2.1 Measurability and simple approximation . . . 14
2.2 Almost uniform convergence and continuity . . . 18
2.3 Convergence in measure . . . 20
Exercises . . . 22

3 The Lebesgue integral 24
3.1 Simple and bounded functions . . . 24
3.2 Nonnegative functions and monotone limits . . . 28
3.3 Integrable functions and dominated convergence . . . 31
3.4 Comparison with Riemann integration . . . 33
3.5 Integrals depending on a parameter . . . 35
Exercises . . . 35

4 The space L^1 and repeated integration 37
4.1 Completeness and approximation in L^1 . . . 37
4.2 Sections of sets and Fubini's theorem . . . 40
4.3 Tonelli's theorem and changes of variables . . . 43
4.4 Convolution . . . 45
Exercises . . . 46

5 Distribution functions and rearrangements **47**
5.1 Distribution functions and level sets 47
5.2 Integration over the range 48
5.3 Symmetric decreasing rearrangement 51
5.4 Weak integral bounds 53
Exercises 53

6 L^p spaces **55**
6.1 Norms and integral inequalities 55
6.2 Completeness, density, and endpoints 59
6.3 Interpolation and convolution estimates 61
6.4 Uniform integrability and convergence 62
Exercises 63

7 Maximal functions and differentiation **65**
7.1 A covering lemma and the maximal function 65
7.2 Strong bounds above the endpoint 67
7.3 Lebesgue points and differentiation of averages 68
7.4 The indefinite integral and absolute continuity 70
7.5 Why arbitrary rectangles are different 71
Exercises 72

8 Approximation and the Fourier transform **74**
8.1 Approximate identities 74
8.2 Mollifiers and smooth approximation 76
8.3 The Fourier transform on L^1 78
8.4 Gaussian kernels and inversion 78
8.5 An application to the heat equation 80
Exercises 82

9 Weak derivatives and Sobolev spaces **83**
9.1 Test functions and weak differentiation 83
9.2 The Sobolev norm and elementary properties 86
9.3 Local and global smooth approximation 87
9.4 Zero boundary values and the interval case 90
Exercises 92

10 Sobolev inequalities and compactness **93**
10.1 Scaling and the Sobolev inequality 93
10.2 Poincaré inequalities and ball averages 95
10.3 Morrey’s inequality and the critical case 97

10.4 Compactness from translations 99
Exercises 101

11 Hilbert spaces and Fourier series 103
11.1 Inner products and orthogonal projection 103
11.2 Orthonormal systems and expansions 105
11.3 Fourier series on the circle 108
11.4 Poisson integrals and radial limits 111
11.5 The Fourier transform on $L^2(\mathbb{R}^n)$ 112
11.6 Weak convergence and a Dirichlet problem 114
Exercises 116

12 Further measure theory 118
12.1 Measure spaces and abstract integration 118
12.2 Outer measures and extension 121
12.3 Absolute continuity and the Radon–Nikodym theorem 124
12.4 The dual of L^p 127
12.5 Stieltjes measures and differentiation on the line 128
12.6 Bounded variation and absolute continuity 133
Exercises 140

Bibliography 142

Notation

Throughout, n is a positive integer and $\mathbb{N} = \{1, 2, \ldots\}$. Lebesgue exterior measure is denoted by m_* and Lebesgue measure by m. We write m_n when the dimension needs to be indicated. For a measurable set E, its measure is written $m(E)$. For a rectangle R, $|R|$ denotes the product of its side lengths; Chapter 1 proves $m(R) = |R|$. The same volume notation is used for cubes. For scalars, vectors, and multiindices, vertical bars denote absolute value, Euclidean norm, and order, respectively.

The open ball with centre x and radius r is $B(x, r)$, and $\omega_n = m(B(0, 1))$. Complements are taken in the ambient space under discussion. The symmetric difference is $E \triangle F = (E \setminus F) \cup (F \setminus E)$. The indicator of E is $\mathbf{1}_E$. The notation $U \Subset \Omega$ means that $\overline{U}$ is compact and contained in Ω.

For $1 \leq p \leq \infty$, the conjugate exponent p' satisfies $1/p + 1/p' = 1$, with $1/\infty = 0$. Inner products are linear in the first argument. Thus the complex L^2 inner product is $\langle f, g \rangle = \int f\overline{g}$. The same conjugation convention is used in duality pairings.

The maximal function is written Mf. The notation f^* is reserved for symmetric decreasing rearrangement; the star in this notation is unrelated to the subscript in m_*. The Fourier transform is

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x \cdot \xi} \, dx.$$

On $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$, integration is normalized by $dt/(2\pi)$. A constant C may change from one estimate to the next, and its relevant dependence is indicated. A nonnegative integral may be infinite. We use $0 \cdot \infty = 0$ in simple integrals, but never assign a value to $\infty - \infty$.

For a measurable simple function, a finite measure support condition means that its nonzero level sets have finite measure. It does not mean that the topological closure of its nonzero set has finite measure. For a continuous function, its support has the usual topological meaning.

Chapter 1

Lebesgue measure

The construction of measure begins with rectangles. Their volumes are prescribed, whereas the measure of a general set must be obtained by approximation. We first establish the properties of exterior measure for arbitrary sets. Measurability is then imposed through approximation by open sets with a small excess.

1.1 Cubes and exterior measure

A closed rectangle is a set $R = \prod_{i=1}^{n}[a_i, b_i]$, where $a_i \leq b_i$, and its volume is $|R| = \prod_{i=1}^{n}(b_i - a_i)$. A cube has equal side lengths. We allow degenerate rectangles, whose volume is zero. Rectangles are called almost disjoint when their interiors are pairwise disjoint. Common boundary points are allowed.

Lemma 1.1. *If a rectangle R is covered by finitely many rectangles $R_1, \dots, R_N$, then $|R| \leq \sum_{j=1}^{N} |R_j|$. If their union is R and they are almost disjoint, equality holds.*

Proof. In each coordinate, list all endpoints of the intervals defining $R, R_1, \dots, R_N$. Consecutive endpoints determine a finite grid in a large rectangle containing these sets. The distributive law shows that the volume of any original rectangle is the sum of the volumes of the grid cells it contains. Cells lying in a coordinate face have a zero side length and contribute zero.

Consider a grid cell with positive volume whose interior lies in R. Choose a point in its interior. Some R_j contains that point. Since no endpoint of any R_j lies strictly between consecutive grid coordinates, the whole interior of the cell lies in that same rectangle. Thus every positive-volume cell of R is counted at least once in $\sum_j |R_j|$. This proves the inequality. For an almost disjoint decomposition, a positive-volume cell cannot be counted twice, since its interior would then lie in the interiors of two rectangles. Every such cell is therefore counted exactly once. □

Lemma 1.2. *Every open set $O \subset \mathbb{R}^n$ is a countable union of almost disjoint closed cubes. It is also a countable disjoint union of half open cubes.*

Proof. For $k = 0, 1, \ldots$, consider the half open dyadic cubes

$$Q_{k,j} = 2^{-k}(j + [0,1)^n), \quad j \in \mathbb{Z}^n.$$

At generation zero, accept every cube whose closure is contained in O. At generation $k \geq 1$, accept a cube if its closure is contained in O and it is not contained in a cube already accepted. There are only countably many candidates.

Two dyadic cubes from these generations are either disjoint or one contains the other. To verify this, compare their coordinate intervals: the endpoints of a coarser interval are endpoints in every finer grid. A finer interval meeting a coarser one is therefore contained in it, with the half open convention removing ambiguity at endpoints. Accepted cubes are consequently disjoint.

Fix $x \in O$. Choose $r > 0$ with $B(x, r) \subset O$. For sufficiently large k, the unique half open cube $Q_{k,j}$ containing x has diameter less than r, and hence $\overline{Q_{k,j}} \subset B(x, r) \subset O$. It is either accepted or contained in a previously accepted cube. Thus accepted half open cubes cover O. Their closures also cover O, remain inside O, and have disjoint interiors. This proves both assertions. □

Definition 1.3. For an arbitrary set $E \subset \mathbb{R}^n$, its Lebesgue exterior measure is

$$m_*(E) = \inf\Big\{\sum_{j=1}^{\infty} |Q_j| : E \subset \bigcup_{j=1}^{\infty} Q_j, \quad Q_j \text{ closed cubes}\Big\}. \tag{1.1}$$

The empty covering is permitted for $E = \varnothing$.

Exterior measure is defined for every set and takes values in $[0, \infty]$. A finite covering is included by appending zero-volume cubes, or by treating the sum as finite. Every set has a countable cube covering, since $\mathbb{R}^n$ does.

Proposition 1.4. *Exterior measure is monotone and countably subadditive:*

$$E \subset F \implies m_*(E) \leq m_*(F), \quad m_*\Big(\bigcup_j E_j\Big) \leq \sum_j m_*(E_j).$$

Every countable set has exterior measure zero.

Proof. A covering of F is a covering of E, so the infimum for E is no larger. For subadditivity, the assertion is automatic if $\sum_j m_*(E_j) = \infty$. Otherwise each term is finite. Fix $\varepsilon > 0$. For every j, choose a covering $(Q_{j,k})_{k\geq 1}$ of E_j with

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

The family indexed by $(j,k) \in \mathbb{N}^2$ is countable and covers $\bigcup_j E_j$. Hence its total volume is an upper bound for the exterior measure of that union. Summing the preceding inequalities gives an upper bound $\sum_j m_*(E_j) + \varepsilon$. Letting $\varepsilon \downarrow 0$ proves the claim.

A point is covered by a cube of arbitrarily small positive side length, so its exterior measure is zero. Countable subadditivity applies to a countable union of points. Equivalently, for an enumeration (x_j), cover x_j by a cube of volume less than $\varepsilon 2^{-j}$ and sum. □

Proposition 1.5. *The exterior measure of an open or closed rectangle, or a rectangle with any choice of open and closed faces, equals its volume.*

Proof. Let Q first be a closed cube. The covering by Q gives $m_*(Q) \leq |Q|$. Conversely, take an arbitrary countable covering $Q \subset \bigcup_j Q_j$. If $\sum_j |Q_j| = \infty$, then $|Q| \leq \sum_j |Q_j|$ is immediate. Otherwise fix $\varepsilon > 0$ and enlarge each Q_j slightly to an open cube U_j satisfying $|U_j| < |Q_j| + \varepsilon 2^{-j}$. Such an enlargement is possible by continuity of the volume as a function of the side length, including when the original cube is a point.

The open cubes cover the compact set Q, so finitely many suffice. Their closures still cover Q and have the same volumes. Lemma 1.1 yields

$$|Q| \leq \sum_{j\in J} |U_j| \leq \sum_{j=1}^{\infty} |Q_j| + \varepsilon.$$

Take the infimum over coverings and then let $\varepsilon \downarrow 0$. This gives $|Q| \leq m_*(Q)$.

For a closed rectangle R, the same compactness argument, using cube coverings of R, gives $|R| \leq m_*(R)$. To prove the reverse bound, cover R by all closed grid cubes of side length $1/k$ meeting it. There are at most $\prod_{i=1}^n (k(b_i - a_i) + 3)$ such cubes. Consequently

$$m_*(R) \leq \prod_{i=1}^{n} \left(b_i - a_i + \frac{3}{k}\right).$$

Let $k \to \infty$. This argument also covers degenerate rectangles, since the product on the right then tends to zero.

An open nondegenerate rectangle contains closed rectangles with volumes tending to its own and is contained in its closure. Monotonicity proves the assertion for open rectangles. Finally, the boundary of a rectangle is a finite union of degenerate rectangles and is therefore null by subadditivity. If two sets differ by a null set, subadditivity in both directions gives equal exterior measures. Adding or removing faces therefore does not change the answer. □

Remark 1.6. Countability in (1.1) is essential. A finite union of closed intervals containing $\mathbb{Q} \cap [0,1]$ is closed and hence contains its closure $[0,1]$. The sum of the interval lengths is then at least one. Nevertheless $m_*(\mathbb{Q} \cap [0,1]) = 0$. Restricting to finite coverings produces outer Jordan content, not Lebesgue exterior measure.

Proposition 1.7. *For every $E \subset \mathbb{R}^n$,*

$$m_*(E) = \inf\{m_*(O) : E \subset O,\ O \text{ open}\}.$$

Proof. Monotonicity gives $m_*(E) \leq m_*(O)$ for every admissible O. If $m_*(E) = \infty$, this already proves equality. If $m_*(E) < \infty$, fix $\varepsilon > 0$ and choose cubes covering E with total volume less than $m_*(E) + \varepsilon/2$. Enlarge the jth cube to an open cube with added volume less than $\varepsilon 2^{-j-1}$. The union O is open, contains E, and satisfies

$$m_*(O) \leq \sum_j |U_j| < m_*(E) + \varepsilon.$$

Since ε is arbitrary, the infimum is no larger than $m_*(E)$. □

Lemma 1.8. *If* $\operatorname{dist}(E,F) > 0$, *then* $m_*(E \cup F) = m_*(E) + m_*(F)$.

Proof. Subadditivity supplies the upper bound. Put $\delta = \operatorname{dist}(E,F) > 0$ and take a cube covering of $E \cup F$. Subdivide each cube into finitely many congruent closed cubes with diameter less than δ. The total volume is unchanged, and the resulting family remains countable. A member of this family cannot meet both E and F, because two points in the same cube have distance less than δ.

The cubes meeting E cover E, and the cubes meeting F cover F. These two index families are disjoint. Thus the total volume of the original covering is at least $m_*(E) + m_*(F)$. This uses only addition of nonnegative extended numbers, so it also applies when an exterior measure is infinite. Taking the infimum proves the lower bound. □

Proposition 1.9. *If (Q_j) are almost disjoint closed cubes, then $m_*(\bigcup_j Q_j) = \sum_j |Q_j|$.*

Proof. The upper bound follows from subadditivity. Ignore degenerate cubes, whose union is null. Fix N and shrink each of $Q_1, \dots, Q_N$ concentrically by a factor $1 - \delta$, where $0 < \delta < 1$. The resulting compact cubes Q'_j lie in the respective interiors and are disjoint. Each pair has positive distance: the continuous distance function attains its minimum on the product of two compact sets, and a zero minimum would give an intersection. Since the family is finite, its minimum pairwise distance is positive.

Apply Lemma 1.8 repeatedly. Monotonicity gives

$$m_*\Big(\bigcup_j Q_j\Big) \geq \sum_{j=1}^{N} |Q'_j| = (1-\delta)^n \sum_{j=1}^{N} |Q_j|.$$

First let $\delta \downarrow 0$, with N fixed. Then let $N \to \infty$. No continuity theorem for measure is being used here: only a limit of finite numerical sums followed by the definition of a nonnegative series is needed. □

1.2 Measurable sets and regularity

Definition 1.10. A set $E \subset \mathbb{R}^n$ is Lebesgue measurable if for every $\varepsilon > 0$ there is an open set $O \supset E$ with $m_*(O \setminus E) < \varepsilon$. For a measurable set, define $m(E) = m_*(E)$.

The excess set $O \setminus E$ is part of the definition. Merely finding open $O \supset E$ with $m_*(O)$ close to $m_*(E)$ is not sufficient: Proposition 1.7 gives that approximation for every set.

A collection $\mathcal{M}$ of subsets of $\mathbb{R}^n$ is a σ-algebra if it contains $\mathbb{R}^n$, is closed under complements, and is closed under countable unions. These properties also imply closure under countable intersections, since $\bigcap_j E_j = (\bigcup_j E_j^c)^c$. Finite intersections and differences then belong to the same collection.

We will use the following elementary consequence of compactness. If K is a nonempty compact set, F is a nonempty closed set, and $K \cap F = \varnothing$, then $\operatorname{dist}(K, F) > 0$. To see this, the function $x \mapsto \operatorname{dist}(x, F)$ is continuous: the triangle inequality gives $|\operatorname{dist}(x,F) - \operatorname{dist}(y,F)| \leq |x - y|$. It is positive at every $x \in K$, because F is closed. Its minimum on K is attained and is therefore positive. When either set is empty, any separation estimate involving their union is immediate.

Theorem 1.11. *The measurable sets form a σ-algebra containing the open sets, the closed sets, and every subset of a set of exterior measure zero.*

Proof. An open set satisfies the definition with $O = E$. If $m_*(E) = 0$, Proposition 1.7 gives open $O \supset E$ with $m_*(O) < \varepsilon$. Then $m_*(O \setminus E) < \varepsilon$, so E is measurable. A subset of E also has exterior measure zero by monotonicity and is measurable by the same argument.

We next prove closure under countable unions. Let E_j be measurable and fix $\varepsilon > 0$. Choose open $O_j \supset E_j$ with $m_*(O_j \setminus E_j) < \varepsilon 2^{-j}$. Set $E = \bigcup_j E_j$ and $O = \bigcup_j O_j$. If $x \in O \setminus E$, then $x \in O_j$ for some j and $x \notin E_j$. Hence

$$O \setminus E \subset \bigcup_j (O_j \setminus E_j), \quad m_*(O \setminus E) < \varepsilon.$$

Thus E is measurable.

Before taking complements, we prove measurability of compact sets. Let K be compact. It is bounded, so $m_*(K) < \infty$. Fix $\varepsilon > 0$ and choose open $O \supset K$ with $m_*(O) < m_*(K) + \varepsilon/2$. By Lemma 1.2, write $O \setminus K = \bigcup_j Q_j$ as an almost disjoint union of closed cubes. For fixed N, the compact set $\bigcup_{j=1}^N Q_j$ is disjoint from K, so its distance from K is positive. Lemma 1.8 and Proposition 1.9 imply

$$m_*(O) \geq m_*(K) + \sum_{j=1}^N |Q_j|.$$

Both $m_*(O)$ and $m_*(K)$ are finite. Subtracting $m_*(K)$ and then letting $N \to \infty$ gives

$$m_*(O \setminus K) = \sum_j |Q_j| \leq m_*(O) - m_*(K) < \varepsilon.$$

This proves that K is measurable. Every closed set is the union of its compact intersections with the closed balls $\overline{B}(0, j)$, so closure under countable unions proves measurability of every closed set.

Now let E be measurable. For each j, choose open $O_j \supset E$ with $m_*(O_j \setminus E) < 1/j$, and put $F = \bigcup_j O_j^c$. The set F is measurable, since every O_j^c is closed, and $F \subset E^c$. Moreover,

$$E^c \setminus F \subset O_j \setminus E \quad \text{for every } j.$$

Its exterior measure is therefore at most $1/j$ for every j, and is zero. Thus $E^c = F \cup (E^c \setminus F)$ is measurable. We have proved closure under complements and countable unions; closure under countable intersections follows from De Morgan's laws. □

Every rectangle is now measurable: a closed rectangle is closed, an open rectangle is open, and a rectangle with some faces removed is a difference of closed sets. Proposition 1.5 therefore gives $m(R) = m_*(R) = |R|$. From this point onwards, the measure of any such set may be used without confusing its measurability with its prescribed volume.

Proposition 1.12. *If E is measurable, then for every $\varepsilon > 0$ there is a closed $F \subset E$ with $m(E \setminus F) < \varepsilon$. If E is bounded, F can be chosen compact.*

Proof. Apply Definition 1.10 to E^c. Choose open $O \supset E^c$ with $m(O \setminus E^c) < \varepsilon$ and set $F = O^c$. Then F is closed, $F \subset E$, and $E \setminus F = O \setminus E^c$. If E is bounded, its closed subset F is compact by the Heine–Borel theorem. □

Theorem 1.13. *For pairwise disjoint measurable sets (E_j),*

$$m\Big(\bigcup_j E_j\Big) = \sum_j m(E_j).$$

Proof. The upper bound is subadditivity. First suppose every E_j is bounded; a common bound is not required. Fix $\varepsilon > 0$ and choose compact $K_j \subset E_j$ such that $m(E_j \setminus K_j) < \varepsilon 2^{-j}$. Subadditivity applied to $E_j = K_j \cup (E_j \setminus K_j)$ gives $m(K_j) > m(E_j) - \varepsilon 2^{-j}$. For each fixed N, the compact sets $K_1, \dots, K_N$ are disjoint, so the separated-set lemma gives

$$m\Big(\bigcup_j E_j\Big) \geq \sum_{j=1}^{N} m(K_j) \geq \sum_{j=1}^{N} m(E_j) - \varepsilon.$$

Let $N \to \infty$ and then $\varepsilon \downarrow 0$.

For arbitrary E_j, choose a disjoint measurable partition (S_k) of $\mathbb{R}^n$ by bounded shells, for example $S_1 = [-1, 1]^n$ and $S_k = [-k, k]^n \setminus [-k+1, k-1]^n$ for $k \geq 2$. The sets $E_j \cap S_k$ form a countable disjoint family of bounded measurable sets. The bounded case, applied first to this family and then to the shell decomposition of each E_j, yields

$$m\Big(\bigcup_j E_j\Big) = \sum_{j,k} m(E_j \cap S_k) = \sum_j m(E_j).$$

Regrouping is legitimate because all terms are nonnegative. □

For measurable sets, $E_j \uparrow E$ means that $E_j \subset E_{j+1}$ for every j and $E = \bigcup_j E_j$. Similarly, $E_j \downarrow E$ means that $E_{j+1} \subset E_j$ and $E = \bigcap_j E_j$. The arrows include both the nesting and the specified union or intersection.

Corollary 1.14. *If $E_j \uparrow E$, then $m(E_j) \to m(E)$. If $E_j \downarrow E$ and $m(E_{j_0}) < \infty$ for some j_0, then $m(E_j) \to m(E)$.*

Proof. For increasing sets, put $G_1 = E_1$ and $G_j = E_j \setminus E_{j-1}$ for $j \geq 2$. They are measurable and disjoint, $E = \bigcup_j G_j$, and $E_N = \bigcup_{j \leq N} G_j$. Countable additivity identifies $m(E_N)$ with the partial sums of the series $m(E) = \sum_j m(G_j)$.

For decreasing sets, discard the first $j_0 - 1$ terms. Thus assume $m(E_1) < \infty$. The sets $E_1 \setminus E_j$ increase to $E_1 \setminus E$. The increasing case gives convergence of their measures. Finite additivity within E_1 permits the identities

$$m(E_j) = m(E_1) - m(E_1 \setminus E_j), \quad m(E) = m(E_1) - m(E_1 \setminus E).$$

All quantities being subtracted are finite, so the asserted limit follows. □

The decreasing-set assertion is false without the finiteness hypothesis. The sets (j, ∞) decrease to the empty set but all have infinite measure.

Theorem 1.15. *For every measurable $E \subset \mathbb{R}^n$,*

$$m(E) = \inf_{O \supset E,\ O \text{ open}} m(O) = \sup_{K \subset E,\ K \text{ compact}} m(K).$$

If $m(E) < \infty$, then for every $\varepsilon > 0$ there is a finite union R of closed cubes with $m(E \triangle R) < \varepsilon$.

Proof. The open approximation follows from Proposition 1.7, since open sets are measurable. To obtain compact approximation when $m(E) < \infty$, choose closed $F \subset E$ with $m(E \setminus F) < \varepsilon/2$. The compact sets $K_j = F \cap \overline{B}(0, j)$ increase to F. Continuity from below gives $m(K_j) \to m(F)$, so some K_j satisfies $m(F \setminus K_j) < \varepsilon/2$. Consequently $m(E \setminus K_j) < \varepsilon$.

If $m(E) = \infty$, the finite-measure sets $E \cap \overline{B}(0, j)$ have measures increasing to infinity. Given any $A > 0$, first choose j with measure greater than $A + 1$, then choose a compact subset of that intersection losing less than one. This compact subset has measure greater than A. The supremum is therefore infinite.

For the last assertion, take a cube covering $E \subset G = \bigcup_j Q_j$ with $\sum_j |Q_j| < m(E) + \varepsilon/2$. The sum is finite. Choose N with $\sum_{j>N} |Q_j| < \varepsilon/2$ and put $R = \bigcup_{j \leq N} Q_j$. Then $E \setminus R \subset \bigcup_{j>N} Q_j$, so $m(E \setminus R) < \varepsilon/2$. Also $R \setminus E \subset G \setminus E$ and

$$m(G \setminus E) = m(G) - m(E) \leq \sum_j |Q_j| - m(E) < \varepsilon/2.$$

Here subtraction is permitted because $m(G) < \infty$. Adding the two errors proves the result. □

Definition 1.16. The Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ is the smallest σ-algebra containing the open sets. A G_δ set is a countable intersection of open sets, and an F_σ set is a countable union of closed sets.

The finite approximation in Theorem 1.15 can be chosen to use rectangles with rational endpoints. Indeed, for $m(E) < \infty$ and $\varepsilon > 0$, first choose $R = \bigcup_{j=1}^{N} Q_j$ with $m(E \triangle R) < \varepsilon/2$. Enlarge each closed cube Q_j slightly to a closed rectangle S_j with rational endpoints, so that $Q_j \subset S_j$ and $m(S_j) - m(Q_j) < \varepsilon/(2N)$. Density of the rationals and continuity of the product of the side lengths permit these choices, including for degenerate cubes. Set $S = \bigcup_{j=1}^{N} S_j$. Since $R \subset S$ and $S \setminus R \subset \bigcup_j (S_j \setminus Q_j)$, finite additivity and subadditivity give $m(R \triangle S) < \varepsilon/2$. The inclusion $E \triangle S \subset (E \triangle R) \cup (R \triangle S)$ now proves $m(E \triangle S) < \varepsilon$.

Proposition 1.17. *Every measurable set has representations $E = G \setminus N = F \cup N'$, where G is G_δ, F is F_σ, and N, N' are null. If A is an arbitrary set of finite exterior measure, it has a G_δ superset G with $m(G) = m_*(A)$.*

Proof. For measurable E, choose open $O_j \supset E$ with $m(O_j \setminus E) < 1/j$ and set $G = \bigcap_j O_j$. Then $E \subset G$ and $G \setminus E \subset O_j \setminus E$ for every j, so $N = G \setminus E$ is null. Apply this result to E^c and take complements to obtain the F_σ representation.

For arbitrary A of finite exterior measure, choose open $U_j \supset A$ with $m(U_j) < m_*(A) + 1/j$. The Borel set $G = \bigcap_j U_j$ contains A, and monotonicity gives $m_*(A) \le m(G) \le m_*(A) + 1/j$ for every j. Thus $m(G) = m_*(A)$. This last construction alone does not make A measurable. □

In particular, Lebesgue measurable sets are precisely Borel sets modified by subsets of Borel null sets. Lebesgue measure is the completion of its restriction to the Borel sets.

1.3 Invariance and examples

Proposition 1.18. *Translation, positive dilation, and coordinate reflection preserve measurability. For measurable E, $h \in \mathbb{R}^n$, and $a > 0$,*

$$m(E + h) = m(E), \quad m(aE) = a^n m(E).$$

Coordinate reflection preserves measure.

Proof. Each transformation takes cubes to cubes. Their volumes are respectively unchanged, multiplied by a^n, or unchanged. Transforming any covering gives one inequality for exterior measure; applying the inverse transformation gives the opposite inequality. Thus these formulas hold for exterior measure on arbitrary sets.

If open $O \supset E$ has a small excess, its image is open and contains the image of E. Since the maps are one-to-one, the image of $O \setminus E$ equals the excess of the two images. The exterior-measure formula therefore transfers Definition 1.10 to the image. In the dilation case choose the original error less than ε/a^n. Applying the inverse proves preservation in both directions. □

It follows that $m(B(x,r)) = \omega_n r^n$, where $0 < \omega_n < \infty$ by comparison with cubes. A sphere is null: for $0 < \delta < r$, it lies in $B(x, r+\delta) \setminus B(x, r-\delta)$, whose measure tends to zero as $\delta \downarrow 0$.

Proposition 1.19. *If $T : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz with constant L, then $m_*(T(E)) \le C_n L^n m_*(E)$. In particular, it maps null sets to null sets. A homeomorphism whose map and inverse are locally Lipschitz preserves Lebesgue measurability in both directions.*

Proof. If $L = 0$, the image has at most one point. Assume $L > 0$. A cube of side length s has diameter $\sqrt{n}\,s$. Its image has diameter at most $L\sqrt{n}\,s$ and is contained in a closed cube of side length $2L\sqrt{n}\,s$. Given a cube covering of E, these image-containing cubes cover $T(E)$ and have total volume at most $(2\sqrt{n})^n L^n$ times the original total. Infimizing proves the estimate.

For the last assertion, write a measurable set as a Borel set modified by a null set, using Proposition 1.17. A homeomorphism maps Borel sets to Borel sets, since its inverse is continuous. A locally Lipschitz map is Lipschitz on each member of a countable covering by sufficiently small neighbourhoods: choose such a neighbourhood at each point and then a countable subcover from a countable Euclidean base. The image of the null part is therefore a countable union of null sets. The image is measurable. Apply the same reasoning to the inverse. □

Example 1.20 (The Cantor set)**.** Let $C_0 = [0,1]$. At each stage remove the open middle third of every remaining interval, and let $C = \bigcap_{k\ge 0} C_k$. The set C_k consists of 2^k intervals of length 3^{-k}. Hence C is compact and $m(C) \le (2/3)^k$ for every k, so $m(C) = 0$.

The set is uncountable. For every binary sequence $(\varepsilon_j) \in \{0,1\}^{\mathbb{N}}$, the number $x = \sum_{j\ge 1} 2\varepsilon_j 3^{-j}$ belongs to C: at stage k, its first k digits specify one of the remaining intervals. If two sequences first differ at k, the difference of the leading terms is $2 \cdot 3^{-k}$, whereas the absolute value of the remaining difference is at most $\sum_{j>k} 2 \cdot 3^{-j} = 3^{-k}$. The two numbers are different. Binary sequences are uncountable by the diagonal argument, so C is uncountable as well.

Example 1.21 (A compact nowhere dense set of positive measure)**.** At stage $k \ge 1$, remove an open interval of length 4^{-k} from the middle of each of the

2^{k-1} surviving intervals. Let ℓ_k be their common length after stage k. The recurrence is $\ell_k = (\ell_{k-1} - 4^{-k})/2$, with $\ell_0 = 1$. Direct induction gives

$$\ell_k = 2^{-k-1} + 2^{-2k-1}.$$

In particular, $\ell_{k-1} > 4^{-k}$, so every prescribed removal is possible. The stage-k set has total length $2^k \ell_k = 1/2 + 2^{-k-1}$. Continuity from above gives measure $1/2$ for the limiting compact set K.

Since $\ell_k \to 0$, no nonempty open interval can be contained in K: an interval contained in K would be contained in one surviving component at every stage, which would force its length to be at most ℓ_k for all k. Thus K has empty interior. Being closed, it is nowhere dense.

Theorem 1.22. *There exists a subset of* $[0, 1]$ *that is not Lebesgue measurable.*

Proof. Declare $x \sim y$ when $x - y \in \mathbb{Q}$. Choose one representative in $[0, 1]$ from every equivalence class meeting $[0, 1]$, and let V be the set of representatives. This step uses the axiom of choice. The translates $V + q$, for $q \in \mathbb{Q} \cap [-1, 1]$, are disjoint. Indeed, $v + q = w + r$ implies $v - w = r - q \in \mathbb{Q}$, hence $v = w$ by the representative choice, and then $q = r$.

Their union U contains $[0, 1]$: every $x \in [0, 1]$ differs from its representative by a rational number in $[-1, 1]$. Also $U \subset [-1, 2]$. If V were measurable, translation invariance and countable additivity would give $m(U) = \sum_{q\in\mathbb{Q}\cap[-1,1]} m(V)$. If $m(V) = 0$, this contradicts $m(U) \geq 1$. If $m(V) > 0$, the series is infinite, contradicting $m(U) \leq 3$. These exhaust the possibilities, since $V \subset [0, 1]$ has finite exterior measure. ◻

Theorem 1.23 (Carathéodory criterion)**.** *A set* $E \subset \mathbb{R}^n$ *is measurable if and only if*

$$m_*(A) = m_*(A \cap E) + m_*(A \setminus E) \quad \textit{for every } A \subset \mathbb{R}^n. \tag{1.2}$$

Proof. Suppose E is measurable. If $m_*(A) < \infty$, choose a measurable hull $G \supset A$ with $m(G) = m_*(A)$ by Proposition 1.17. The two measurable pieces of G give

$$m_*(A) = m(G \cap E) + m(G \setminus E) \geq m_*(A \cap E) + m_*(A \setminus E).$$

Subadditivity gives the reverse inequality. If $m_*(A) = \infty$, subadditivity already forces the sum on the right to be infinite, so equality still holds.

Conversely, assume (1.2). If $m_*(E) < \infty$, choose a measurable hull $G \supset E$ of equal measure. Applying the splitting identity with $A = G$ gives $m(G) = m_*(E) + m_*(G \setminus E)$. Both $m(G)$ and $m_*(E)$ are finite and equal, so $G \setminus E$ is null. Thus E is measurable by completeness.

For general E, put $Q_j = [-j, j]^n$ and $H_j = E \cap Q_j$. We verify the splitting criterion for H_j without presupposing its measurability. Since Q_j is measurable, the direction already proved gives

$$\begin{aligned} m_*(A) &= m_*(A \cap E) + m_*(A \setminus E) \\ &= m_*(A \cap E \cap Q_j) + m_*((A \cap E) \setminus Q_j) + m_*(A \setminus E) \\ &\geq m_*(A \cap H_j) + m_*(A \setminus H_j). \end{aligned}$$

The last inequality is subadditivity on the two pieces outside H_j. The opposite inequality is always true by subadditivity. Hence H_j satisfies the criterion and has finite exterior measure, so the finite case makes it measurable. Finally $E = \bigcup_j H_j$ is measurable. □

Exercises

Exercise 1.1. Prove that a compact set and a disjoint closed set have positive distance. Give disjoint closed subsets of $\mathbb{R}^2$ with distance zero.

Exercise 1.2. Prove that the definition of m_* is unchanged when closed cubes are replaced by open cubes or by rectangles. Include degenerate rectangles.

Exercise 1.3. For bounded $E \subset \mathbb{R}$, define outer Jordan content using finite closed-interval coverings. Prove that $J_*(E) = J_*(\overline{E})$ and compute it for $\mathbb{Q} \cap [0, 1]$.

Exercise 1.4. For measurable E_j, prove $m(\liminf E_j) \leq \liminf m(E_j)$. Prove the reverse limsup inequality when $m(\bigcup_j E_j) < \infty$, and show why this assumption matters.

Exercise 1.5. Approximate a finite-measure set in symmetric difference by finite unions of rectangles with rational endpoints. Can the approximating sets always be contained in the original set?

Exercise 1.6. Construct a dense open subset of $\mathbb{R}$ with measure less than a prescribed positive number.

Exercise 1.7. Prove that the Cantor set has no isolated points. Construct a compact nowhere dense subset of $[0, 1]$ of any prescribed measure in $(0, 1)$.

Exercise 1.8. Prove that every positive-measure set contains a compact subset of positive measure. Explain why it need not contain a nonempty open set.

Exercise 1.9. Prove that every coordinate hyperplane is null, and use a Lipschitz linear map to deduce the result for every affine hyperplane.

Exercise 1.10. Show that a bi-Lipschitz homeomorphism preserves null sets in both directions. Construct an increasing homeomorphism of $[0,1]$ that does not preserve null sets.

Exercise 1.11. Let $G \supset A$ be measurable with $m(G) = m_*(A) < \infty$. Show that every measurable subset of $G \setminus A$ is null. Does this imply that A is measurable?

Exercise 1.12. For finite-measure sets, prove $|m(E) - m(F)| \leq m(E \triangle F)$. Show that symmetric-difference measure defines a metric on classes modulo null sets.

Exercise 1.13. Show that there are Lebesgue measurable subsets of the Cantor set that are not Borel. One approach compares the cardinality of Borel sets with that of all subsets of the Cantor set.

Exercise 1.14. If $m(E) > 0$, prove that $E - E$ contains an interval about zero in dimension one. First establish continuity of $h \mapsto m(K \cap (K + h))$ for compact K of positive measure.

Chapter 2

Measurable functions and convergence

We now pass from sets to functions. All measurability statements in this chapter refer to Lebesgue measure. Completeness is used when functions are modified on null sets. This point will need separate treatment when abstract measure spaces are introduced.

2.1 Measurability and simple approximation

Definition 2.1. Let $E \subset \mathbb{R}^n$ be measurable. An extended real valued function $f : E \to [-\infty, \infty]$ is measurable if $\{x \in E : f(x) > a\}$ is measurable for every $a \in \mathbb{R}$. A complex valued function is measurable when its real and imaginary parts are measurable.

All level sets are taken in the domain E. Complements below are consequently relative to E. The definition also gives measurability of $\{f \le a\}$, and

$$\{f < a\} = \bigcup_{k \ge 1} \{f \le a - 1/k\}, \quad \{f \ge a\} = E \setminus \{f < a\}.$$

It follows that $\{f = a\}$ is measurable. The infinite level sets are $\{f = \infty\} = \bigcap_k \{f > k\}$ and $\{f = -\infty\} = \bigcap_k \{f < -k\}$. Testing only rational a is enough, because

$$\{f > a\} = \bigcup_{r \in \mathbb{Q},\ r > a} \{f > r\}.$$

For the reverse inclusion implicit in this identity, if $f(x) > a$, choose a rational r strictly between a and $f(x)$.

Proposition 2.2. *A finite real valued function f on E is measurable if and only if the inverse image of every open subset of $\mathbb{R}$ is measurable. In this case the inverse image of every Borel set is measurable.*

Proof. The inverse image of an interval (a, b) is $\{f > a\} \cap \{f < b\}$. Every open subset of $\mathbb{R}$ is a countable union of open intervals, so measurable level sets give measurable inverse images of all open sets. Conversely, take the open set (a, ∞).

Let $\mathcal{A}$ be the collection of subsets $B \subset \mathbb{R}$ for which $f^{-1}(B)$ is measurable. Since inverse images commute with complements and countable unions, $\mathcal{A}$ is a σ-algebra. It contains all open sets and therefore contains the Borel σ-algebra. □

A continuous function on a measurable set E, with its relative topology, is measurable: the inverse image of an open set has the form $E \cap O$ with O open in $\mathbb{R}^n$. If f is measurable and Φ is Borel measurable on its range, then $\Phi \circ f$ is measurable. For several finite measurable functions $f_1, \ldots, f_k$, the map $x \mapsto (f_1(x), \ldots, f_k(x))$ has measurable inverse images of rectangles. Since open subsets of $\mathbb{R}^k$ are countable unions of rational rectangles, continuous functions of this vector are measurable. In particular, sums, products, absolute values, maxima, and minima are measurable. A quotient is measurable on the measurable set where its denominator is nonzero.

For extended values, expressions must first be defined. We do not form $\infty - \infty$. For a real function, set $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$. At every finite value, $f = f^+ - f^-$ and $|f| = f^+ + f^-$. At an infinite value, only one of these two nonnegative parts is infinite.

Definition 2.3. A property holds almost everywhere on E if it holds outside a measurable null subset of E. We abbreviate this by a.e.

Here and below, Borel measurability means that the inverse image of every open set is Borel in a Borel domain. It is a stronger requirement on the level sets than Lebesgue measurability. For real measurable f, g, a useful direct comparison is

$$\{f > g\} = \bigcup_{r \in \mathbb{Q}} \big(\{f > r\} \cap \{g < r\}\big).$$

For a point where $f > g$, choose a rational number strictly between the two values; this remains possible if one of the values is infinite. Conversely, membership in any set on the right gives $f > r > g$. Thus $\{f > g\}$ is measurable. In particular, for finite functions, $\{f + g > a\} = \{f > a - g\}$ is measurable once scalar multiplication and translation have been obtained from continuous composition. Products follow from $fg = ((f+g)^2 - f^2 - g^2)/2$, and division follows on the measurable set $\{g \neq 0\}$ by composing g with the continuous reciprocal map. These identities do not assign values to $\infty - \infty$.

Proposition 2.4. *If f is measurable and $g = f$ almost everywhere, then g is measurable.*

Proof. Choose a measurable null set N outside which the functions agree. For each threshold a, the symmetric difference between $\{f > a\}$ and $\{g > a\}$ is contained in N. Every subset of N is measurable by completeness, so $\{g > a\}$ is measurable. Apply this separately to real and imaginary parts for complex functions. □

In particular, a measurable function finite almost everywhere can be redefined to be zero on its infinite level sets. Countably many almost-everywhere assertions can be made simultaneous by removing the union of their exceptional null sets. The analogous statement for uncountably many assertions does not follow from countable additivity.

Proposition 2.5. *For measurable extended real functions f_j, the functions $\sup_j f_j$, $\inf_j f_j$, $\limsup_j f_j$, and $\liminf_j f_j$ are measurable. A pointwise limit, or an almost everywhere limit with arbitrary values assigned on the exceptional set, is measurable.*

Proof. The identity $\{\sup_j f_j > a\} = \bigcup_j \{f_j > a\}$ proves the assertion for the supremum. Since $\inf_j f_j = -\sup_j(-f_j)$, it also proves the assertion for the infimum. Now use

$$\limsup_j f_j = \inf_N \sup_{j \geq N} f_j, \quad \liminf_j f_j = \sup_N \inf_{j \geq N} f_j.$$

When a limit exists, it equals both functions. For an almost everywhere limit, first define it to equal a measurable limsup where appropriate, and then use Proposition 2.4 on the exceptional set. □

Definition 2.6. A simple function is a finite linear combination $\phi = \sum_{j=1}^{N} a_j \mathbf{1}_{E_j}$ of indicators of measurable sets. Its canonical representation uses the distinct nonzero values a_j and the pairwise disjoint sets $E_j = \{\phi = a_j\}$. A step function is a finite linear combination of indicators of rectangles.

A simple function need not be constant on intervals. Its level sets may have positive measure and empty interior. The zero level set is usually omitted from the canonical representation, but it must be included when a proof requires a partition of the whole domain.

Theorem 2.7. *Every nonnegative measurable function is an increasing pointwise limit of nonnegative simple functions. Every finite measurable function is a pointwise limit of simple functions. Bounded measurable functions admit uniform simple approximation.*

Proof. For $f \geq 0$, define, when $f(x) < \infty$,

$$\phi_k(x) = \min\{k, 2^{-k}\lfloor 2^k f(x)\rfloor\},$$

and set $\phi_k(x) = k$ when $f(x) = \infty$. The function takes values among $0, 2^{-k}, 2 \cdot 2^{-k}, \dots, k$. Every level set is obtained from measurable level sets of f, so ϕ_k is simple and measurable.

For $t \geq 0$, the inequality $\lfloor 2^{k+1}t\rfloor \geq 2\lfloor 2^k t\rfloor$ implies $2^{-k-1}\lfloor 2^{k+1}t\rfloor \geq 2^{-k}\lfloor 2^k t\rfloor$. Increasing the truncation level from k to $k+1$ preserves this inequality. Thus $0 \leq \phi_k \leq \phi_{k+1} \leq f$. If $f(x) < \infty$, then for $k > f(x)$ the truncation is inactive and $0 \leq f(x) - \phi_k(x) < 2^{-k}$. If $f(x) = \infty$, then $\phi_k(x) = k \to \infty$. This proves pointwise monotone convergence at every point.

For finite real valued f, apply the construction to f^+ and f^- and subtract. If $|f| \leq M$, then for $k > M$ neither part is truncated and the difference tends uniformly to zero. Finally, approximate the real and imaginary parts of a complex function and add them. The absolute error is bounded by the sum of the two real errors. □

Proposition 2.8. *Every Lebesgue measurable function on $\mathbb{R}^n$ agrees almost everywhere with a Borel measurable function. If the original function is finite almost everywhere, the Borel representative may be chosen finite everywhere.*

Proof. Assume first $f \geq 0$ and use the approximations ϕ_k from Theorem 2.7. Each of the finitely many level sets in ϕ_k differs from a Borel set by a null set. Replacing those indicators gives a Borel simple function ψ_k equal to ϕ_k outside a null set N_k. The replacing Borel sets need not be disjoint: the resulting finite sum is still Borel, and outside N_k it has the required values.

Outside $N = \bigcup_k N_k$, the sequence ψ_k equals ϕ_k at every index. Hence the Borel function $\limsup_k \psi_k$ agrees with f there. For a signed function, perform the construction for its two nonnegative parts. The possible undefined difference occurs on a Borel set where both reconstructed parts are infinite. That set is null whenever the original difference is defined, and assigning zero there gives a Borel function. Real and imaginary parts treat complex functions. If the original function is finite a.e., the infinite level sets of the resulting representative are Borel and null; redefine the representative to be zero on them. □

A continuous inverse image always preserves Borel measurability, but not necessarily Lebesgue measurability. In particular, a continuous map can have a large inverse image of a null set. Later proofs involving $f(x-y)$ will handle the Borel representative and its exceptional set separately.

2.2 Almost uniform convergence and continuity

Theorem 2.9 (Egorov). *Let E be measurable with $m(E) < \infty$, and let finite measurable f_j converge almost everywhere on E to a finite function f. For every $\varepsilon > 0$ there is a compact $K \subset E$ such that $m(E \setminus K) < \varepsilon$ and $f_j \to f$ uniformly on K.*

Proof. Remove a measurable null set containing all points where convergence fails or a value under consideration is infinite, and call the remaining set E'. The limit is measurable by Proposition 2.5. For positive integers N, k, define

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

These sets are measurable. For fixed k, they increase with N. Their union is E', because at every $x \in E'$ the numerical sequence converges. Continuity from below and $m(E') < \infty$ imply $m(E' \setminus E_{N,k}) \to 0$. Therefore one can choose N_k with

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

The dependence of N_k on k is allowed. Set $F = \bigcap_{k \geq 1} E_{N_k,k}$. Subadditivity gives $m(E' \setminus F) < \varepsilon/2$.

We check uniform convergence on F. Given $\eta > 0$, choose k with $1/k < \eta$. For every $j \geq N_k$ and every $x \in F$, the definition of $E_{N_k,k}$ gives $|f_j(x) - f(x)| \leq 1/k < \eta$. The same N_k works for all $x \in F$.

Finally, by compact inner approximation choose $K \subset F$ with $m(F \setminus K) < \varepsilon/2$. Uniform convergence persists on the smaller set. The removed set $E \setminus E'$ is null, so $m(E \setminus K) < \varepsilon$. □

The finite-measure hypothesis cannot be omitted. On $\mathbb{R}$, the functions $\mathbf{1}_{(j,\infty)}$ converge pointwise to zero. If A has finite measure, then $(j, \infty) \setminus A$ is nonempty for every j, so the convergence is not uniform on $\mathbb{R} \setminus A$.

Lemma 2.10. *If ϕ is simple on a finite-measure set E, then for every $\varepsilon > 0$ there is a compact $K \subset E$ with $m(E \setminus K) < \varepsilon$ such that $\phi|_K$ is continuous.*

Proof. Include the zero level set and write E as a disjoint union of its finitely many level sets $E_1, \ldots, E_N$. Choose compact $K_j \subset E_j$ with $m(E_j \setminus K_j) < \varepsilon/N$.

Empty level sets cause no difficulty and may be omitted. The union $K = \bigcup_j K_j$ is compact, and the measure loss is less than ε.

For each nonempty K_j, its distance from the union of the other nonempty compact sets is positive, unless there are no other sets. Thus K_j is relatively open in K. The restriction of ϕ is constant on each of these finitely many relatively open pieces, so it is continuous in the relative topology of K. □

Theorem 2.11 (Lusin). *Let E be measurable with finite measure, and let f be finite almost everywhere on E. Then f is measurable if and only if, for every $\varepsilon > 0$, there is a compact $K \subset E$ with $m(E \setminus K) < \varepsilon$ for which $f|_K$ is continuous.*

Proof. Suppose f is measurable. Work on the measurable full-measure set where it is finite. Choose simple functions $\phi_j \to f$ pointwise. Egorov's theorem provides compact K_0 losing less than $\varepsilon/2$ on which this convergence is uniform. For each $j \geq 1$, Lemma 2.10 gives compact K_j with loss less than $\varepsilon 2^{-j-1}$ such that $\phi_j|_{K_j}$ is continuous. Put $K = \bigcap_{j \geq 0} K_j$. This is a closed subset of the compact set K_0 and is compact. The sum of all losses is less than ε.

Every $\phi_j|_K$ is continuous, and convergence on K is uniform. To check continuity of $f|_K$ at $x \in K$, fix $\eta > 0$, choose j with $\sup_K |f - \phi_j| < \eta/3$, and then use continuity of $\phi_j|_K$ to make $|\phi_j(x) - \phi_j(y)| < \eta/3$. The triangle inequality gives $|f(x) - f(y)| < \eta$. Thus the limit is continuous.

Conversely, choose compact $K_j \subset E$ with $m(E \setminus K_j) < 1/j$ and continuous restrictions. The set $N = E \setminus \bigcup_j K_j$ is measurable and satisfies $m(N) \leq 1/j$ for every j, hence is null. For a real threshold a, the set $\{f > a\} \cap K_j$ is relatively open in K_j and therefore Borel. Now

$$\{f > a\} = \bigcup_j (\{f > a\} \cap K_j) \;\cup\; (\{f > a\} \cap N).$$

The last set is a subset of a null set and is measurable. This proves measurability. Real and imaginary parts give the complex version. □

Continuity of $f|_K$ is not continuity of f on its original domain at every point of K. For example, $\mathbf{1}_{\mathbb{Q}}$ is discontinuous at every point of $\mathbb{R}$, but its restriction to a compact set avoiding $\mathbb{Q}$ is identically zero.

For an infinite-measure set, the analogous approximation uses a closed set instead of a compact set. Indeed, partition E into $E_k = E \cap \{k - 1 \leq |x| < k\}$ and apply Lusin on each piece with error $\varepsilon 2^{-k}$. The resulting compact sets K_k are disjoint and locally finite: a bounded neighbourhood meets only finitely many shells. Their union is closed. Near any point of this union, only finitely many disjoint compact sets occur, and their positive separations show that the restrictions fit together continuously. The total measure loss is less than ε.

A compact set cannot have a finite-measure complement in a set of infinite measure.

For clarity, the locally finite union in the preceding argument has both required properties. Every bounded neighbourhood meets only finitely many of the compact sets K_j, because $K_j \subset \{j-1 \le |x| < j\}$. A point outside their union has a neighbourhood meeting only finitely many K_j; shrinking that neighbourhood to avoid these finitely many closed sets proves that the union is closed. At a point of K_j, shrink a bounded neighbourhood to avoid every other K_i meeting it. On the resulting neighbourhood inside the union, the function agrees with its continuous restriction to K_j. This proves continuity relative to the entire union. No assertion of ambient continuity on the original measurable domain is being made.

2.3 Convergence in measure

Definition 2.12. Finite measurable functions f_j converge in measure on E to f if

$$m(\{x \in E : |f_j(x) - f(x)| > \varepsilon\}) \longrightarrow 0 \quad \text{for every } \varepsilon > 0.$$

They are Cauchy in measure if the same condition holds for $|f_j - f_k|$ as $j, k \to \infty$.

This is a global condition. When $m(E) = \infty$, convergence in measure on every bounded subset is a different condition, called local convergence in measure.

Proposition 2.13. *On a finite-measure set, almost everywhere convergence implies convergence in measure. On any measurable set, a limit in measure is unique up to a null set.*

Proof. Assume $m(E) < \infty$ and $f_j \to f$ a.e. Fix $\varepsilon > 0$ and an arbitrary $\eta > 0$. Egorov gives a compact K with $m(E \setminus K) < \eta$ on which convergence is uniform. For all sufficiently large j, the error set $\{|f_j - f| > \varepsilon\}$ is contained in $E \setminus K$. Its limsup measure is therefore at most η. Let $\eta \downarrow 0$.

For uniqueness, if both f and g are limits in measure, then for every $\varepsilon > 0$,

$$\{|f - g| > 2\varepsilon\} \subset \{|f - f_j| > \varepsilon\} \cup \{|f_j - g| > \varepsilon\}.$$

Take measures and let $j \to \infty$. The set on the left is null. Taking its countable union over $\varepsilon = 1/k$ shows that $\{f \neq g\}$ is null. □

The first assertion fails on $\mathbb{R}$ for $f_j = \mathbf{1}_{(j,\infty)}$. Each nontrivial error set has infinite measure.

Example 2.14 (The typewriter sequence)**.** For each $k \geq 0$, list the indicators of $[j2^{-k}, (j+1)2^{-k})$, $0 \leq j < 2^k$, and concatenate these finite lists. Their supports have measures tending to zero, so the sequence converges in measure to zero on $[0, 1)$. Every point belongs to exactly one interval at each generation, and to none of the other intervals at that generation. Thus every point has value one infinitely often and value zero infinitely often. There is no pointwise limit anywhere in $[0, 1)$.

Lemma 2.15 (Borel–Cantelli)**.** *If A_j are measurable and $\sum_j m(A_j) < \infty$, then $m(\limsup_j A_j) = 0$, where $\limsup_j A_j = \bigcap_N \bigcup_{j \geq N} A_j$.*

Proof. For every N, subadditivity gives $m(\limsup A_j) \leq \sum_{j \geq N} m(A_j)$. The tail of this convergent numerical series tends to zero. A point outside the limsup belongs to only finitely many of the sets. □

Theorem 2.16. *If $f_j \to f$ in measure, some subsequence converges to f almost everywhere. No finiteness assumption on $m(E)$ is needed.*

Proof. Choose indices j_k recursively, with $j_{k+1} > j_k$, such that

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

Convergence in measure makes every such choice possible. The exceptional sets have summable measures. Outside their null limsup, there is, for each point x, an index $k(x)$ such that $|f_{j_k}(x) - f(x)| \leq 2^{-k}$ for all $k \geq k(x)$. This gives the pointwise limit on the complement of one null set. □

We also record the set inequality used in the next proof. For measurable A_k, put $H_N = \bigcap_{k \geq N} A_k$. Then $H_N \uparrow \liminf_k A_k$ and $m(H_N) \leq \inf_{k \geq N} m(A_k)$. Continuity from below gives

$$m(\liminf_k A_k) = \lim_N m(H_N) \leq \lim_N \inf_{k \geq N} m(A_k) = \liminf_k m(A_k).$$

This remains valid on an infinite-measure domain, since no subtraction is used.

Theorem 2.17. *A sequence of finite measurable functions is Cauchy in measure if and only if it converges in measure to a finite measurable function.*

Proof. If $f_j \to f$ in measure, then

$$\{|f_j - f_k| > \varepsilon\} \subset \{|f_j - f| > \varepsilon/2\} \cup \{|f_k - f| > \varepsilon/2\},$$

which proves the Cauchy condition.

Conversely, choose increasing indices j_k such that

$$m(\{|f_{j_{k+1}} - f_{j_k}| > 2^{-k}\}) < 2^{-k}.$$

For instance, first choose increasing thresholds N_k for the Cauchy condition with error and measure bounds 2^{-k}, and take $j_k \geq N_k$. Increase the choices if necessary so that both successive indices meet the relevant threshold. Borel–Cantelli shows that outside a null set the successive numerical differences are eventually bounded by 2^{-k}. Hence their series is absolutely convergent, and the telescoping sequence $f_{j_k}(x)$ has a finite limit $f(x)$. Define $f = 0$ on the exceptional set. The resulting function is measurable.

It remains to prove convergence in measure of the entire sequence; pointwise convergence of the chosen subsequence alone is not enough on a set of infinite measure. Fix $\varepsilon, \eta > 0$. The Cauchy condition gives N such that

$$m(\{|f_j - f_l| > \varepsilon/2\}) < \eta \quad (j, l \geq N).$$

For fixed $j \geq N$ and every point where $f_{j_k} \to f$, the inequality $|f_j - f| > \varepsilon$ implies $|f_j - f_{j_k}| > \varepsilon/2$ for all sufficiently large k. Thus, apart from a null set,

$$\{|f_j - f| > \varepsilon\} \subset \liminf_k \{|f_j - f_{j_k}| > \varepsilon/2\}.$$

For measurable sets B_k, continuity from below applied to $\bigcap_{k \geq K} B_k$ gives $m(\liminf B_k) \leq \liminf m(B_k)$. Therefore $m(\{|f_j - f| > \varepsilon\}) \leq \eta$ for every $j \geq N$. Since η is arbitrary, this is convergence in measure. □

Corollary 2.18. *On a finite-measure set, $f_j \to f$ in measure if and only if every subsequence has a further subsequence converging to f almost everywhere.*

Proof. Convergence in measure passes to subsequences, so one direction follows from Theorem 2.16. If convergence in measure fails, there are $\varepsilon, \eta > 0$ and a subsequence for which $m(\{|f_j - f| > \varepsilon\}) \geq \eta$ at every index. A further subsequence converging almost everywhere would converge in measure by Proposition 2.13, contradicting this lower bound. □

Exercises

Exercise 2.1. Express the set where a sequence of finite measurable functions has a finite limit using countable unions and intersections of measurable sets.

Exercise 2.2. If f is finite a.e. on a finite-measure set, prove $m(\{|f| > M\}) \to 0$. Show that this can fail on an infinite-measure set.

Exercise 2.3. Prove that an arbitrary supremum of continuous functions is lower semicontinuous and hence Borel measurable. Explain why an arbitrary supremum of measurable functions need not be measurable.

Exercise 2.4. Give uniform simple approximations to a bounded complex measurable function, with explicit bounds for the error and the number of values.

Exercise 2.5. Under Egorov's hypotheses, construct increasing compact sets K_l whose union has full measure and on each of which the convergence is uniform.

Exercise 2.6. Apply Lusin's theorem to $\mathbf{1}_A$. Explain the difference between relative continuity on K and ambient continuity at points of K.

Exercise 2.7. Show that sums preserve convergence in measure. On finite-measure sets, show the same for products, and identify the step using finiteness.

Exercise 2.8. Construct a sequence converging to zero in measure on $[0, 1)$ such that no tail converges uniformly outside a null set.

Exercise 2.9. After the integral is defined, prove that $d(f, g) = \int_0^1 \min(1, |f - g|)$ is a complete metric on a.e. classes and induces convergence in measure.

Exercise 2.10. On a finite-measure set, prove that $f_j \to f$ in measure implies $\Phi(f_j) \to \Phi(f)$ in measure for every continuous $\Phi : \mathbb{R} \to \mathbb{R}$.

Exercise 2.11. Prove that local convergence in measure on $\mathbb{R}^n$ gives an a.e. convergent subsequence by a diagonal choice over balls.

Exercise 2.12. Does $m(A_j) \to 0$ imply $m(\limsup A_j) = 0$? Compare with Borel–Cantelli.

Exercise 2.13. Construct two Lebesgue measurable functions on $[0, 1]$ whose composition is not Lebesgue measurable, using a homeomorphism taking a null Cantor set to a set of positive measure.

Exercise 2.14. Show that convergence in measure on a finite-measure set gives a subsequence converging almost uniformly. No uniform bound on the functions is needed.

Chapter 3

The Lebesgue integral

The integral is constructed in four stages. The order matters: the bounded convergence theorem is first proved on a finite-measure set, before nonnegative integrals are allowed to be infinite. Monotone convergence then controls the passage to the whole space, and an integrable majorant reduces dominated convergence to the bounded case.

3.1 Simple and bounded functions

Definition 3.1. If $\phi = \sum_{j=1}^{N} a_j \mathbf{1}_{E_j}$ is a nonnegative simple function in canonical form, define

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

For a measurable set A, write $\int_A \phi = \int \phi \mathbf{1}_A$. The convention $0 \cdot \infty = 0$ is used for the zero level.

Proposition 3.2. *The simple integral is independent of a noncanonical representation. It is monotone, additive, and positively homogeneous. For pairwise disjoint measurable A_k,*

$$\int_{\bigcup_k A_k} \phi = \sum_k \int_{A_k} \phi.$$

Proof. To handle a representation whose sets need not be disjoint, take all the finitely many sets $E_1, \dots, E_N$ that occur and form the atoms

$$D_\sigma = \bigcap_{j=1}^{N} E_j^{\sigma_j}, \quad E_j^1 = E_j, \quad E_j^0 = E_j^c, \quad \sigma \in \{0,1\}^N.$$

Empty atoms may be discarded. The remaining atoms are measurable and pairwise disjoint, and their union is the ambient space. On D_σ, the function $\sum_j a_j \mathbf{1}_{E_j}$ has the constant value $c_\sigma = \sum_{j:\sigma_j=1} a_j$. Finite additivity gives

$$\sum_j a_j m(E_j) = \sum_j a_j \sum_{\sigma:\sigma_j=1} m(D_\sigma) = \sum_\sigma c_\sigma m(D_\sigma).$$

In this nonnegative assertion all coefficients are nonnegative, so these rearrangements are valid even when a measure is infinite. Regrouping atoms with the same value gives precisely the canonical integral. For two representations, use the atoms formed from all the sets in both representations; the constant values then agree on every atom. This proves independence.

For two simple functions use the same finite atom partition. Their sum and a nonnegative scalar multiple have on each atom the sum and the corresponding multiple of their values. The finite-sum formula proves additivity and positive homogeneity. If one simple function is bounded above by the other, its value on each nonempty atom is bounded above by the other value, giving monotonicity. For a signed or complex representation, require every set carrying a nonzero coefficient to have finite measure. On the union of these finitely many sets, the same atom refinement involves only finite quantities. It proves independence and linearity, and the triangle inequality gives $|\int \phi| \le \int |\phi|$. This restriction prevents expressions such as the difference of two infinite integrals, even when their formal pointwise difference would be zero.

Finally, for pairwise disjoint measurable A_k, countable additivity on each fixed level gives

$$\int_{\bigcup_k A_k} \phi = \sum_{j=1}^N a_j \sum_k m(E_j \cap A_k) = \sum_k \sum_{j=1}^N a_j m(E_j \cap A_k) = \sum_k \int_{A_k} \phi.$$

The rearrangement is again of nonnegative terms. For signed simple functions of finite-measure support, apply the result to the positive and negative parts separately. No subtraction of infinite integrals is needed. □

A signed simple function whose support has finite measure has a finite integral defined by the same finite sum, or equivalently by integrating its positive and negative parts and subtracting.

Auxiliary lemma (Bounded pointwise simple approximation). Let $m(E) < \infty$ and let ϕ_j be real simple functions on E with $|\phi_j| \le M$. If $\phi_j \to f$ almost everywhere, then $(\int_E \phi_j)$ converges in $\mathbb{R}$. If $f = 0$ almost everywhere, this limit is zero.

Proof. When $M = 0$ or $m(E) = 0$, all integrals are zero. Otherwise fix $\varepsilon > 0$. Egorov's theorem gives a measurable $K \subset E$ such that $m(E \setminus K) < \varepsilon/(8M)$ and the convergence on K is uniform. In particular, (ϕ_j) is uniformly Cauchy on K. Choose J so that $\sup_K |\phi_j - \phi_k| < \varepsilon/(2(m(E)+1))$ whenever $j, k \geq J$. Simple integration and the domain decomposition $E = K \cup (E \setminus K)$ give

$$\Big|\int_E \phi_j - \int_E \phi_k\Big| \leq m(K) \sup_K |\phi_j - \phi_k| + 2Mm(E \setminus K) < \varepsilon.$$

Thus the scalar integrals form a Cauchy sequence, and completeness of $\mathbb{R}$ gives their limit. If $f = 0$, uniform convergence on the same K makes $\sup_K |\phi_j|$ arbitrarily small. The inequality $|\int_E \phi_j| \leq m(K) \sup_K |\phi_j| + Mm(E \setminus K)$ then makes the limit zero. We have used only simple integration, not an integral of f or a convergence theorem for that integral. □

Definition 3.3. Let f be bounded and real measurable on a measurable set E of finite measure. Choose uniformly bounded simple functions ϕ_j converging to f almost everywhere on E and define

$$\int_E f = \lim_{j\to\infty} \int_E \phi_j.$$

For complex f, integrate its real and imaginary parts.

Proposition 3.4. *This limit exists and is independent of the approximating sequence. The bounded integral is linear and monotone, and*

$$\Big|\int_E f\Big| \leq \int_E |f| \leq m(E) \sup_E |f|.$$

If $0 \leq f \leq M$, its integral is the supremum of the integrals of simple ϕ satisfying $0 \leq \phi \leq f$.

Proof. The preceding lemma proves existence. If (ψ_j) is another uniformly bounded simple approximation to f, then $\phi_j - \psi_j$ is uniformly bounded, is simple, and converges to zero almost everywhere. The zero-limit assertion of the lemma gives $\int_E \phi_j - \int_E \psi_j \to 0$. This proves independence. In particular, we are free to use the uniform dyadic approximations supplied by Theorem 2.7.

For real f, g and real constants a, b, the simple sequence $a\phi_j + b\psi_j$ converges to $af + bg$ and has a common finite bound. The simple identity $\int_E (a\phi_j + b\psi_j) = a\int_E \phi_j + b\int_E \psi_j$ passes to the limit and gives linearity. For monotonicity, choose uniform approximants with errors at most $\delta_j \downarrow 0$. If $f \leq g$, then $\phi_j \leq \psi_j + 2\delta_j$ on E, so

$$\int_E \phi_j \leq \int_E \psi_j + 2\delta_j m(E).$$

Let $j \to \infty$. A bounded function that is zero almost everywhere has integral zero by the lemma, so the same arguments apply to almost everywhere inequalities and show invariance under bounded null-set modifications.

The inequalities $-|f| \le f \le |f|$ give the real absolute-value bound, and $|f| \le \sup_E |f|$ gives its second part. Defining the integral componentwise for complex functions gives complex linearity. For $z = \int_E f \neq 0$, choose θ with $e^{-i\theta} z = |z|$. Real monotonicity gives

$$|z| = \int_E \mathrm{Re}(e^{-i\theta} f) \le \int_E |f|.$$

The case $z = 0$ is immediate.

We will also use finite additivity in the domain. If $E = A \cup B$ is a disjoint measurable decomposition, restrict the same uniformly bounded simple approximants to A and B. The identity $\int_E \phi_j = \int_A \phi_j + \int_B \phi_j$ passes to the limit. A finite disjoint decomposition follows by repetition. Extending a function by zero to a larger finite-measure set therefore leaves its integral unchanged.

If $0 \le f \le M$, monotonicity bounds every lower simple integral by $\int_E f$. The lower dyadic approximants lie below f and converge uniformly once the truncation level exceeds M. Their integrals approach $\int_E f$, proving the reverse inequality for the supremum. □

We also integrate an essentially bounded measurable function on E by first choosing a bounded representative: set it equal to zero on a measurable null set outside which a finite bound holds. Any two such representatives are bounded and agree almost everywhere, so the preceding proposition gives the same integral. This extends the bounded definition without changing any of its properties.

Theorem 3.5 (Bounded convergence)**.** *Suppose $m(E) < \infty$, the measurable functions f_j satisfy $|f_j| \le M$ a.e. on E, and $f_j \to f$ almost everywhere. Then*

$$\int_E |f_j - f| \longrightarrow 0, \quad \int_E f_j \longrightarrow \int_E f.$$

Proof. Remove the countable union of the null sets where the bounds fail, together with the null set where convergence fails. On the remaining set, $|f| \le M$. Redefine all functions to be zero on this common exceptional set; their bounded integrals do not change, since a bounded function supported on a null set has integral zero by the preceding estimate.

If $M = 0$, there is nothing to prove. Otherwise fix $\varepsilon > 0$. Egorov gives a compact $K \subset E$ such that $m(E \setminus K) < \varepsilon/(4M)$ and convergence on K is uniform.

For large j, make the uniform error less than $\varepsilon/(2(m(E)+1))$. Splitting the domain gives

$$\begin{aligned}\int_E |f_j - f| &\le m(K)\sup_K |f_j - f| + 2Mm(E \setminus K) \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon.\end{aligned}$$

This proves the first limit. The inequality $|\int_E f_j - \int_E f| \le \int_E |f_j - f|$ proves the second. □

The proof uses only the bounded integral on finite-measure sets and Egorov's theorem. Neither monotone nor dominated convergence has been used to establish it.

3.2 Nonnegative functions and monotone limits

Definition 3.6. For a nonnegative measurable f on $\mathbb{R}^n$, define

$$\int f = \sup\Big\{\int g : 0 \le g \le f,\ g \in \mathcal{B}\Big\},$$

where $\mathcal{B}$ is the class of bounded measurable functions g with $m(\{g \neq 0\}) < \infty$. For a measurable domain E, extend f by zero and write $\int_E f = \int f\mathbf{1}_E$.

The value can be infinite. If f is bounded and vanishes outside a finite-measure set, then f itself is an admissible lower function, and bounded monotonicity bounds every other admissible integral above by $\int f$. Thus this definition agrees with the bounded integral.

Equivalently, the same supremum may be taken only over nonnegative simple lower functions whose nonzero sets have finite measure. One inequality follows because such simple functions are allowed in the displayed definition. For the other, fix an admissible bounded g and let $\phi_k \le g$ be its lower dyadic approximants on $\{g \neq 0\}$, extended by zero. They are admissible simple lower functions for f and converge uniformly to g. Their integrals tend to $\int g$ by the bounded definition. Hence the simple supremum is at least $\int g$ for every such g; taking the supremum in g proves equivalence.

This definition also agrees with the nonnegative simple integral when a level set has infinite measure. For $\phi = \sum_{j=1}^{L} a_j \mathbf{1}_{E_j}$, the truncated simple functions $\phi\mathbf{1}_{[-k,k]^n}$ have integrals increasing to $\sum_{j=1}^{L} a_j m(E_j)$ as $k \to \infty$, by continuity from below on each fixed level. Conversely, any admissible bounded lower function $g \leq \phi$ has $\int g \leq \sum_j a_j m(E_j)$ by bounded comparison on its finite-measure nonzero set and finite domain splitting. These two inequalities identify the definitions.

Monotonicity follows directly from inclusion of the families being supremized. Also, changing f on a null set does not change its integral. Indeed, for an admissible simple ϕ, deleting its values on that null set leaves its simple integral unchanged and gives an admissible lower approximation off the exceptional set. Applying the same argument in the opposite direction gives equality.

Theorem 3.7 (Monotone convergence)**.** *If* $0 \leq f_j \uparrow f$ *almost everywhere, then* $\int f_j \uparrow \int f$.

Proof. After a null modification, assume the monotonicity and convergence hold everywhere. Monotonicity of the integral gives $I = \lim_j \int f_j \leq \int f$. To prove the reverse inequality, fix a simple function ϕ with finite-measure support such that $0 \leq \phi \leq f$, and fix $0 < c < 1$. Put

$$A_j = \{x : f_j(x) \geq c\phi(x)\}.$$

These sets are measurable and increasing. Every point where $\phi(x) > 0$ eventually belongs to A_j: the limit $f(x) \geq \phi(x)$ is strictly greater than $c\phi(x)$. The points where $\phi = 0$ already belong to every A_j. Thus $A_j \uparrow \mathbb{R}^n$.

Write $\phi = \sum_{l=1}^{N} a_l \mathbf{1}_{E_l}$ canonically. Every nonzero level has finite measure. Continuity from below gives $m(E_l \cap A_j) \to m(E_l)$, and hence $\int \phi \mathbf{1}_{A_j} \to \int \phi$. Since $f_j \geq c\phi\mathbf{1}_{A_j}$, the definition yields

$$I \geq c \int \phi.$$

Let $c \uparrow 1$, and then take the supremum over the admissible simple ϕ. This gives $I \geq \int f$. If $\int f = \infty$, the same argument says that I exceeds every finite lower simple integral, so $I = \infty$ as required. □

Proposition 3.8. *For nonnegative measurable* f, g *and* $a \geq 0$, *the nonnegative integral satisfies* $\int (f+g) = \int f + \int g$ *and* $\int af = a \int f$, *with the stated convention when* $a = 0$. *Moreover,*

$$\int \sum_{j=1}^{\infty} f_j = \sum_{j=1}^{\infty} \int f_j \quad (f_j \geq 0).$$

Proof. Use increasing simple approximations to f and g, and multiply their kth members by $\mathbf{1}_{[-k,k]^n}$. The resulting sequences s_k, t_k remain increasing, have finite-measure support, and converge respectively to f, g. The simple identity gives $\int(s_k + t_k) = \int s_k + \int t_k$. Apply monotone convergence to all three sequences. Homogeneity follows in the same way. Finally, the finite partial sums of a nonnegative series increase to its sum; monotone convergence and finite additivity give the series identity. □

In particular, integration is countably additive over disjoint measurable domains. For any nonnegative f, one also has the exhaustion formula

$$\int f = \lim_{N\to\infty} \int_{[-N,N]^n} \min(f, N). \tag{3.1}$$

Each integral on the right is a bounded integral on a finite-measure set. Thus the third stage is explicitly obtained from the second stage by monotone exhaustion.

Proposition 3.9. *If $f \geq 0$ is measurable, then $\int f = 0$ if and only if $f = 0$ a.e. If $\int f < \infty$, then $f < \infty$ a.e. For $t > 0$,*

$$m(\{f > t\}) \leq \frac{1}{t} \int f. \tag{3.2}$$

Proof. The pointwise inequality $f \geq t\mathbf{1}_{\{f>t\}}$ and monotonicity give (3.2). If the integral is zero, every set $\{f > 1/k\}$ is null, and their union is $\{f > 0\}$. Conversely, a null modification of zero has integral zero.

If the integral is finite, $m(\{f = \infty\}) \leq m(\{f > k\}) \leq k^{-1} \int f$ for every k, so the infinite level set is null. No finiteness of the whole domain is needed. □

Theorem 3.10 (Fatou)**.** *For nonnegative measurable f_j,*

$$\int \liminf_j f_j \leq \liminf_j \int f_j.$$

Proof. Set $f = \liminf_j f_j$, which is nonnegative and measurable. Fix a bounded nonnegative measurable $g \leq f$ whose nonzero set E has finite measure, and put $g_j = \min(g, f_j)$. Each g_j is measurable, vanishes outside E, and is bounded by the same constant as g. At a point where $g(x) = 0$ one has $g_j(x) = 0$. If $g(x) > 0$, then for every $\eta > 0$ the inequality $\liminf_j f_j(x) \geq g(x)$ implies $f_j(x) > g(x) - \eta$ for all sufficiently large j. Hence $g(x) - \eta < g_j(x) \leq g(x)$ eventually, proving $g_j(x) \to g(x)$.

Bounded convergence therefore gives $\int g_j \to \int g$. Since $g_j \leq f_j$, monotonicity gives $\int g_j \leq \int f_j$ for every j. Taking the lower limit yields $\int g \leq \liminf_j \int f_j$. The right side is independent of the chosen lower function g.

Taking the supremum over all such g gives the desired inequality, including when $\int f = \infty$. This proof follows the bounded approximation argument and does not require an interchange of an unbounded limit and integral. □

3.3 Integrable functions and dominated convergence

Definition 3.11. A measurable real or complex function f is integrable when $\int |f| < \infty$. For real f, define $\int f = \int f^+ - \int f^-$. For complex f, integrate the real and imaginary parts. The resulting equivalence classes modulo equality almost everywhere form L^1.

Both integrals in the real definition are finite. An expression whose positive and negative integrals are both infinite is not an integrable signed function and is not assigned an integral by cancellation. Integrable functions can be made finite everywhere by a null modification.

Proposition 3.12. *Integrable functions form a vector space, their integral is linear, and* $|\int f| \leq \int |f|$.

Proof. The inequalities $|f + g| \leq |f| + |g|$ and $|af| = |a||f|$ prove integrability. For real functions, the pointwise identity

$$(f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+$$

contains only nonnegative terms. Integrate using Proposition 3.8 and subtract the finite quantities to obtain additivity. Homogeneity for positive scalars is already known, and changing sign handles negative scalars. Real and imaginary parts give complex linearity. The real absolute-value estimate follows from $-|f| \leq f \leq |f|$, and the complex estimate follows by rotating the integral as in Proposition 3.4. □

Proposition 3.13. *If* $f \in L^1(\mathbb{R}^n)$, *then:*

(i) *for every* $\varepsilon > 0$, *some ball* B *satisfies* $\int_{B^c} |f| < \varepsilon$;

(ii) *for every* $\varepsilon > 0$, *some* $\delta > 0$ *satisfies* $\int_A |f| < \varepsilon$ *whenever* A *is measurable and* $m(A) < \delta$.

Proof. Apply monotone convergence to $|f|\mathbf{1}_{B(0,N)} \uparrow |f|$. Because the full integral is finite, subtracting the interior integral shows that the exterior integral tends to zero.

For the second assertion, let $g = |f|$ and use the magnitude truncations $g\mathbf{1}_{\{g \leq N\}} \uparrow g$ after redefining the null infinite level set. Choose $N \geq 1$ with

$\int_{\{g>N\}} g < \varepsilon/2$, and then choose $\delta = \varepsilon/(2N)$. For every measurable A with $m(A) < \delta$,

$$\int_A g \le \int_{\{g>N\}} g + \int_{A\cap\{g\le N\}} g < \varepsilon/2 + Nm(A) < \varepsilon.$$

The cutoff N is chosen before δ, and depends on f and ε, not on A. □

Theorem 3.14 (Dominated convergence)**.** *Let measurable f_j converge almost everywhere to f. Suppose $|f_j| \le g$ a.e. for every j, where $g \ge 0$ is integrable. Then $f \in L^1$ and*

$$\int |f_j - f| \longrightarrow 0, \quad \int f_j \longrightarrow \int f.$$

Proof. Remove one null set containing the exceptional sets for all bounds, the convergence, and the finiteness of g. On its complement, $|f| \le g$ by passage to the pointwise limit. Define every function to be zero on the exceptional set. Then f is measurable and integrable by domination.

Set $E_N = \{x : |x| < N,\ g(x) \le N\}$. These sets have finite measure, increase to the whole space after the null modification, and satisfy $|f_j|, |f| \le N$ on E_N. Monotone convergence applied to $g\mathbf{1}_{E_N}$ and finiteness of $\int g$ give $\int_{E_N^c} g \to 0$.

Fix $\varepsilon > 0$. First choose N with $2\int_{E_N^c} g < \varepsilon/2$. Keep this N fixed. The bounded convergence theorem applied on E_N gives an index J such that $\int_{E_N} |f_j - f| < \varepsilon/2$ whenever $j \ge J$. Therefore

$$\int |f_j - f| \le \int_{E_N} |f_j - f| + 2\int_{E_N^c} g < \varepsilon.$$

The integral limit follows from $|\int f_j - \int f| \le \int |f_j - f|$. □

This proof follows the two truncations separately: spatial and value truncation reduce to a finite-measure bounded problem; the integrable majorant controls the discarded part uniformly in j.

Corollary 3.15. *If $\sum_j \int |f_j| < \infty$, then $\sum_j f_j$ converges absolutely almost everywhere to an integrable function f, converges in L^1, and $\int f = \sum_j \int f_j$.*

Proof. The nonnegative series theorem gives $\int \sum_j |f_j| = \sum_j \int |f_j| < \infty$. Thus $g = \sum_j |f_j|$ is finite a.e. The partial sums converge absolutely there and are dominated by g. Dominated convergence gives both the L^1 convergence and the integral identity. Alternatively, the L^1 error is at most the tail $\sum_{j>N} \int |f_j|$. □

Example 3.16. On $(0,1)$, $f_j = j\mathbf{1}_{(0,1/j)}$ converges a.e. to zero but has integral one. The functions are not dominated by one integrable majorant: such a majorant would contradict dominated convergence. On $\mathbb{R}$, $g_j = \mathbf{1}_{(j,j+1)}$ also converges pointwise to zero with integral one. The first sequence concentrates its mass, while the second moves it to infinity. Both possibilities are excluded by a common integrable majorant.

3.4 Comparison with Riemann integration

Theorem 3.17. *A Riemann integrable function on a compact interval is Lebesgue measurable and integrable, and its two integrals agree.*

Proof. Let f be real and Riemann integrable on $[a,b]$. It is bounded, say $|f| \leq M$. Choose partitions P_j, each refining the preceding one, for which the upper and lower Darboux sums have difference tending to zero. Such nested partitions are obtained by taking common refinements of any sequence whose gaps tend to zero.

On every open partition interval, define l_j and u_j as the infimum and supremum of f on the corresponding closed interval. Assign any bounded values at the finitely many endpoints of that partition. These functions are measurable step functions apart from finitely many endpoint values; their Lebesgue integrals equal the lower and upper Darboux sums. Let D be the countable union of all partition endpoints. At every $x \notin D$, refinement gives $l_j(x) \uparrow l(x)$ and $u_j(x) \downarrow u(x)$, with $l(x) \leq f(x) \leq u(x)$. Set $l = u = 0$ on D. These are measurable bounded functions, and convergence of l_j, u_j to them holds a.e.

Bounded convergence implies $\int(u-l) = \lim_j(\int u_j - \int l_j) = 0$. Since $u - l \geq 0$ off the null set D, it follows that $u = l$ a.e. The sandwich then gives $f = l$ a.e. Thus f is Lebesgue measurable and integrable by completeness and boundedness. Finally,

$$\int l_j \leq \int f \leq \int u_j,$$

where any failure of the pointwise inequalities at finitely many endpoints has zero Lebesgue integral. The two bounds tend to the Riemann integral, identifying the values.

Only finitely many values of each individual step function were assigned at its own partition endpoints. The countable union D was used solely for almost everywhere convergence and Lebesgue measurability; no assertion that an arbitrary countable modification preserves Riemann integrability is needed. □

We use one elementary compactness fact in the next proof. A finite open cover $U_1, \dots, U_N$ of a compact metric set K has a number $\delta > 0$ such that every nonempty subset of K of diameter less than δ lies in one U_i. For each $x \in K$, choose $r_x > 0$ so that $B(x, 2r_x) \cap K$ lies in one member of the cover. Finitely many balls $B(x, r_x)$ cover K. Let δ be the minimum of their radii. If $A \subset K$ has diameter less than δ, choose $a \in A$ and a selected ball $B(x, r_x)$ containing a. For $b \in A$, $d(b, x) \le d(b, a) + d(a, x) < \delta + r_x \le 2r_x$. Thus A lies in the same member of the cover. Such a δ is called a Lebesgue number.

Theorem 3.18 (Lebesgue's criterion)**.** *A bounded real function on $[a, b]$ is Riemann integrable if and only if its discontinuity set is null.*

Proof. For $x \in [a, b]$, define

$$\omega_f(x) = \inf_{r>0} \sup\{|f(y) - f(z)| : y, z \in [a, b] \cap (x - r, x + r)\}.$$

The function is continuous at x precisely when $\omega_f(x) = 0$. For $\eta > 0$, the set $D_\eta = \{x : \omega_f(x) \ge \eta\}$ is closed relative to $[a, b]$. Indeed, if the oscillation in one neighbourhood is less than η, then every sufficiently close point has a smaller neighbourhood contained in it, and also has oscillation less than η. The discontinuity set is $D = \bigcup_{k\ge1} D_{1/k}$.

If f is Riemann integrable, choose a partition with Darboux gap less than ε. Let J run over the subintervals whose oscillations are at least η. Then $\eta \sum_J m(J) \le \sum_J \operatorname{osc}_J(f) m(J) < \varepsilon$. Apart from the finite partition endpoints, these intervals cover D_η: a point inside any other subinterval has local oscillation less than η. Therefore $m(D_\eta) \le \varepsilon/\eta$. Let $\varepsilon \downarrow 0$ and then take the countable union in k.

Conversely, suppose D is null and $|f| \le M$. Fix $\eta, \varepsilon > 0$. The compact null set D_η has a finite open-interval cover whose total length is less than ε. Let O be their union. Every point of the compact set $[a, b] \setminus O$ has a relative open neighbourhood on which the oscillation of f is less than η. Together with O, these neighbourhoods cover $[a, b]$. Choose a finite subcover and a Lebesgue number $\delta > 0$ for it. A partition of mesh less than δ has every subinterval contained in one member of the cover.

Assign each subinterval either to O or to a small-oscillation neighbourhood containing it. The subintervals assigned to O have disjoint interiors and total length at most $m(O) < \varepsilon$. Their Darboux contribution is at most $2M\varepsilon$. All other subintervals contribute at most $\eta(b - a)$. Thus the Darboux gap is at most $2M\varepsilon + \eta(b - a)$. Since both parameters can be made arbitrarily small, f is Riemann integrable. ∎

The function $\mathbf{1}_{\mathbb{Q}\cap[0,1]}$ has Lebesgue integral zero and is nowhere continuous, so it is not Riemann integrable. Equality almost everywhere preserves the Lebesgue integral, but does not preserve Riemann integrability.

3.5 Integrals depending on a parameter

Proposition 3.19. *Let $I \subset \mathbb{R}$ be open and let $F(t,x)$ be measurable in x for each $t \in I$. Fix $t_0 \in I$.*

(i) *Suppose $F(t,x) \to F(t_0,x)$ as $t \to t_0$ for a.e. x, and $|F(t,x)| \le g(x)$ for t near t_0, outside one null set, with $g \in L^1$. Then $t \mapsto \int F(t,x)\,dx$ is continuous at t_0.*

(ii) *Suppose $F(t_0,\cdot) \in L^1$, the function $t \mapsto F(t,x)$ is differentiable on a fixed neighbourhood of t_0 outside one null set, and $|\partial_t F(t,x)| \le g(x)$ there with $g \in L^1$. Then*
$$\frac{d}{dt}\int F(t,x)\,dx\Big|_{t=t_0} = \int \partial_t F(t_0,x)\,dx.$$

Proof. For the first assertion, take any sequence $t_j \to t_0$ contained eventually in the neighbourhood. Dominated convergence gives convergence of the integrals. Sequential continuity characterizes continuity of scalar functions.

For the second assertion, the mean value theorem bounds the real difference quotient by g. Applying it separately to real and imaginary parts gives the bound $2g$ for a complex function. The same estimate shows $|F(t,x)| \le |F(t_0,x)| + 2|t-t_0|g(x)$, and hence integrability for nearby t. For each nonzero h, the difference quotient is measurable in x, and its pointwise limit is $\partial_t F(t_0,x)$ off the common null set. The derivative is therefore measurable. Dominated convergence along any sequence $h \to 0$ gives the displayed identity. □

A derivative estimate only at t_0 does not bound nearby difference quotients. The neighbourhood and the common integrable majorant are essential parts of the hypothesis.

Exercises

Exercise 3.1. Prove directly that the nonnegative integral can also be defined by the exhaustion in (3.1). Show that the answer is unchanged for any increasing finite-measure exhaustion of the space.

Exercise 3.2. If $f \ge 0$ and $\int_A f = 0$ for every bounded measurable A, prove $f = 0$ a.e. Extend the assertion to locally integrable signed functions.

Exercise 3.3. Construct nonnegative functions tending to zero a.e. whose integrals tend to infinity, and another sequence whose integrals have no limit.

Exercise 3.4. If $f_j, f \geq 0$, $f_j \to f$ a.e., and $\int f_j \to \int f < \infty$, prove $\int |f_j - f| \to 0$. Use $|f_j - f| = f_j + f - 2\min(f_j, f)$.

Exercise 3.5. Prove the dominated version of Fatou's lemma for real functions bounded below by one integrable function.

Exercise 3.6. Show that absolute convergence of $\sum_j \int |f_j|$ permits termwise integration, and give an example showing why conditional numerical convergence of $\sum_j \int f_j$ is insufficient.

Exercise 3.7. Prove that $\int_{\{|f|>N\}} |f| \to 0$ for $f \in L^1$. Does the converse hold for a finite-valued measurable function on an infinite-measure set?

Exercise 3.8. Construct an integrable function on $(0, 1)$ that is unbounded on every nonempty subinterval.

Exercise 3.9. Show that a nonnegative measurable function has a measurable subgraph and that its integral is the area of that subgraph once repeated integration is established.

Exercise 3.10. Give a bounded function whose Lebesgue integral exists but whose improper Riemann integral does not. Distinguish this from conditional improper integrability.

Exercise 3.11. Prove the bounded convergence theorem directly from convergence in measure on a finite-measure set, assuming a uniform bound.

Exercise 3.12. Give a differentiable parameter family showing that an integrable bound for $\partial_t F(t_0, \cdot)$ alone does not justify differentiation under the integral.

Exercise 3.13. Prove Jensen's inequality for a finite convex function on $\mathbb{R}$ and an integrable real function on a probability space, assuming the composed function is integrable. Establish the supporting-line inequality first.

Exercise 3.14. If $f_j \to f$ in measure and $|f_j| \leq g \in L^1$, prove $f_j \to f$ in L^1 by the subsequence principle.

Chapter 4

The space L^1 and repeated integration

We first prove completeness and approximation in L^1. Repeated integration is then established by following the geometry of measurable sets. In this proof, the assertion about almost every section is obtained before the interchange of integrals is used.

4.1 Completeness and approximation in L^1

Functions in $L^1(E)$ are identified when they agree almost everywhere, and the norm is $\|f\|_1 = \int_E |f|$. Without this identification, a nonzero function supported on a null set would have norm zero. The integral inequalities from Chapter 3 give the norm properties.

Theorem 4.1 (Riesz–Fischer)**.** *The normed space $L^1(E)$ is complete.*

Proof. Let (f_j) be Cauchy in L^1. Choose an increasing subsequence of indices j_k with

$$\|f_{j_{k+1}} - f_{j_k}\|_1 \le 2^{-k}.$$

This is possible by choosing the indices beyond successive Cauchy thresholds. Select measurable representatives finite everywhere after null modifications. Define

$$g = |f_{j_1}| + \sum_{k=1}^{\infty} |f_{j_{k+1}} - f_{j_k}|.$$

The series theorem for nonnegative functions gives $\int g \le \|f_{j_1}\|_1 + \sum_k 2^{-k} < \infty$. Thus g is finite outside a measurable null set N.

For $x \notin N$, the telescoping series

$$f_{j_1}(x) + \sum_{k=1}^{\infty} (f_{j_{k+1}}(x) - f_{j_k}(x))$$

converges absolutely. Denote its sum by $f(x)$ and set $f = 0$ on N. This is a measurable function, and $|f| \le g$ a.e., so $f \in L^1$. The series tail gives, for each K,

$$|f - f_{j_K}| \le \sum_{k \ge K} |f_{j_{k+1}} - f_{j_k}| \quad \text{a.e.}$$

Integrating yields $\|f - f_{j_K}\|_1 \le \sum_{k \ge K} 2^{-k} \to 0$.

To return to the whole sequence, fix $\varepsilon > 0$ and choose J such that $\|f_j - f_l\|_1 < \varepsilon/2$ whenever $j, l \ge J$. Choose K with $j_K \ge J$ and $\|f_{j_K} - f\|_1 < \varepsilon/2$. Then for every $j \ge J$,

$$\|f_j - f\|_1 \le \|f_j - f_{j_K}\|_1 + \|f_{j_K} - f\|_1 < \varepsilon.$$

This proves completeness. □

Corollary 4.2. *If $f_j \to f$ in L^1, then some subsequence converges to f almost everywhere.*

Proof. Choose j_k with $\|f_{j_k} - f\|_1 < 2^{-k}$. Then $\int \sum_k |f_{j_k} - f| < \infty$, so the numerical series is finite a.e. Its summands therefore tend to zero at almost every point. □

Proposition 4.3. *Simple functions with finite-measure nonzero sets, step functions with bounded rectangular supports, and $C_c(\mathbb{R}^n)$ are dense in $L^1(\mathbb{R}^n)$.*

Proof. Let $f \in L^1$. Truncate both its location and its magnitude by setting

$$f_N = f \mathbf{1}_{\{|x| \le N,\ |f| \le N\}}.$$

The error tends to zero in L^1 by dominated convergence, since $|f - f_N| \le |f|$. On the finite-measure set $[-N, N]^n$, approximate the bounded f_N uniformly by a simple function. The L^1 error is bounded by the uniform error times the measure of this cube. This proves simple-function density.

Consider one nonzero level set A of such a simple function. It has finite measure. By Theorem 1.15, choose a finite union R of closed cubes with $m(A \triangle R)$ arbitrarily small. The identity $\|\mathbf{1}_A - \mathbf{1}_R\|_1 = m(A \triangle R)$ controls the error for one indicator. Also $\mathbf{1}_R$ really is a step function even if the cubes overlap: for $R = \bigcup_{j=1}^{N} Q_j$, inclusion–exclusion gives

$$\mathbf{1}_R = \sum_{\varnothing \ne J \subset \{1,\dots,N\}} (-1)^{\#J+1} \mathbf{1}_{\bigcap_{j \in J} Q_j}.$$

Here $\#J$ is the number of indices in J. Each nonempty intersection is a closed rectangle, possibly degenerate. To verify the identity at a point lying in exactly $r \geq 1$ cubes, its right side is $\sum_{k=1}^{r}(-1)^{k+1}\binom{r}{k} = 1$; at a point in none of the cubes it is zero. Finally, if $s = \sum_{l=1}^{L} a_l \mathbf{1}_{A_l}$, choose R_l with $m(A_l \triangle R_l) < \varepsilon/(L(1+|a_l|))$. Then

$$\left\| s - \sum_{l=1}^{L} a_l \mathbf{1}_{R_l} \right\|_1 \leq \sum_{l=1}^{L} |a_l| m(A_l \triangle R_l) < \varepsilon.$$

This proves step-function density.

To approximate $\mathbf{1}_A$ by a continuous function, choose compact $K \subset A$ with $m(A \setminus K) < \eta/2$. If $K = \varnothing$, zero already gives the desired estimate when the loss is small. Otherwise choose a bounded open $O \supset K$ with $m(O \setminus K) < \eta/2$. Such an O is obtained from open regularity and then intersection with a sufficiently large ball containing K. The distance $\operatorname{dist}(K, O^c)$ is positive. Set

$$\phi(x) = \frac{\operatorname{dist}(x, O^c)}{\operatorname{dist}(x, O^c) + \operatorname{dist}(x, K)}.$$

The denominator never vanishes, $0 \leq \phi \leq 1$, $\phi = 1$ on K, and $\phi = 0$ on O^c. Since O is bounded, $\phi \in C_c$. Moreover,

$$\|\mathbf{1}_A - \phi\|_1 \leq m(A \setminus K) + m(O \setminus K) < \eta.$$

Approximate each of the finitely many indicators in a simple function with an error small enough after multiplication by its coefficient. This proves density of C_c. □

Proposition 4.4. *For integrable f and $h \in \mathbb{R}^n$, define $\tau_h f(x) = f(x-h)$. Then*

$$\int \tau_h f = \int f, \quad \|\tau_h f\|_1 = \|f\|_1, \quad \|\tau_h f - f\|_1 \to 0 \quad (h \to 0).$$

Also, for $a > 0$, $\int f(ax)\,dx = a^{-n} \int f(x)\,dx$, and coordinate reflections preserve the integral.

Proof. For indicators, the integral identities are exactly the measure invariances from Proposition 1.18. They extend by finite linearity to simple functions, by increasing simple approximation and monotone convergence to nonnegative functions, and by positive/negative and real/imaginary decompositions to integrable functions. Applying the translation identity to $|f|$ proves the norm identity.

If $\phi \in C_c$, then it is uniformly continuous on $\mathbb{R}^n$. For $|h| \leq 1$, all supports of $\tau_h \phi - \phi$ lie in one fixed bounded set A. Thus $\|\tau_h \phi - \phi\|_1 \leq m(A) \sup_x |\phi(x-h) - \phi(x)| \to 0$.

For general f, choose $\phi \in C_c$ with $\|f - \phi\|_1 < \varepsilon$. Translation isometry gives

$$\|\tau_h f - f\|_1 \le 2\varepsilon + \|\tau_h \phi - \phi\|_1.$$

Take $h \to 0$ and then $\varepsilon \downarrow 0$. □

4.2 Sections of sets and Fubini's theorem

Write a point of $\mathbb{R}^{n+d}$ as (x, y) with $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^d$. For a set $A \subset \mathbb{R}^{n+d}$, define $A^y = \{x : (x, y) \in A\}$. For a function f, write $f^y(x) = f(x, y)$. A Borel set has a Borel section for every y, since $x \mapsto (x, y)$ is continuous. A Lebesgue measurable set need not have a measurable section for every y.

Example 4.5. Choose the nonmeasurable $V \subset [0, 1]$ supplied by Theorem 1.22. The set $V \times \{0\}$ is contained in the degenerate rectangle $[0, 1] \times \{0\}$, whose outer measure is zero by Proposition 1.5. Hence $V \times \{0\}$ is Lebesgue measurable by completeness. Its section at $y = 0$ is V, which is not measurable. This is the reason that exceptional parameters cannot be omitted from the theorem.

Lemma 4.6. *For open $O \subset \mathbb{R}^{n+d}$, the function $y \mapsto m_n(O^y)$ is measurable and*

$$\int_{\mathbb{R}^d} m_n(O^y)\, dy = m_{n+d}(O),$$

with infinity allowed.

Proof. First let $Q = I \times J$ be a half open rectangle, with $I \subset \mathbb{R}^n$ and $J \subset \mathbb{R}^d$. Then $Q^y = I$ for $y \in J$ and is empty otherwise. Consequently $m_n(Q^y) = m_n(I)\mathbf{1}_J(y)$, and its integral is $m_n(I)m_d(J) = m_{n+d}(Q)$.

By Lemma 1.2, write $O = \bigcup_{k \ge 1} Q_k$ as a disjoint union of half open dyadic cubes. Each cube has the product form just considered. For every y, the sections Q_k^y are disjoint, so countable additivity in $\mathbb{R}^n$ gives

$$m_n(O^y) = \sum_k m_n(Q_k^y).$$

The summands are measurable functions of y. Monotone convergence for their partial sums yields

$$\int m_n(O^y)\, dy = \sum_k m_{n+d}(Q_k) = m_{n+d}(O).$$

The final equality is countable additivity in the full space. No theorem about iterated integrals has been used. □

The hyperplane nullity used in this section does not require Fubini. A bounded part of a coordinate hyperplane lies in a degenerate rectangle, and so has outer measure zero by Proposition 1.5. The whole hyperplane is a countable union of such bounded parts. Subadditivity therefore makes it null. If $N \subset \mathbb{R}^n$ is null, then $N \times \mathbb{R}^d$ is null as well: for fixed $R > 0$ and $\varepsilon > 0$, cover N by rectangles Q_j with $\sum_j m_n(Q_j) < \varepsilon/(2R)^d$. The product rectangles cover $N \times [-R, R]^d$ and have total volume less than ε. Thus this bounded product is null; taking the union over integer R proves the assertion. This covering argument justifies the product-null-set assertions without presupposing a section theorem.

For completeness, an arbitrary subset of the boundary of a rectangle also has null sections for almost every parameter. A face normal to an x coordinate produces sections contained in an x-hyperplane and hence null. A face normal to a y coordinate can have a nonempty section only when that parameter lies in a fixed y-hyperplane. The union of these finitely many exceptional parameter hyperplanes is null. This also justifies replacing closed, open, and half open faces in a geometric section argument.

Lemma 4.7. *Suppose $G = \bigcap_j O_j$, where O_j are decreasing open sets and $m_{n+d}(O_1) < \infty$. Then $y \mapsto m_n(G^y)$ is measurable, integrable, and*

$$\int m_n(G^y)\, dy = m_{n+d}(G).$$

Proof. By Lemma 4.6, $g_j(y) = m_n(O_j^y)$ is measurable and $g_1 \in L^1(\mathbb{R}^d)$. Thus g_1 is finite outside a null set N. For $y \notin N$, the measurable sections O_j^y decrease to G^y and have finite first measure. Continuity from above gives $g_j(y) \downarrow m_n(G^y)$. The section G^y is Borel for every y, since G is Borel. The pointwise limit of the g_j agrees with its section measure outside N; arbitrary changes on N preserve Lebesgue measurability.

Since $0 \le g_j \le g_1$ a.e., dominated convergence gives $\int m_n(G^y) = \lim_j \int g_j$. The open-set formula identifies the right side with $\lim_j m_{n+d}(O_j) = m_{n+d}(G)$, by continuity from above in the full space. □

Lemma 4.8. *If $Z \subset \mathbb{R}^{n+d}$ is null, then Z^y is measurable and null for almost every $y \in \mathbb{R}^d$.*

Proof. Choose open $U_j \supset Z$ with $m_{n+d}(U_j) < 2^{-j}$ and put $O_j = \bigcap_{l \le j} U_l$. Then $G = \bigcap_j O_j$ is a Borel null set containing Z, and O_1 has finite measure. The preceding lemma gives $\int m_n(G^y)\, dy = 0$. Thus $m_n(G^y) = 0$ outside a measurable null set of parameters. At such a parameter, $Z^y \subset G^y$ is a subset of a null set in $\mathbb{R}^n$, and completeness makes it measurable and null. We make no claim about the remaining parameters. $\square$

Lemma 4.9. *If $A \subset \mathbb{R}^{n+d}$ is measurable and has finite measure, then A^y is measurable for almost every y. After assigning zero on an exceptional null set, the function $y \mapsto m_n(A^y)$ is integrable and its integral is $m_{n+d}(A)$.*

Proof. Choose a G_δ hull $G \supset A$ with $m_{n+d}(G) = m_{n+d}(A)$, formed as an intersection of decreasing open sets of finite first measure. Then $Z = G \setminus A$ is null. Outside the union of the exceptional sets from the preceding two lemmas, G^y is measurable with finite measure, Z^y is null, and

$$A^y = G^y \setminus Z^y, \quad m_n(A^y) = m_n(G^y).$$

Hence the section-measure function agrees a.e. with the integrable section-measure function for G, and has the same integral, namely $m_{n+d}(G) = m_{n+d}(A)$. $\square$

Theorem 4.10 (Fubini). *If $f \in L^1(\mathbb{R}^{n+d})$, then for almost every y, f^y is measurable and integrable on $\mathbb{R}^n$. The function $g(y) = \int_{\mathbb{R}^n} f(x, y)\, dx$, defined to be zero on an exceptional null set, is measurable and integrable, and*

$$\int_{\mathbb{R}^{n+d}} f(x, y)\, d(x, y) = \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^n} f(x, y)\, dx \Big)\, dy. \tag{4.1}$$

The analogous identity holds in the other order.

Proof. First consider a nonnegative simple function whose nonzero level sets have finite measure. Lemma 4.9, applied to its finitely many level sets, shows that its sections are measurable and integrable outside a null set and that the formula holds by finite linearity. Taking the finite union of the exceptional parameter sets makes all these statements simultaneous.

Now let $f \geq 0$ be integrable. Choose simple functions $s_j \uparrow f$ with finite-measure nonzero sets, using dyadic approximation and bounded spatial truncation. For each j, apply the simple case and remove its exceptional parameter set N_j. Outside $N = \bigcup_j N_j$, every section s_j^y is measurable and the section sequence increases pointwise to f^y. Thus f^y is measurable there. Monotone convergence in x gives

$$g_j(y) = \int s_j(x, y)\, dx \uparrow g(y) = \int f(x, y)\, dx \quad (y \notin N).$$

Assign zero to these functions on N. They are measurable functions of y. Monotone convergence in y, the simple identities, and monotone convergence in the full space now give

$$\int g = \lim_j \int g_j = \lim_j \int s_j = \int f < \infty.$$

Consequently g is finite almost everywhere, so f^y is integrable outside a possibly larger null set. This proves the nonnegative integrable case.

For real integrable f, apply that result to f^+, f^-, and $|f|$, and remove the union of the corresponding exceptional sets. The sections of $|f|$ have finite integral there, so subtraction of the two section integrals is legitimate. The resulting function g is measurable and satisfies $|g(y)| \le \int |f(x,y)|\,dx$ a.e. The nonnegative case shows that the right side is integrable. Subtracting the two finite full integrals gives (4.1). Real and imaginary parts prove the complex case. Interchanging the coordinate groups in the geometric argument proves the other order. □

4.3 Tonelli's theorem and changes of variables

Theorem 4.11 (Tonelli). *For nonnegative measurable f on $\mathbb{R}^{n+d}$, its sections are measurable for almost every parameter, the section integral is measurable after a null-set definition, and*

$$\int f = \int \Big(\int f(x,y)\,dx \Big)\,dy = \int \Big(\int f(x,y)\,dy \Big)\,dx.$$

Each quantity is allowed to be infinite.

Proof. Set $f_j = \min(f, j)\mathbf{1}_{[-j,j]^{n+d}}$. These functions are integrable and increase pointwise to f. For each j, Fubini gives measurable sections outside a null parameter set and a measurable section-integral function. Remove the union of these countably many exceptional sets. For every remaining parameter, monotone convergence in x gives $\int f_j(x,y)\,dx \uparrow \int f(x,y)\,dx$. Monotone convergence in y then yields

$$\int \Big(\int f\,dx \Big)\,dy = \lim_j \int \Big(\int f_j\,dx \Big)\,dy = \lim_j \int f_j = \int f.$$

The other order follows in the same manner, now with a common exceptional set in the other parameter space. □

For a measurable signed or complex function, applying Tonelli first to $|f|$ is the appropriate way to verify the integrability hypothesis of Fubini. If the resulting nonnegative iterated integral is finite, then the full absolute integral is finite and Fubini applies. Existence of signed iterated integrals without this absolute estimate is not sufficient.

Example 4.12. Let $I_k = (2^{-k}, 2^{-k+1})$ and let s_k equal 1 on the left half of I_k, -1 on its right half, and zero elsewhere. Define on $(0,1)^2$

$$f(x,y) = \sum_{k\geq 1} m(I_k)^{-2} s_k(x) s_k(y).$$

At each point at most one summand is nonzero. Every fixed row or column meets at most one square $I_k \times I_k$, and the positive and negative halves have equal lengths. Thus both signed iterated integrals exist and equal zero. However, on each square the positive integral and the negative integral are both $1/2$. Tonelli gives $\int f^+ = \int f^- = \infty$, so the signed double integral is undefined. The zero iterated values do not justify subtracting these two infinities.

Proposition 4.13. *If $T : \mathbb{R}^n \to \mathbb{R}^n$ is invertible and linear, then measurable sets remain measurable and*

$$m(T(E)) = |\det T| m(E).$$

For nonnegative measurable f, or for integrable f,

$$\int_{\mathbb{R}^n} f(Tx)\, dx = |\det T|^{-1} \int_{\mathbb{R}^n} f(y)\, dy.$$

Proof. Both T and T^{-1} are Lipschitz, so measurability follows from Proposition 1.19. Gaussian elimination expresses T as a finite composition of coordinate permutations, nonzero coordinate scalings, and shears. For a scaling of one coordinate by $a \neq 0$, the one-dimensional integral scales by $|a|^{-1}$, by Proposition 4.4; Tonelli then gives the same factor for a nonnegative integral in n dimensions. A coordinate permutation leaves the integral unchanged by Tonelli. A shear replacing x_i by $x_i + cx_j$ leaves the integral unchanged: fix the other coordinates and use translation invariance in x_i, then integrate the parameters by Tonelli.

These factors are the absolute determinants of the elementary matrices, inverted. Multiplicativity of the determinant gives the stated formula for their composition. Apply it to $f = \mathbf{1}_{T(E)}$ to obtain the set formula. For integrable f, first apply the nonnegative formula to $|f|$ to establish integrability of $f \circ T$, and then apply it to its positive/negative or real/imaginary parts. □

4.4 Convolution

We first address a measurability issue. If $N \subset \mathbb{R}^n$ is null, then $N \times \mathbb{R}^d$ is null in $\mathbb{R}^{n+d}$. On $N \times [-R, R]^d$, cover N by rectangles with total n-volume less than η. Their products with $[-R, R]^d$ have total $(n+d)$-volume less than $(2R)^d\eta$, so this product is null. Taking a countable union in R proves the assertion.

If f is Lebesgue measurable, choose a Borel representative f_0 equal to f outside a null set N. Then $f_0(x-y)$ is Borel. The exceptional set $\{(x, y) : x-y \in N\}$ is the image of $N \times \mathbb{R}^n$ under the invertible linear map $(z, y) \mapsto (z+y, y)$, so it is null. Therefore $f(x-y)$ is Lebesgue measurable on the product space. This is a null-set argument, not a general assertion about continuous compositions of Lebesgue measurable functions.

Definition 4.14. The convolution is $(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy$ wherever this integral is absolutely convergent. Values may be assigned arbitrarily on an exceptional null set when the integral exists almost everywhere.

Theorem 4.15. *If $f, g \in L^1(\mathbb{R}^n)$, then $f * g$ exists almost everywhere, belongs to L^1, and*

$$\|f * g\|_1 \le \|f\|_1 \|g\|_1, \qquad \int (f * g) = \Big(\int f\Big)\Big(\int g\Big).$$

Also $f * g = g * f$ almost everywhere.

Proof. The preceding argument makes $|f(x-y)||g(y)|$ jointly measurable. Tonelli and translation invariance give

$$\int\int |f(x-y)||g(y)|\,dy\,dx = \int |g(y)|\Big(\int |f(x-y)|\,dx\Big)\,dy = \|f\|_1\|g\|_1 < \infty.$$

Thus the inner absolute integral is finite for almost every x. Fubini makes the convolution measurable, and the pointwise triangle inequality followed by the displayed estimate gives its L^1 bound. Applying Fubini to the signed or complex integrand gives the integral identity. Finally, wherever the defining integral is absolutely convergent, the translation and reflection substitution $z = x - y$ gives

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

This is the expression for $g * f$, and both are defined almost everywhere. $\square$

Exercises

Exercise 4.1. Prove that a Cauchy sequence in L^1 has a subsequence dominated in absolute value by one integrable function.

Exercise 4.2. Give a sequence converging in L^1 that does not converge pointwise a.e. Use the typewriter sequence.

Exercise 4.3. Prove translation continuity in L^1 using step-function approximation instead of continuous-function approximation.

Exercise 4.4. Show that every section of a Borel set is Borel, and explain why completion of the product measure changes this statement.

Exercise 4.5. Give a direct cube-covering proof that a null set in $\mathbb{R}^{n+d}$ has null sections for a.e. parameter, with the exceptional set estimated by a summable sequence of covers.

Exercise 4.6. For nonnegative f, identify its integral with the measure of $\{(x,t) : 0 < t < f(x)\}$.

Exercise 4.7. In Example 4.12, compute the absolute integral on every row and column and the positive integral on each square.

Exercise 4.8. Construct an example where the two signed iterated integrals are finite but unequal.

Exercise 4.9. Prove associativity of L^1 convolution by first establishing absolute integrability on the relevant product space.

Exercise 4.10. If $f \in L^1$ and $g \in L^\infty$, prove $\|f * g\|_\infty \le \|f\|_1 \|g\|_\infty$. Specify how representatives are handled.

Exercise 4.11. If g is bounded and uniformly continuous and $f \in L^1$, prove that $f * g$ has a bounded uniformly continuous representative.

Exercise 4.12. Use elementary shears and coordinate scalings to derive the determinant formula for the matrix $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

Exercise 4.13. If $f_j \to f$ and $g_j \to g$ in L^1, prove $f_j * g_j \to f * g$ in L^1.

Exercise 4.14. Let A, B have finite measure. Show that $\mathbf{1}_A * \mathbf{1}_B$ has integral $m(A)m(B)$, and identify its values as measures of intersections.

Chapter 5

Distribution functions and rearrangements

The integral can be organized by the values of a function rather than by points in its domain. Distribution functions record the measures of level sets. We first establish their continuity properties, with the relevant finiteness assumptions, and then use them to compare functions with different spatial arrangements.

5.1 Distribution functions and level sets

Definition 5.1. For a measurable real function f on E, its distribution function is $\omega_{f,E}(a) = m(\{x \in E : f(x) > a\})$. For a real or complex function on $\mathbb{R}^n$, we also write

$$\mu_f(t) = m(\{|f| > t\}), \quad t > 0.$$

The function $\omega_{f,E}$ is decreasing. Its values can be infinite when $m(E) = \infty$. An identity involving differences of its values requires that these differences be defined.

Proposition 5.2. *Let f be measurable and finite a.e. on a finite-measure set E. Then*

(i) $\omega(a+) = \omega(a)$ *and* $\omega(a-) = m(\{f \geq a\})$*;*

(ii) $\omega(a) - \omega(b) = m(\{a < f \leq b\})$ *for* $a < b$*;*

(iii) $\omega(a-) - \omega(a) = m(\{f = a\})$*;*

(iv) $\omega(a) \to 0$ *as* $a \to \infty$*, and* $\omega(a) \to m(E)$ *as* $a \to -\infty$*.*

Proof. Take $a_j \downarrow a$ with $a_j > a$. The sets $\{f > a_j\}$ increase to $\{f > a\}$. Indeed,

$f(x) > a$ leaves room to choose a threshold a_j below $f(x)$. Continuity from below gives $\omega(a_j) \to \omega(a)$, proving right continuity.

If $a_j \uparrow a$ with $a_j < a$, the sets $\{f > a_j\}$ decrease to $\{f \geq a\}$. Their first measure is at most $m(E) < \infty$, so continuity from above gives the left-limit formula. Since $\{f > b\} \subset \{f > a\}$, the measurable disjoint decomposition

$$\{f > a\} = \{a < f \leq b\} \cup \{f > b\}$$

gives the second assertion by subtraction of finite measures. The third follows by splitting $\{f \geq a\}$ into its strict and equality levels.

Finally, $\{f > j\} \downarrow \{f = \infty\}$, which is null, and finite measure permits continuity from above. The sets $\{f > -j\}$ increase to $E \setminus \{f = -\infty\}$, whose measure is $m(E)$. This proves the two endpoint limits. □

Right continuity remains valid on an infinite-measure domain, as an identity of extended values. The left-limit and large-threshold limits hold as soon as one of the sets in the decreasing family has finite measure. For $f \in L^p$, $p > 0$, Chebyshev applied to $|f|^p$ gives $\mu_f(t) \leq t^{-p}\|f\|_p^p$, so this condition holds at every positive threshold. Merely assuming that f is finite a.e. on an infinite-measure domain is not enough: $f(x) = x$ on $(0, \infty)$ has infinite distribution at every finite threshold.

Corollary 5.3. *Under the finite-measure hypotheses, ω is continuous at a precisely when $m(\{f = a\}) = 0$. It is constant on an open interval (a, b) precisely when $m(\{a < f < b\}) = 0$.*

Proof. The continuity assertion follows from the jump formula. If the distribution is constant on (a, b), then for any $a < c < d < b$, the set $\{c < f \leq d\}$ is null. Take a countable increasing exhaustion of (a, b) by such intervals. Conversely, if f takes values in (a, b) only on a null set, then $\omega(c) - \omega(d) = m(\{c < f \leq d\}) = 0$ for $a < c < d < b$. □

5.2 Integration over the range

Theorem 5.4 (Layer-cake formula)**.** *If $f \geq 0$ is measurable, then*

$$\int f = \int_0^\infty m(\{f > t\})\, dt.$$

For a real or complex measurable function and $0 < p < \infty$,

$$\int |f|^p = p \int_0^\infty t^{p-1} \mu_f(t)\, dt. \tag{5.1}$$

Proof. For each numerical value $s \in [0, \infty]$, one has $s = \int_0^\infty \mathbf{1}_{\{t<s\}}\, dt$. Hence $f(x) = \int_0^\infty \mathbf{1}_{\{t<f(x)\}}\, dt$ pointwise. The integrand is jointly Lebesgue measurable. To justify this without assuming a composition principle that may fail for Lebesgue functions, replace f by a Borel representative. The subgraph of the representative is Borel, since $(x, t) \mapsto f_0(x) - t$ is Borel after treating infinite levels separately. The original subgraph differs only on $N \times (0, \infty)$ for a null set N, and this product is null by Chapter 4.

Tonelli therefore allows the order of integration to be exchanged, giving the first formula. For the second, use the numerical identity $s^p = \int_0^s pt^{p-1}\, dt$ and the same argument. Nonnegativity permits infinity on either side and avoids any cancellation. □

More generally, if $\Phi(0) = 0$ and Φ is continuously differentiable with $\Phi' \geq 0$ on $[0, \infty)$, then Tonelli gives

$$\int \Phi(f) = \int_0^\infty \Phi'(t) m(\{f > t\})\, dt \quad (f \geq 0).$$

For a difference of two such expressions, both relevant integrals must be finite before they are subtracted.

The range formula also has a Riemann–Stieltjes form. We make the endpoint convention explicit. If α is increasing on $[a, b]$ and ψ is continuous, the Riemann–Stieltjes sums are

$$\sum_{i=1}^N \psi(\xi_i)(\alpha(t_i) - \alpha(t_{i-1})), \quad a = t_0 < \cdots < t_N = b, \quad \xi_i \in [t_{i-1}, t_i].$$

Here is the convergence argument. Let $\omega_\psi(\delta) = \sup\{|\psi(s) - \psi(t)| : s, t \in [a, b],\ |s - t| \leq \delta\}$; uniform continuity gives $\omega_\psi(\delta) \to 0$. If a partition P' refines P, split every increment of α in the sum for P into its subincrements over P'. The old tag and each new tag belonging to that interval differ by at most the mesh $\|P\|$. Since the subincrements of α are nonnegative and sum to $\alpha(b) - \alpha(a)$,

$$|S(P, \xi) - S(P', \xi')| \leq \omega_\psi(\|P\|)\big(\alpha(b) - \alpha(a)\big).$$

For two arbitrary partitions P, Q, compare both with their common refinement. If both meshes are below δ, their sums differ by at most $2\omega_\psi(\delta)(\alpha(b) - \alpha(a))$. This tends to zero independently of all tags. Completeness of $\mathbb{R}$, or of $\mathbb{C}$ for complex ψ, gives a common limit, denoted $\int_a^b \psi\, d\alpha$. If α is constant, every increment is zero and the conclusion is immediate. For a decreasing integrator use a minus sign. With a right continuous integrator, these increments count its jump at b and do not count a jump at a coming from the left; thus the convention corresponds to the range interval $(a, b]$.

Proposition 5.5. *Let f be finite measurable on a finite-measure set E, let $a < b$, and put $E_{a,b} = \{a < f \le b\}$. Then*

$$\int_{E_{a,b}} f = -\int_a^b t \, d\omega_{f,E}(t).$$

If $a < f \le b$ everywhere on E, this is a formula for $\int_E f$.

Proof. For a partition $a = t_0 < \cdots < t_N = b$, let $A_i = \{t_{i-1} < f \le t_i\}$. These are disjoint and have union $E_{a,b}$. Monotonicity and finite additivity of the integral give

$$\sum_i t_{i-1} m(A_i) \le \int_{E_{a,b}} f \le \sum_i t_i m(A_i).$$

The difference of the bounds is at most $\max_i (t_i - t_{i-1}) m(E_{a,b})$, which tends to zero with the mesh. By Proposition 5.2, $m(A_i) = \omega(t_{i-1}) - \omega(t_i)$. Thus the two bounds are the lower and upper tagged Stieltjes sums for the continuous function t and the increasing integrator $-\omega$. Their common limit proves the identity. □

For a bounded real function, choose a strictly below its lower bound and b at least its upper bound. This avoids losing the level set $\{f = a\}$. For example, a constant function $f = c$ on a set of measure M has a distribution with a jump of size M at c, and the range formula gives cM when $a < c \le b$. On an infinite-measure domain, a signed distribution may be infinite at all negative thresholds, so a global Stieltjes formula must not be written by subtracting infinite distribution values. The nonnegative layer-cake formulas remain valid without that restriction.

The unbounded range can also be treated without concealing a convergence assumption. Suppose $m(E) < \infty$ and f is finite and measurable. Apply the same tagged-sum proof to the continuous range function $|t|$: its error on an interval is at most the length of that interval. It gives

$$\int_{\{a<f\le b\}} |f| = -\int_a^b |t| \, d\omega(t).$$

For $a = -N$ and $b = N$, the sets $\{-N < f \le N\}$ increase to E. Monotone convergence shows that f is integrable exactly when

$$\sup_N \Big(-\int_{-N}^N |t| \, d\omega(t) \Big) < \infty.$$

Under this condition, $|f|$ dominates every signed truncation and dominated convergence gives

$$\int_E f = \lim_{N\to\infty} \int_{\{-N<f\le N\}} f = -\lim_{N\to\infty} \int_{-N}^{N} t\, d\omega(t).$$

The same limit is obtained by letting the lower endpoint tend independently to $-\infty$ and the upper endpoint to ∞, since the omitted absolute integral tends to zero. A merely conditionally convergent signed range limit would not establish Lebesgue integrability; the absolute-integral condition is essential.

Definition 5.6. Two nonnegative measurable functions are equimeasurable if their strict positive superlevel sets have equal measures at every positive threshold.

Corollary 5.7. *Equimeasurable nonnegative functions have equal integrals and equal L^p quantities for every $0 < p < \infty$. They also have the same essential supremum.*

Proof. The integral assertions follow from Theorem 5.4. The essential supremum is the infimum of the thresholds above which the distribution is zero, so it also depends only on that distribution. □

For finite real functions on finite-measure domains of equal measure, equality of the distributions at every real threshold also gives equality of the distributions of the negative parts. Indeed, the measure of $\{f < -t\}$ is the domain measure minus $m(\{f \ge -t\})$, and the latter is a left limit of the distribution. Thus signed integrability and the signed integral are preserved as well.

5.3 Symmetric decreasing rearrangement

Definition 5.8. For a measurable set A of finite measure, let A^* be the open ball centred at zero with measure $m(A)$; if $m(A) = 0$, take $A^* = \varnothing$. Let $f \ge 0$ be measurable on $\mathbb{R}^n$ with $m(\{f > t\}) < \infty$ for every $t > 0$. Define

$$f^*(x) = \int_0^\infty \mathbf{1}_{\{f>t\}^*}(x)\, dt.$$

For real or complex f, its rearrangement means the rearrangement of $|f|$.

The superlevel sets used here are nested as t increases. Their centred balls are nested in the same order. The finiteness condition on positive levels is sometimes called vanishing at infinity in measure; it does not require pointwise decay along every sequence tending to infinity.

Theorem 5.9. *The function f^* is nonnegative, radial, and decreasing in the radius. It is lower semicontinuous and measurable, and for every $t > 0$,*

$$\{x : f^*(x) > t\} = \{f > t\}^*. \tag{5.2}$$

Consequently it is equimeasurable with f, has the same L^p quantities, and preserves order: $f \le g$ a.e. implies $f^* \le g^*$.

Proof. Put $\mu(t) = m(\{f > t\})$ and $v(x) = \omega_n|x|^n$. The condition $x \in \{f > s\}^*$ is exactly $v(x) < \mu(s)$. For fixed x, the set of $s > 0$ satisfying this condition is an initial interval, possibly empty or unbounded, because μ is decreasing. Its length is $f^*(x)$. This immediately gives nonnegativity, radial symmetry, and decrease in $|x|$.

We check (5.2). If $f^*(x) > t$, the initial interval has length greater than t, so some $s > t$ satisfies $v(x) < \mu(s) \le \mu(t)$. Hence $x \in \{f > t\}^*$. Conversely, suppose $v(x) < \mu(t)$. Right continuity of μ gives some $s > t$ sufficiently close to t with $v(x) < \mu(s)$. The initial interval then contains $(0, s)$, so $f^*(x) \ge s > t$. This proves equality, including values at the boundaries of the centred balls.

Every strict positive superlevel set of f^* is open. Since $f^* \ge 0$, its strict superlevel sets at nonpositive thresholds are either all of $\mathbb{R}^n$ or the union of the positive superlevel sets. They are open as well. This is lower semicontinuity, and hence Borel measurability. The level-set equality gives equimeasurability and Corollary 5.7 gives the norm assertions.

Finally, if $f \le g$ a.e., then $m(\{f > t\}) \le m(\{g > t\})$ for every $t > 0$. The corresponding centred balls are contained in one another. Integrating their indicators in t gives $f^* \le g^*$. □

Theorem 5.10 (Rearrangement inequality)**.** *For nonnegative measurable f, g whose positive levels have finite measure,*

$$\int_{\mathbb{R}^n} f(x)g(x)\,dx \le \int_{\mathbb{R}^n} f^*(x)g^*(x)\,dx.$$

Proof. For numerical nonnegative values, the product equals the double integral of the two level indicators. Tonelli gives

$$\int fg = \int_0^\infty \int_0^\infty m(\{f > s\} \cap \{g > t\})\,ds\,dt.$$

The measure of an intersection is at most the smaller measure, so the integrand is at most $\min(\mu_f(s), \mu_g(t))$. The two rearranged superlevel sets are centred balls; one contains the other, and their intersection has exactly this minimum measure. Applying the same Tonelli identity to f^*, g^* proves the inequality. All integrands are nonnegative, so the argument also covers infinite integrals. □

5.4 Weak integral bounds

Definition 5.11. For $0 < p < \infty$, define the weak L^p quantity

$$\|f\|_{p,\infty} = \sup_{t>0} t\, \mu_f(t)^{1/p}.$$

A finite value means $\mu_f(t) \le (\|f\|_{p,\infty}/t)^p$ at every threshold.

Chebyshev gives $\|f\|_{p,\infty} \le \|f\|_p$ for $f \in L^p$. The converse is false. The function $f(x) = |x|^{-n/p}$ on $\mathbb{R}^n$ has $\mu_f(t) = \omega_n t^{-p}$ and finite weak L^p quantity, but (5.1) makes its pth power integral infinite.

Proposition 5.12. *On a set E of finite measure, weak L^q is contained in L^p whenever $0 < p < q < \infty$.*

Proof. Let $M = m(E)$ and $A = \|f\|_{q,\infty}$. If $M = 0$ or $A = 0$, the assertion is immediate. Otherwise $\mu_f(t) \le \min(M, A^q t^{-q})$. Split the layer integral at $t_0 = AM^{-1/q}$:

$$\begin{aligned}\int_E |f|^p &\le pM \int_0^{t_0} t^{p-1}\, dt + pA^q \int_{t_0}^{\infty} t^{p-q-1}\, dt \\ &= \frac{q}{q-p} A^p M^{1-p/q} < \infty.\end{aligned}$$

The tail integral is finite exactly because $p < q$. □

Exercises

Exercise 5.1. Prove the right- and left-continuity statements for distribution functions under the weakest finite-level hypotheses needed in each case.

Exercise 5.2. Compute the distributions of $x^{-\alpha}$ on $(0,1)$ and on $(1,\infty)$, and recover all their L^p ranges.

Exercise 5.3. Show that a finite measurable function on a finite-measure set has at most countably many levels of positive measure.

Exercise 5.4. Derive the formula for $\int_{\{f>t\}} f$ in terms of the distribution of a nonnegative function, including the boundary term $t\mu_f(t)$.

Exercise 5.5. Verify the Riemann–Stieltjes range formula for a function with finitely many values, paying attention to atoms at the two endpoints.

Exercise 5.6. Prove that equimeasurability preserves $\int \Phi(f)$ for every increasing continuously differentiable Φ with $\Phi(0) = 0$ and $\Phi' \geq 0$.

Exercise 5.7. Show that the rearrangement of an indicator is the indicator of its centred ball, apart from the harmless convention on the boundary.

Exercise 5.8. Prove that the rearrangement of a radial decreasing lower semicontinuous nonnegative function equals the original function under the strict-superlevel convention.

Exercise 5.9. For $m(A) < \infty$, prove $\int_A f \leq \int_{A^*} f^*$ for nonnegative f with finite positive levels.

Exercise 5.10. Show that weak L^p is closed under addition, with an explicit quasi-triangle bound obtained from level sets.

Exercise 5.11. Give a function in weak $L^1(0,1)$ that is not integrable, and a function in every $L^p(0,1)$, $p < \infty$, that is unbounded.

Exercise 5.12. Prove that $\|f\|_{p,\infty} = \|f^*\|_{p,\infty}$.

Exercise 5.13. Show that a finite L^p norm forces $t^p \mu_f(t) \to 0$ both as $t \to \infty$ and as $t \downarrow 0$.

Exercise 5.14. Prove that rearrangement preserves the measure of each positive equality level by taking jumps of the distribution.

Chapter 6

L^p spaces

The spaces L^p distinguish the size of a function from the measure of the region on which it is large. The finite-measure hypothesis produces inclusions between these spaces, but on the whole space neither inclusion generally holds. We keep the endpoint cases separate whenever the proof changes.

6.1 Norms and integral inequalities

Definition 6.1. For $0 < p < \infty$, let $L^p(E)$ consist of measurable functions, modulo equality a.e., with $\|f\|_p = (\int_E |f|^p)^{1/p} < \infty$. Define $\|f\|_\infty = \operatorname{ess\,sup}_E |f|$, the infimum of $M \geq 0$ for which $|f| \leq M$ a.e. The space $L^\infty(E)$ consists of the classes with finite essential supremum.

If the infimum in the essential supremum is finite, the bound holds with that infimum. Indeed, $|f| \leq \|f\|_\infty + 1/k$ outside a null set for every k; remove their countable union and let $k \to \infty$. In this chapter the normed-space assertions concern $p \geq 1$. For $0 < p < 1$, the displayed quantity need not satisfy the triangle inequality.

For an open set $\Omega \subset \mathbb{R}^n$, the notation $f \in L^p_{\mathrm{loc}}(\Omega)$ means that f is measurable and belongs to $L^p(K)$ on every compact $K \subset \Omega$. For $p = \infty$ this means essential boundedness on each such compact set. When $1 \leq p \leq \infty$, local L^p membership implies local L^1 membership by the finite-measure inequality proved below. Statements involving L^p norms for $0 < p < 1$ will be called quantities rather than norms. A Banach space means a complete normed vector space; separable means that it has a countable dense subset.

Lemma 6.2 (Young's numerical inequality). *For $1 < p < \infty$ and $p' = p/(p-1)$,*

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'} \quad (a, b \geq 0).$$

Equality holds exactly when $b = a^{p-1}$.

Proof. A useful general form comes from a continuous strictly increasing unbounded function $\phi : [0, \infty) \to [0, \infty)$ with $\phi(0) = 0$, and its inverse ψ. In the rectangle $[0, a] \times [0, \phi(a)]$, the regions below $y = \phi(x)$ and to the left of $x = \psi(y)$ divide its area, up to the graph. Tonelli, or integration of the one-dimensional sections, gives

$$\int_0^a \phi(x)\, dx + \int_0^{\phi(a)} \psi(y)\, dy = a\phi(a).$$

If $b \geq \phi(a)$, subtract ab after adding the remaining inverse integral; the difference is $\int_{\phi(a)}^{b} (\psi(y) - a)\, dy \geq 0$. If $b < \phi(a)$, it is $\int_b^{\phi(a)} (a - \psi(y))\, dy \geq 0$. Strict increase makes the difference positive unless $b = \phi(a)$.

Take $\phi(x) = x^{p-1}$ and $\psi(y) = y^{p'-1}$. Their integrals give the asserted powers and denominators. □

Theorem 6.3 (Hölder). *For conjugate exponents $1 \leq p, p' \leq \infty$,*

$$\int_E |fg| \leq \|f\|_p \|g\|_{p'}.$$

Proof. If either norm is zero, the corresponding function vanishes a.e., so the product is zero a.e. When $p = 1$, the bound $|g| \leq \|g\|_\infty$ a.e. gives the assertion by integration. The case $p = \infty$ is the same with the factors interchanged.

For $1 < p < \infty$ and positive finite norms, set $F = |f|/\|f\|_p$ and $G = |g|/\|g\|_{p'}$. Young's numerical inequality gives $FG \leq F^p/p + G^{p'}/p'$. Integrating yields $\int FG \leq 1/p + 1/p' = 1$. Multiplication by the two original norms proves the result. If a norm is infinite and the other is positive, the extended upper bound is automatic; in applications requiring a finite integral, both specified norms are finite. □

The case $p = 2$ is the Cauchy–Schwarz inequality. We will need its multiple-factor version with all exponent choices explicit. If $1 \leq p_j \leq \infty$ and $\sum_{j=1}^{k} 1/p_j = 1$, then

$$\int \prod_{j=1}^{k} |f_j| \leq \prod_{j=1}^{k} \|f_j\|_{p_j}.$$

Assume the norms on the right are finite; a zero norm makes the assertion immediate. Factors with exponent ∞ can first be bounded by their essential suprema and removed. If only one finite exponent remains, it equals one and the assertion is just integration of that last factor. For two finite exponents it is Hölder. Inductively, when at least three finite exponents remain, set $r = (\sum_{j=1}^{k-1} 1/p_j)^{-1}$. Then $1/r + 1/p_k = 1$ and $p_j/r \geq 1$ for $j < k$. The induction hypothesis applied to $|f_j|^r$, whose exponents p_j/r have reciprocal sum one, gives

$$\int \prod_{j=1}^{k-1} |f_j|^r \leq \prod_{j=1}^{k-1} \Big(\int |f_j|^{p_j} \Big)^{r/p_j}.$$

Taking the rth root bounds $\| \prod_{j<k} f_j \|_r$. Hölder with this product and f_k proves the desired inequality. Thus later applications to three or more factors do not require an additional inequality to be assumed.

Theorem 6.4 (Minkowski)**.** *For* $1 \leq p \leq \infty$, $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

Proof. The cases $p = 1$ and $p = \infty$ follow directly from the pointwise triangle inequality and integration or essential bounds. Assume $1 < p < \infty$. The convexity inequality $(a + b)^p \leq 2^{p-1}(a^p + b^p)$ first shows that $f + g \in L^p$, so the quantities below are finite. If $\|f + g\|_p = 0$, the assertion is immediate. Otherwise, Hölder gives

$$\begin{aligned} \|f + g\|_p^p &\leq \int |f + g|^{p-1}|f| + \int |f + g|^{p-1}|g| \\ &\leq \|f + g\|_p^{p-1}(\|f\|_p + \|g\|_p), \end{aligned}$$

since $(p - 1)p' = p$. Divide by the positive finite factor $\|f + g\|_p^{p-1}$. □

Proposition 6.5. *If* $m(E) < \infty$ *and* $1 \leq p < q \leq \infty$, *then*

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

Also $\lim_{p\to\infty}\|f\|_p = \|f\|_\infty$ *for every measurable* f *on a finite-measure set, with extended values allowed.*

Proof. A null set E gives zero on both sides, so assume $0 < m(E) < \infty$. If $q < \infty$, Hölder applied to $|f|^p \cdot 1$ with exponents q/p and $q/(q - p)$ gives

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

Taking the pth root yields the stated power $m(E)^{1/p-1/q}$. If $q = \infty$, the almost everywhere bound $|f|^p \leq \|f\|_\infty^p$ gives $\int_E |f|^p \leq m(E)\|f\|_\infty^p$ and the same formula. An infinite right side is interpreted only as an automatic upper bound.

For the limit, a null domain gives zero norms and can be omitted. If $M = \|f\|_\infty < \infty$, the upper estimate $\|f\|_p \le Mm(E)^{1/p}$ gives limsup at most M. For any $0 < a < M$, the set $A = \{|f| > a\}$ has positive finite measure, and $\|f\|_p \ge am(A)^{1/p}$. Taking liminf gives a bound at least a, and then $a \uparrow M$ proves equality. If $M = \infty$, the same lower bound holds for every finite $a > 0$, so the norms tend to infinity. □

On infinite-measure domains, the inclusion generally fails in both directions. For $p < q < \infty$, choose $1/q < \alpha < 1/p$. Then $x^{-\alpha}$ lies in $L^p(0,1)$ but not $L^q(0,1)$, while on $(1,\infty)$ it lies in L^q but not L^p. Extend these functions by zero to view them on $\mathbb{R}$. The constant one belongs to $L^\infty(\mathbb{R})$ but to no finite $L^p(\mathbb{R})$.

Proposition 6.6 (Dual norm formula)**.** *For $1 \le p \le \infty$ and a finite a.e. measurable f on a Lebesgue measurable set E,*

$$\|f\|_p = \sup\Big\{\Big|\int_E f\overline{g}\Big| : \|g\|_{p'} \le 1,\ f\overline{g} \in L^1\Big\}.$$

The norm on the left may be infinite.

Proof. Hölder gives the upper bound when $\|f\|_p$ is finite. If $\|f\|_p = 0$, then $f = 0$ almost everywhere and every admissible pairing is zero, proving the formula in that case. For $1 < p < \infty$ and $0 < \|f\|_p < \infty$, set

$$g = \|f\|_p^{1-p}|f|^{p-2}f,$$

with zero at the zeros of f. Since $(p-1)p' = p$, one has $\|g\|_{p'} = 1$ and $\int f\overline{g} = \|f\|_p$. For $p = 1$, take $g = f/|f|$ off the zero set and zero on it; its essential bound is at most one and the pairing is $\int |f|$.

To include infinite norms at finite p, let $A_N = E \cap [-N,N]^n \cap \{|f| \le N\}$ and $f_N = f\mathbf{1}_{A_N}$. The nonnegative functions $|f_N|^p$ increase to $|f|^p$ a.e., so $\|f_N\|_p \uparrow \|f\|_p$. Apply the preceding maximizing test to f_N, extending the test by zero outside A_N. Its pairing with f equals its pairing with f_N and is finite. The supremum is at least every $\|f_N\|_p$.

For $p = \infty$, let $0 < a < \|f\|_\infty$. The set $\{x \in E : |f(x)| > a\}$ has positive measure. Its subsets $A_N = E \cap [-N,N]^n \cap \{a < |f| \le N\}$ increase to it except for the null infinite-value set. Continuity from below therefore gives some N with $0 < m(A_N) < \infty$. On $A = A_N$, the function $|f|$ is bounded above by N. Define $g = (f/|f|)\mathbf{1}_A/m(A)$. Then $\|g\|_1 = 1$, the pairing is finite, and $\int f\overline{g} = m(A)^{-1}\int_A |f| > a$. Let a tend to the essential supremum, also when it is infinite. □

This is a norm characterization, not yet a description of every bounded linear functional. The latter statement requires measure-theoretic representation and will be proved in Chapter 12.

6.2 Completeness, density, and endpoints

Theorem 6.7. *For $1 \le p \le \infty$, $L^p(E)$ is a Banach space.*

Proof. The norm properties follow from Minkowski, homogeneity, and equality modulo null sets. Suppose first $p < \infty$ and (f_j) is Cauchy. For each k choose N_k so that $\|f_j - f_l\|_p \le 2^{-k}$ whenever $j, l \ge N_k$. Choose strictly increasing $j_k \ge \max(N_1, \ldots, N_k)$. Then $d_k = f_{j_{k+1}} - f_{j_k}$ satisfies $\|d_k\|_p \le 2^{-k}$. Select measurable representatives and set $g_N = |f_{j_1}| + \sum_{k=1}^N |d_k|$. Minkowski gives $\|g_N\|_p \le \|f_{j_1}\|_p + 1 =: C$. Since $g_N \uparrow g$, monotone convergence applied to g_N^p yields $\int g^p \le C^p$. Thus g is finite outside a measurable null set.

At every point of its finite set, the series $f_{j_1} + \sum_{k\ge1} d_k$ converges absolutely. Let f be this pointwise sum there, and set $f = 0$ on the exceptional set. It is measurable, $|f| \le g$ almost everywhere, and $f \in L^p$. For fixed K, the finite tails satisfy

$$\Big\|\sum_{k=K}^N d_k\Big\|_p \le \sum_{k=K}^N \|d_k\|_p \le \sum_{k=K}^N 2^{-k}.$$

Their pointwise limit is $f - f_{j_K}$ outside the same null set. Fatou applied to the pth power therefore gives $\|f - f_{j_K}\|_p^p \le (\sum_{k\ge K} 2^{-k})^p$. Taking a pth root proves subsequence convergence.

For the full sequence, given $\varepsilon > 0$ choose J with $\|f_j - f_l\|_p < \varepsilon/2$ for $j, l \ge J$, then choose K with $j_K \ge J$ and $\|f_{j_K} - f\|_p < \varepsilon/2$. For every $j \ge J$, $\|f_j - f\|_p \le \|f_j - f_{j_K}\|_p + \|f_{j_K} - f\|_p < \varepsilon$.

For $p = \infty$, choose the same subsequence using the essential supremum norm. The inequalities $|d_k| \le 2^{-k}$ and $|f_{j_1}| \le \|f_{j_1}\|_\infty$ hold outside a countable union of measurable null sets. On its complement, every tail of the telescoping series is bounded by the numerical tail $\sum_{k\ge K} 2^{-k}$, so the convergence is uniform. Define its bounded measurable limit f there and set it to zero on the exceptional set. Then $\|f - f_{j_K}\|_\infty \le \sum_{k\ge K} 2^{-k} \to 0$. The same $\varepsilon/2$ comparison with a fixed subsequence term proves convergence of the whole sequence. □

Proposition 6.8. *For $1 \le p < \infty$, bounded simple functions with finite-measure nonzero sets and $C_c(\mathbb{R}^n)$ are dense in $L^p(\mathbb{R}^n)$. Translations satisfy $\|\tau_h f - f\|_p \to 0$ as $h \to 0$.*

Proof. Truncate f in location and magnitude. The pth power of the error is bounded by $|f|^p$ and tends pointwise to zero, so dominated convergence gives an L^p approximation. Uniform simple approximation on the remaining bounded set then gives an arbitrarily small L^p error.

For an indicator of a finite-measure set, the continuous cutoff construction in Proposition 4.3 has $0 \le \phi \le 1$, equals one on a compact inner set, and vanishes outside a nearby bounded open set. Therefore $|\mathbf{1}_A - \phi|^p \le \mathbf{1}_{A\setminus K} + \mathbf{1}_{O\setminus K}$, giving arbitrarily small L^p error. Finite summation and Minkowski give continuous-function density.

For $\phi \in C_c$, uniform continuity and one common bounded support for small translations give $\|\tau_h\phi - \phi\|_p \to 0$. Translation invariance of $\int |f|^p$ makes translations isometries. The estimate $\|\tau_h f - f\|_p \le 2\|f - \phi\|_p + \|\tau_h\phi - \phi\|_p$ completes the proof. □

For $p < \infty$, the separability assertion can be seen from a concrete approximation. Start with $s = \sum_{l=1}^{L} a_l \mathbf{1}_{A_l}$ with all $m(A_l) < \infty$. The rational-rectangle approximation established after Theorem 1.15 gives finite unions R_l of bounded rational rectangles such that $m(A_l \triangle R_l)^{1/p}$ is as small as prescribed. Since $\|\mathbf{1}_{A_l} - \mathbf{1}_{R_l}\|_p = m(A_l \triangle R_l)^{1/p}$, Minkowski controls the total error by $\sum_l |a_l| m(A_l \triangle R_l)^{1/p}$. Inclusion–exclusion, as in Proposition 4.3, expresses each $\mathbf{1}_{R_l}$ as a finite sum of indicators of rational rectangles. Finally replace each coefficient a by a number $b \in \mathbb{Q} + i\mathbb{Q}$ so close that $|a - b| m(R)^{1/p}$ is below the chosen error for that term. The resulting finite rational sums form a countable family: there are countably many finite choices of rational endpoints and rational coefficients. The three approximations make this family dense. Thus $L^p(\mathbb{R}^n)$ is separable. Restricting the family to a fixed measurable E gives a dense family in $L^p(E)$, since every function on E can be extended by zero. The space $L^\infty(0,1)$ is not separable: the uncountable family $\mathbf{1}_{(0,t)}$, $0 < t < 1$, has pairwise distance one. A countable dense set would have to approximate each member within $1/3$, but one centre cannot approximate two different members. Translation continuity also fails in $L^\infty(\mathbb{R})$ for an interval indicator: every sufficiently small nonzero translation changes its value by one on a set of positive measure.

The space L^∞ is an algebra with $\|fg\|_\infty \le \|f\|_\infty \|g\|_\infty$. For finite p, $L^p(0,1)$ is not closed under multiplication: $f(x) = x^{-1/(2p)}$ belongs to L^p, while f^2 does not.

6.3 Interpolation and convolution estimates

Proposition 6.9. *Let $1 \le p < r < q \le \infty$ and $1/r = \theta/p + (1-\theta)/q$, with $0 < \theta < 1$. Then*
$$\|f\|_r \le \|f\|_p^{\theta}\|f\|_q^{1-\theta}.$$

Proof. For $q < \infty$, write $|f|^r = |f|^{\theta r}|f|^{(1-\theta)r}$ and apply Hölder with exponents $p/(\theta r)$ and $q/((1-\theta)r)$. Their reciprocals sum to one by the assumed identity. Taking the rth root gives the result. For $q = \infty$, use $|f|^{(1-\theta)r} \le \|f\|_\infty^{(1-\theta)r}$ and note that $\theta r = p$. □

Theorem 6.10 (Young's convolution inequality). *If $1 \le p, q, r \le \infty$ and $1 + 1/r = 1/p + 1/q$, then*
$$\|f * g\|_r \le \|f\|_p\|g\|_q.$$
The defining convolution integral is absolutely convergent almost everywhere.

Proof. Assume $f \in L^p$ and $g \in L^q$, as required for a finite bound. If either is zero almost everywhere, the convolution vanishes almost everywhere, by the product-null-set argument in Chapter 4. We may therefore suppose both norms are positive. Assume first $r < \infty$. The exponent relation implies $p, q \le r$, and in particular $p, q < \infty$. For fixed x, factor the absolute integrand as
$$|f(x-y)||g(y)| = \left(|f(x-y)|^p|g(y)|^q\right)^{1/r}|f(x-y)|^{1-p/r}|g(y)|^{1-q/r}.$$
Apply generalized Hölder with reciprocal exponents $1/r$, $1/p - 1/r$, and $1/q - 1/r$. These are nonnegative and sum to one. A zero reciprocal means an L^∞ factor equal to one, so the cases $p = r$ or $q = r$ are included. The result is
$$\left(\int |f(x-y)||g(y)|\,dy\right)^r \le \|f\|_p^{r-p}\|g\|_q^{r-q}\int |f(x-y)|^p|g(y)|^q\,dy.$$
For $r = 1$, necessarily $p = q = 1$, and this is just the original nonnegative integral estimate. If a norm vanishes, the convolution is zero a.e.; otherwise the formula has no ambiguous zero powers.

Tonelli and translation invariance show that the integral of the last inner integral in x is $\|f\|_p^p\|g\|_q^q$. It is finite, so the absolute convolution is finite a.e. Integrating the bound and taking the rth root gives the assertion.

If $r = \infty$, the exponents p, q are conjugate. Hölder, for each fixed x, bounds the absolute convolution by $\|f\|_p\|g\|_q$, using translation and reflection invariance of the L^p norm. This includes $(p, q) = (1, \infty)$ and $(\infty, 1)$. Measurability follows from the joint-measurability argument in Chapter 4 and measurable truncated integrals. □

6.4 Uniform integrability and convergence

Definition 6.11. A family $\mathcal{F} \subset L^1(E)$ is uniformly integrable if

$$\lim_{M\to\infty} \sup_{f\in\mathcal{F}} \int_{\{|f|>M\}} |f| = 0.$$

In this section, $m(E) < \infty$.

Proposition 6.12. *On a finite-measure set, uniform integrability is equivalent to L^1 boundedness together with uniform absolute continuity of the integrals: for every $\varepsilon > 0$ there is $\delta > 0$ such that $\int_A |f| < \varepsilon$ for every $f \in \mathcal{F}$ whenever $m(A) < \delta$.*

Proof. From uniform integrability choose M making all tails at most one. Then $\|f\|_1 \le Mm(E) + 1$, proving boundedness. For a prescribed ε, choose M making all tails less than $\varepsilon/2$, and take $\delta = \varepsilon/(2M)$, increasing M to at least one. Splitting at M gives $\int_A |f| < \varepsilon/2 + Mm(A) < \varepsilon$ uniformly.

Conversely, let $C = \sup_{f\in\mathcal{F}} \|f\|_1 < \infty$. Given ε, choose δ from uniform absolute continuity. Chebyshev gives $m(\{|f| > M\}) \le C/M < \delta$ for all f when $M > C/\delta$. Apply the uniform small-set estimate to these level sets to obtain the tail condition. □

Theorem 6.13 (Vitali convergence)**.** *On a finite-measure set, $f_j \to f$ in L^1 if and only if $f_j \to f$ in measure and (f_j) is uniformly integrable.*

Proof. If $f_j \to f$ in L^1, Chebyshev gives convergence in measure and the norms are bounded. For uniform absolute continuity, fix $\varepsilon > 0$. Choose J so that $\|f_j - f\|_1 < \varepsilon/2$ for $j \ge J$. Choose one $\delta > 0$ making the integrals of $|f|, |f_1|, \dots, |f_{J-1}|$ on every set of measure less than δ smaller than $\varepsilon/2$. This is possible by absolute continuity for finitely many integrable functions. For $j \ge J$, $\int_A |f_j| \le \|f_j - f\|_1 + \int_A |f| < \varepsilon$. Proposition 6.12 gives uniform integrability.

Conversely, choose an a.e. convergent subsequence by Theorem 2.16. Fatou and the uniform norm bound give $f \in L^1$. Moreover, uniform small-set bounds for (f_j) pass to f by Fatou along that subsequence, so the enlarged family including f has uniformly absolutely continuous integrals.

Fix $\varepsilon > 0$ and choose $\delta > 0$ making the integral of each $|f_j|$ and of $|f|$ on a set of measure below δ less than $\varepsilon/3$. Choose $\eta > 0$ with $\eta m(E) < \varepsilon/3$. Convergence in measure gives $m(A_j) < \delta$ for large j, where $A_j = \{|f_j - f| > \eta\}$. Then

$$\int_E |f_j - f| \le \eta m(E) + \int_{A_j} |f_j| + \int_{A_j} |f| < \varepsilon.$$

□

Proposition 6.14. *A family on a finite-measure set is uniformly integrable if and only if there is a nonnegative increasing convex function* Φ *on* $[0,\infty)$ *with* $\Phi(t)/t \to \infty$ *and* $\sup_{f\in\mathcal{F}} \int \Phi(|f|) < \infty$.

Proof. Under uniform integrability choose increasing $a_k \to \infty$ such that $\sup_f \int_{\{|f|>a_k\}} |f| \le 2^{-k}$. Define $\Phi(t) = \sum_{k\ge 1}(t-a_k)_+$. At every finite t only finitely many terms are nonzero, so this is a finite increasing convex function. For each fixed N, $\liminf_{t\to\infty} \Phi(t)/t \ge N$, hence the ratio tends to infinity. Tonelli and $(t-a_k)_+ \le t\mathbf{1}_{\{t>a_k\}}$ give $\sup_f \int \Phi(|f|) \le \sum_k 2^{-k} = 1$.

Conversely, put $C = \sup_f \int \Phi(|f|)$. For every M large enough that $\Phi(t) > 0$ for $t > M$,

$$\int_{\{|f|>M\}} |f| \le C \sup_{t>M} \frac{t}{\Phi(t)}.$$

The last supremum tends to zero because $\Phi(t)/t \to \infty$. This is uniform integrability. □

On infinite-measure domains, control of large values does not prevent mass from escaping to infinity. The translates $\mathbf{1}_{(j,j+1)}$ have uniformly bounded values and satisfy the tail definition, but do not converge to zero in L^1. Additional spatial tightness is needed for a corresponding global convergence statement.

Exercises

Exercise 6.1. Determine all exponents for which $|x|^{-\alpha}$ is integrable to the pth power near zero or infinity in $\mathbb{R}^n$. Use level sets rather than an unproved polar-coordinate formula.

Exercise 6.2. Characterize equality in Hölder for nonnegative functions when $1 < p < \infty$.

Exercise 6.3. Prove the generalized Hölder inequality for finitely many factors by induction, including infinite exponents.

Exercise 6.4. Give a direct proof that L^∞ is complete, starting from a Cauchy sequence rather than a summable subsequence.

Exercise 6.5. Show that L^p convergence for finite p gives convergence in measure and an a.e. convergent subsequence.

Exercise 6.6. Prove that $L^\infty(0,1)$ is not separable using interval indicators.

Exercise 6.7. For $0 < p < 1$, show that $\int |f-g|^p$ defines a complete metric on L^p classes even though $\|\cdot\|_p$ is not a norm.

Exercise 6.8. Determine when the supremum in the L^∞ dual norm formula is attained by an L^1 function.

Exercise 6.9. Prove the interpolation estimate for a sequence whose norms at two exponents are respectively convergent to zero and uniformly bounded.

Exercise 6.10. Verify all endpoint cases of Young's convolution inequality separately.

Exercise 6.11. On a finite-measure set, prove that an L^p bounded family, $p > 1$, is uniformly integrable.

Exercise 6.12. Give examples showing that L^1 boundedness alone does not imply uniform integrability and that uniform integrability alone does not imply convergence in measure.

Exercise 6.13. State and prove a version of Vitali convergence on $\mathbb{R}^n$ with uniform control of the tails outside large balls.

Exercise 6.14. If $f_j \to f$ a.e. and $\|f_j\|_p \to \|f\|_p$ for $1 < p < \infty$, prove L^p convergence under an additional common integrable majorant for $|f_j|^p$. Compare with the corresponding statement without that majorant.

Chapter 7

Maximal functions and differentiation

An integral averages a function over sets. The differentiation problem asks whether sufficiently small averages recover the value of the function. The main estimate is obtained from a finite covering argument. Approximation by continuous functions then transfers a pointwise fact for continuous functions to an almost everywhere fact for integrable functions.

7.1 A covering lemma and the maximal function

Lemma 7.1 (Finite covering lemma)**.** *From a finite family of open balls $B_1, \dots, B_N$ in $\mathbb{R}^n$, one can choose a disjoint subfamily $B_{j_1}, \dots, B_{j_k}$ such that*

$$\bigcup_{i=1}^{N} B_i \subset \bigcup_{l=1}^{k} 3B_{j_l}.$$

Here $3B$ has the same centre as B and three times its radius. Consequently $m(\bigcup_i B_i) \le 3^n \sum_l m(B_{j_l})$.

Proof. Choose a ball of largest radius in the finite family. Retain it and delete it together with every ball meeting it. If balls remain, choose a largest one among those, retain it, and delete all balls meeting it. The procedure stops after at most N steps. Retained balls are disjoint because every ball meeting an earlier retained one was deleted.

A deleted ball B meets the retained ball B' that caused its deletion, and its radius r is no greater than the radius r' of B'. If their centres are c, c', then

$|c - c'| < r + r'$. For $x \in B$,

$$|x - c'| \le |x - c| + |c - c'| < 2r + r' \le 3r'.$$

Thus $B \subset 3B'$. Every original ball is retained or deleted at some step, proving the covering assertion. Subadditivity and the dilation formula give the measure bound. □

Definition 7.2. For $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, define the uncentred Hardy–Littlewood maximal function by

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

where the supremum is over all open balls containing x.

Here L^1_{loc} means integrability on every compact subset, equivalently on every bounded ball. The definition allows $Mf(x) = \infty$. It is unchanged by an a.e. modification of f, because every integral over a ball is unchanged. Subadditivity and homogeneity follow from the triangle inequality under each average.

Theorem 7.3. *For $f \in L^1(\mathbb{R}^n)$, the function Mf is measurable, finite a.e., and*

$$m(\{Mf > \lambda\}) \le \frac{3^n}{\lambda} \|f\|_1 \quad (\lambda > 0). \tag{7.1}$$

Proof. Let $E_\lambda = \{Mf > \lambda\}$. If $x \in E_\lambda$, some ball B containing x has average greater than λ. Every point of that ball also belongs to E_λ, since the same ball is admissible there. Thus E_λ is the union of all balls with average greater than λ, and is open. This proves measurability.

Fix a compact $K \subset E_\lambda$. Choose for each $x \in K$ such a ball B_x, and extract a finite subcover of K. Apply Lemma 7.1 to this finite collection. The retained balls B_{j_l} are disjoint and satisfy $\lambda m(B_{j_l}) < \int_{B_{j_l}} |f|$. Hence

$$m(K) \le 3^n \sum_l m(B_{j_l}) \le \frac{3^n}{\lambda} \sum_l \int_{B_{j_l}} |f| \le \frac{3^n}{\lambda} \|f\|_1.$$

The last inequality uses disjointness, not merely the covering property. Inner regularity of the open set E_λ allows the supremum over compact K, proving (7.1) even if its measure had not initially been known to be finite.

Finally $\{Mf = \infty\} \subset \{Mf > j\}$ for every integer j, so its measure is at most $3^n \|f\|_1 / j$ for all j. It is therefore null. □

The estimate is called weak type $(1,1)$ because it bounds the distribution rather than the L^1 norm. It cannot be replaced by an L^1 norm estimate. Let $f = \mathbf{1}_{B(0,1)}$. For $|x| > 2$, the ball $B(0, 2|x|)$ contains x and the support of f, so

$$Mf(x) \geq \frac{m(B(0,1))}{m(B(0,2|x|))} = 2^{-n}|x|^{-n}.$$

For $k \geq 1$, let $A_k = \{2^k < |x| < 2^{k+1}\}$. Its measure is $\omega_n(2^{(k+1)n} - 2^{kn})$, since the bounding spheres are null. On A_k, the displayed lower bound is at least $2^{-n}2^{-(k+1)n}$. Hence $\int_{A_k} Mf \geq 2^{-n}\omega_n(1-2^{-n})$, a positive number independent of k. The disjointness of these annuli gives $\int Mf \geq \sum_{k\geq 1} 2^{-n}\omega_n(1-2^{-n}) = \infty$.

7.2 Strong bounds above the endpoint

Theorem 7.4. *For* $1 < p \leq \infty$, $\|Mf\|_p \leq C_{n,p}\|f\|_p$.

Proof. The case $p = \infty$ is immediate because every average is bounded by $\|f\|_\infty$. Let $1 < p < \infty$. A function in $L^p(\mathbb{R}^n)$ is locally integrable by Hölder, so its maximal function is defined. For each $\lambda > 0$, decompose

$$f = f\mathbf{1}_{\{|f|>\lambda/2\}} + f\mathbf{1}_{\{|f|\leq\lambda/2\}} = f_\lambda + b_\lambda.$$

The large-value part lies in L^1, since on its support $|f| \leq (2/\lambda)^{p-1}|f|^p$. The bounded part has $Mb_\lambda \leq \lambda/2$. Subadditivity yields $\{Mf > \lambda\} \subset \{Mf_\lambda > \lambda/2\}$. Applying the weak estimate to f_λ gives

$$m(\{Mf > \lambda\}) \leq \frac{2\cdot 3^n}{\lambda}\int_{\{|f|>\lambda/2\}} |f|.$$

Use the layer-cake formula and Tonelli for the nonnegative integrand:

$$\begin{aligned}
\|Mf\|_p^p &\leq 2p3^n \int_0^\infty \lambda^{p-2}\int_{\{|f|>\lambda/2\}} |f(x)|\,dx\,d\lambda \\
&= 2p3^n \int |f(x)| \int_0^{2|f(x)|} \lambda^{p-2}\,d\lambda\,dx \\
&= \frac{p2^p3^n}{p-1}\|f\|_p^p.
\end{aligned}$$

The inner integral is finite because $p > 1$. The estimate itself proves integrability of $(Mf)^p$; it was not assumed in using the nonnegative layer formula. ◻

7.3 Lebesgue points and differentiation of averages

Theorem 7.5 (Differentiation of averages)**.** *For $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$,*

$$\lim_{r\downarrow 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)\,dy = f(x) \quad \textit{for almost every } x.$$

Proof. First let $f \in L^1$ and choose a representative finite everywhere. Write

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

For $g \in C_c$, the averages of g tend to $g(x)$ at every point: their difference is bounded by $\sup_{|y-x|<r} |g(y) - g(x)|$. Splitting $f = (f-g) + g$ therefore gives the pointwise estimate

$$Df(x) \le M(f-g)(x) + |f(x) - g(x)|.$$

Fix $\alpha > 0$. If $Df(x) > 2\alpha$, at least one term on the right exceeds α. The weak maximal estimate and Chebyshev imply

$$m_*(\{Df > 2\alpha\}) \le \frac{3^n + 1}{\alpha} \|f - g\|_1.$$

We use exterior measure here so that measurability of the limsup expression need not be established in advance. By density, $\|f-g\|_1$ can be arbitrarily small. Thus $\{Df > 2\alpha\}$ has exterior measure zero. Taking the countable union over $\alpha = 1/k$ shows that $Df = 0$ outside one null set.

For local integrability, put $f_N = f\mathbf{1}_{B(0,N+1)}$. This is integrable on the whole space. For $x \in B(0,N)$ and $r < 1$, its averages over $B(x,r)$ equal those of f, and its value at x equals $f(x)$. Apply the global result to each f_N and remove the union of their exceptional null sets. Every point of $\mathbb{R}^n$ belongs to some $B(0,N)$, giving the local assertion. □

Definition 7.6. A point x at which $f(x)$ is finite is a Lebesgue point of $f \in L^1_{\mathrm{loc}}$ if

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

Theorem 7.7. *Almost every point is a Lebesgue point of a locally integrable function.*

Proof. For real f, apply Theorem 7.5 to each of the countably many locally integrable functions $|f - a|$, $a \in \mathbb{Q}$. Remove the union of the resulting null sets, together with the null set where f is infinite. At a remaining point x, the differentiation identity holds for every rational a simultaneously.

Given $\varepsilon > 0$, choose a rational a with $|f(x) - a| < \varepsilon$. The triangle inequality gives

$$\frac{1}{m(B(x,r))}\int_{B(x,r)} |f(y) - f(x)|\,dy \le \frac{1}{m(B(x,r))}\int_{B(x,r)} |f(y) - a|\,dy + |a - f(x)|.$$

The limsup is at most $2|f(x) - a| < 2\varepsilon$. Letting $\varepsilon \downarrow 0$ proves the result. For complex f, use the countable dense set $\mathbb{Q} + i\mathbb{Q}$ and the locally integrable real functions $|f - a|$. The same argument applies. □

The countable set of test constants is important. We first obtain one full-measure set on which every rational test works, and only then choose a test value depending on x.

Corollary 7.8 (Lebesgue density). *For a measurable set E, almost every $x \in E$ satisfies*

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

and almost every $x \notin E$ has the same ratio tending to zero.

Proof. The indicator $\mathbf{1}_E$ is locally integrable. Apply the differentiation theorem to it and distinguish its values one and zero. □

A point of density one need not be an interior point. The positive-measure nowhere dense set in Example 1.21 has density one at almost all of its points despite having no interior.

Proposition 7.9. *Let x be a Lebesgue point of f. Suppose measurable sets A_j satisfy $A_j \subset B(x, r_j)$, $r_j \to 0$, and $m(A_j) \ge c\, m(B(x, r_j))$ for one $c > 0$. Then*

$$\frac{1}{m(A_j)}\int_{A_j} f \longrightarrow f(x).$$

In particular, averages over arbitrary balls containing x and of radii tending to zero converge at every Lebesgue point.

Proof. The absolute error is at most

$$\frac{1}{m(A_j)}\int_{A_j} |f(y) - f(x)|\,dy \le \frac{1}{c}\frac{1}{m(B(x,r_j))}\int_{B(x,r_j)} |f(y) - f(x)|\,dy,$$

which tends to zero. A ball $B(c, r)$ containing x is contained in $B(x, 2r)$ and has 2^{-n} times its measure. Thus it satisfies the required comparison, uniformly in its centre. Cubes of side length tending to zero and containing x admit a similar comparison. □

Proposition 7.10. *If $f \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$, $1 \le p < \infty$, then for almost every x,*

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

Proof. Apply differentiation to $|f - a|^p$ for every a in a countable dense set of scalars. These functions are locally integrable because $|f - a|^p \le 2^{p-1}(|f|^p + |a|^p)$. Remove a common null set. At a remaining point, use

$$|f(y) - f(x)|^p \le 2^{p-1}\big(|f(y) - a|^p + |a - f(x)|^p\big).$$

The limsup of its average is at most $2^p|a - f(x)|^p$. Choose dense test values tending to $f(x)$ to conclude. □

The Lebesgue-point set depends on the representative because its definition uses the value at the point. Any two a.e. equal representatives nevertheless have full-measure Lebesgue-point sets, and their averages over every fixed ball agree.

7.4 The indefinite integral and absolute continuity

Definition 7.11. A function $F : [a, b] \to \mathbb{C}$ is absolutely continuous if for every $\varepsilon > 0$ there is $\delta > 0$ such that

$$\sum_{j=1}^{N} |F(b_j) - F(a_j)| < \varepsilon$$

for every finite family of intervals $(a_j, b_j) \subset [a, b]$ with disjoint interiors and $\sum_j (b_j - a_j) < \delta$.

Proposition 7.12. *If $f \in L^1(a, b)$ and $F(x) = c + \int_a^x f(t)\, dt$, then F is absolutely continuous and $F' = f$ almost everywhere.*

Proof. For a disjoint finite interval family,

$$\sum_j |F(b_j) - F(a_j)| \le \sum_j \int_{a_j}^{b_j} |f| = \int_{\bigcup_j (a_j, b_j)} |f|.$$

Absolute continuity of the integral makes this smaller than ε when the total length is sufficiently small. Thus F is absolutely continuous. In particular, it is continuous, since the definition can be applied to one interval.

Extend f by zero outside (a, b) and let $x \in (a, b)$ be a Lebesgue point. For either sign of h, with $x + h \in (a, b)$, the difference quotient is the average of f over the interval joining x to $x + h$. The absolute difference from $f(x)$ is bounded by

$$\frac{1}{|h|} \int_{x-|h|}^{x+|h|} |f(t) - f(x)| \, dt,$$

which is twice a centred average tending to zero. Therefore the two-sided derivative exists and equals $f(x)$. Almost every interior point is a Lebesgue point. □

A Lipschitz function is absolutely continuous, since its increments are bounded by its Lipschitz constant times the interval lengths. The converse to Proposition 7.12, namely the representation of every absolutely continuous function as an integral of its derivative, will be proved in Chapter 12. It must not be replaced by the weaker assertion that a continuous function has an integrable derivative almost everywhere: the Cantor function will distinguish these statements.

7.5 Why arbitrary rectangles are different

Let M_{rec} be defined as the supremum of averages over all axis-parallel rectangles containing the point, without any bound on eccentricity. The ball covering proof does not apply to this family. There is a direct obstruction to its weak $(1, 1)$ bound already in $\mathbb{R}^2$.

Proposition 7.13. *There is no constant C such that $m(\{M_{\text{rec}} f > \lambda\}) \le C\|f\|_1/\lambda$ for all $f \in L^1(\mathbb{R}^2)$ and $\lambda > 0$.*

Proof. Take $f = \mathbf{1}_{(0,1)^2}$. For $x, y > 1$, use rectangles $(0, x+\eta) \times (0, y+\eta)$ and let $\eta \downarrow 0$. Their averages tend to $1/(xy)$, so $M_{\text{rec}} f(x, y) \ge 1/(xy)$. For $0 < \lambda < 1$, the superlevel set therefore contains $\{x > 1, y > 1, xy < 1/\lambda\}$. Its measure is

$$\int_1^{1/\lambda} \left(\frac{1}{\lambda x} - 1\right) dx = \frac{1}{\lambda} \log \frac{1}{\lambda} - \frac{1}{\lambda} + 1.$$

Multiplying by λ gives a quantity tending to infinity as $\lambda \downarrow 0$, while $\|f\|_1 = 1$. This contradicts any proposed weak $(1, 1)$ constant. □

The regular-shrinking proposition only uses sets comparable in measure to a containing ball. Rectangles with arbitrarily large eccentricity do not satisfy that uniform comparison. The preceding counterexample concerns the maximal estimate; it should not be confused with a proof of every possible differentiation failure for that larger family.

Exercises

Exercise 7.1. Compare the centred and uncentred maximal functions and obtain inequalities between them with dimension-dependent constants.

Exercise 7.2. Prove that Mf is lower semicontinuous for locally integrable f with the uncentred-ball definition.

Exercise 7.3. Give the full annulus calculation showing that the maximal function of a nonzero nonnegative compactly supported integrable function is not integrable.

Exercise 7.4. Repeat the weak $(1, 1)$ argument with cubes, using an appropriate finite covering lemma.

Exercise 7.5. Locate every use of compactness, disjointness, and inner regularity in the proof of Theorem 7.3.

Exercise 7.6. Use the differentiation theorem to prove $|f| \leq Mf$ almost everywhere. Explain why this should not be used before differentiation has been established.

Exercise 7.7. Show that averages over cubes not centred at a Lebesgue point still converge when the cubes contain that point and their diameters tend to zero.

Exercise 7.8. Prove the density theorem for the complement of a measurable set without assuming its measure is finite.

Exercise 7.9. Give an integrable function and a null modification for which a specified point is a Lebesgue point of one representative but not the other.

Exercise 7.10. Prove that an indefinite integral is uniformly continuous on a compact interval directly from absolute continuity of the integral.

Exercise 7.11. If a locally integrable function has zero integral over every ball, show that it vanishes a.e.

Exercise 7.12. Prove the L^p Lebesgue-point assertion by approximating in L^p and applying the weak maximal estimate to $|f - g|^p$.

Exercise 7.13. Extend Proposition 7.13 to dimensions greater than two.

Exercise 7.14. Let E have positive measure. Use density points to show that for every $\varepsilon > 0$ there is a ball B with $m(E \cap B) > (1 - \varepsilon)m(B)$.

Chapter 8

Approximation and the Fourier transform

Approximation by convolution has two distinct forms. Norm convergence follows from translation continuity and concentration of the kernels. Convergence at Lebesgue points requires additional control of the spatial shape of the kernels. We prove both statements before applying them to Fourier inversion.

8.1 Approximate identities

Definition 8.1. A family $(K_\delta)_{\delta>0} \subset L^1(\mathbb{R}^n)$ is a good kernel family if

$$\int K_\delta = 1, \quad \sup_{\delta>0} \|K_\delta\|_1 \le A < \infty,$$

and, for every $r > 0$, $\int_{|y|>r} |K_\delta(y)|\, dy \to 0$ as $\delta \downarrow 0$.

The kernels need not be nonnegative. The mass-one identity is used without absolute values, while uniform boundedness and tail estimates concern absolute values.

Theorem 8.2. *For a good kernel family and $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$, one has $K_\delta * f \to f$ in L^p. If f is bounded and uniformly continuous, convergence is uniform. If f is bounded and continuous at a specified point x, then $(K_\delta * f)(x) \to f(x)$.*

Proof. Young's inequality makes the convolution well defined a.e. and bounded in L^p. Since $\int K_\delta = 1$, its error is

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

where the absolute integrals exist. For $p > 1$, Hölder with respect to the weight $|K_\delta(y)|\,dy$ gives

$$|K_\delta * f(x) - f(x)|^p \le \|K_\delta\|_1^{p-1} \int |K_\delta(y)||f(x-y) - f(x)|^p\,dy.$$

The corresponding inequality for $p = 1$ is the integral triangle inequality. Tonelli then yields

$$\|K_\delta * f - f\|_p^p \le A^{p-1} \int |K_\delta(y)|\|\tau_y f - f\|_p^p\,dy.$$

Fix $\varepsilon > 0$. Translation continuity gives $r > 0$ such that $\|\tau_y f - f\|_p < \varepsilon$ when $|y| < r$. On the complementary set, use $\|\tau_y f - f\|_p \le 2\|f\|_p$. Thus

$$\|K_\delta * f - f\|_p^p \le A^p \varepsilon^p + A^{p-1}(2\|f\|_p)^p \int_{|y|\ge r} |K_\delta(y)|\,dy.$$

A sphere is null, so the stated tail condition applies with either strict or nonstrict inequality. Let $\delta \downarrow 0$ and then $\varepsilon \downarrow 0$.

For bounded uniformly continuous f, use the same near/far split before taking the supremum in x: the near part is at most $A \sup_{|y|<r,x} |f(x-y) - f(x)|$, and the far part is at most $2\|f\|_\infty$ times the kernel tail. At one continuity point, the identical argument uses only the local supremum of $|f(x-y) - f(x)|$ at that fixed x. Boundedness controls the far part. □

Definition 8.3. A normalized family is called an approximation to the identity in the pointwise sense used here if, for a constant A independent of δ,

$$|K_\delta(y)| \le A\delta^{-n}, \quad |K_\delta(y)| \le A\delta|y|^{-n-1} \quad (y \ne 0). \tag{8.1}$$

These bounds imply the good-kernel conditions. On $|y| < \delta$, the integral is at most $A\omega_n$. On the annulus $2^k\delta \le |y| < 2^{k+1}\delta$, it is at most

$$A\delta(2^k\delta)^{-n-1}\omega_n(2^{k+1}\delta)^n \le C_n A 2^{-k}.$$

Summing gives a uniform L^1 bound. Splitting $|y| > r$ into annuli based on r gives a tail bound $C_n A\delta/r$, which tends to zero.

Theorem 8.4. *If* (8.1) *holds and* $\int K_\delta = 1$*, then for every Lebesgue point x of $f \in L^1(\mathbb{R}^n)$,*

$$(K_\delta * f)(x) \longrightarrow f(x).$$

The convolution integral is absolutely convergent at such a point for every $\delta > 0$.

Proof. Fix a Lebesgue point x and put

$$A_x(r) = r^{-n} \int_{|y|<r} |f(x-y) - f(x)|\, dy.$$

Then $A_x(r) \to 0$ as $r \downarrow 0$. It is bounded for all $r > 0$. To see this, the limit controls sufficiently small r; on any interval $r_0 \le r \le 1$, the numerator is bounded by $\int_{|y|<1} |f(x-y)|\, dy + \omega_n |f(x)|$, and the denominator is at least r_0^n. For $r \ge 1$, the bound is at most $\|f\|_1 + \omega_n|f(x)|$. Denote a global bound by M_x. Absolute continuity of the integral on annuli also shows that A_x is continuous for $r > 0$, although boundedness and the limit at zero are the properties used below.

Split the error integral into $|y| < \delta$ and the annuli $2^k\delta \le |y| < 2^{k+1}\delta$, $k \ge 0$. The first kernel bound gives an inner contribution at most $AA_x(\delta)$. The second gives for the kth annulus

$$\begin{aligned}\int_{2^k\delta \le |y| < 2^{k+1}\delta} &|K_\delta(y)||f(x-y) - f(x)|\, dy \\ &\le A\delta(2^k\delta)^{-n-1}(2^{k+1}\delta)^n A_x(2^{k+1}\delta) \le C_n A 2^{-k} A_x(2^{k+1}\delta).\end{aligned}$$

The series of these bounds is finite, because $A_x \le M_x$. Since $K_\delta \in L^1$ and $f(x)$ is finite, this also proves absolute integrability of $K_\delta(y)f(x-y)$.

Given $\varepsilon > 0$, first choose N so that $C_n A M_x \sum_{k \ge N} 2^{-k} < \varepsilon/2$. Keep N fixed. Now choose δ sufficiently small that $A_x(\delta)$ and all finitely many quantities $A_x(2^{k+1}\delta)$, $0 \le k < N$, make their weighted sum less than $\varepsilon/2$. This is possible because every displayed radius tends to zero. The full error is less than ε. □

The order of choices in the proof is essential: the large-scale annular tail is made small uniformly first, and only finitely many remaining scales are then sent to zero. Good-kernel concentration alone does not supply the same pointwise conclusion for arbitrary unbounded integrable functions.

8.2 Mollifiers and smooth approximation

Choose a nonnegative function

$$\eta(x) = \begin{cases} c\exp(-1/(1-|x|^2)), & |x| < 1, \\ 0, & |x| \ge 1, \end{cases}$$

where c makes $\int \eta = 1$, and put $\eta_\delta(x) = \delta^{-n}\eta(x/\delta)$. The normalization exists because the unnormalized function is positive on the unit ball and bounded with compact support. It is smooth across $|x| = 1$: each derivative inside the ball is a polynomial in coordinates and reciprocal powers of $1-|x|^2$, multiplied by the exponential, and every such expression tends to zero at the boundary. The exponential dominates every reciprocal power.

The family satisfies (8.1). Indeed, it is bounded by $C\delta^{-n}$ and supported in $|x| \le \delta$; on this support, $\delta^{-n} \le \delta|x|^{-n-1}$ for $x \ne 0$. It is therefore a normalized approximation to the identity.

Proposition 8.5. *Let $\Omega \subset \mathbb{R}^n$ be open and $f \in L^1_{\mathrm{loc}}(\Omega)$. For $\Omega_\delta = \{x \in \Omega : \operatorname{dist}(x, \Omega^c) > \delta\}$, define*

$$f_\delta(x) = \int_\Omega \eta_\delta(x-y) f(y)\, dy, \quad x \in \Omega_\delta.$$

Then $f_\delta \in C^\infty(\Omega_\delta)$ and $D^\alpha f_\delta = (D^\alpha \eta_\delta) * f$ there. If $f \in L^p_{\mathrm{loc}}(\Omega)$, $1 \le p < \infty$, then $f_\delta \to f$ in L^p on every relatively compact open subset. At every Lebesgue point, convergence also holds pointwise for all sufficiently small δ.

Proof. Fix $x_0 \in \Omega_\delta$. In a sufficiently small neighbourhood of x_0, all translated kernel supports lie in one compact subset $K \Subset \Omega$. A difference quotient in x_i of the kernel converges uniformly in y to $\partial_i \eta_\delta(x-y)$ and is bounded by the supremum of that derivative. Thus it is dominated, after multiplication by $|f(y)|$, by an integrable function $C|f|\mathbf{1}_K$. Differentiation under the integral is valid. Repeating this argument for every derivative proves smoothness and the derivative formula.

For norm approximation, fix $V \Subset W \Subset \Omega$. When $\delta < \operatorname{dist}(\overline{V}, W^c)$, the convolution at points of V depends only on $f\mathbf{1}_W$, extended by zero to $\mathbb{R}^n$. This extension is in L^p. Theorem 8.2 gives convergence globally for that extension, and hence on V. The same localization reduces the pointwise assertion to Theorem 8.4; a small neighbourhood of an interior Lebesgue point is unaffected by the cutoff. $\square$

Corollary 8.6. *For $1 \le p < \infty$, $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$.*

Proof. First approximate f by a bounded compactly supported measurable function f_N in L^p. For this fixed function, $\eta_\delta * f_N$ is smooth and supported in the compact δ-neighbourhood of its support. It converges to f_N in L^p as $\delta \downarrow 0$. Choose N and then δ to make the two errors separately small. $\square$

Lemma 8.7. *For compact $K \subset \Omega$ with Ω open, there is $\chi \in C_c^\infty(\Omega)$ such that $0 \le \chi \le 1$ and $\chi = 1$ on a neighbourhood of K.*

Proof. The case $K = \varnothing$ is immediate. Otherwise choose $d > 0$ with the closed $3d$-neighbourhood of K compactly contained in Ω. Let $U = \{x : \mathrm{dist}(x, K) < 2d\}$ and define $\chi = \eta_d * \mathbf{1}_U$. The convolution is smooth and lies between zero and one. If $\mathrm{dist}(x, K) < d$, then $B(x, d) \subset U$, so the full kernel mass is integrated and $\chi(x) = 1$. If $\mathrm{dist}(x, K) > 3d$, the kernel support misses U, so $\chi(x) = 0$. Its support is therefore a compact subset of Ω. □

8.3 The Fourier transform on L^1

Definition 8.8. For $f \in L^1(\mathbb{R}^n)$, define $\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x)e^{-2\pi i x\cdot\xi}\,dx$.

Proposition 8.9. *The Fourier transform is bounded and uniformly continuous, with* $\|\widehat{f}\|_\infty \le \|f\|_1$*, and* $\widehat{f}(\xi) \to 0$ *as* $|\xi| \to \infty$*. For* $f, g \in L^1$*,* $\widehat{f * g} = \widehat{f}\,\widehat{g}$*.*

Proof. The modulus of the exponential is one, so the defining integral is absolutely convergent and has the stated bound. For a frequency increment h,

$$|\widehat{f}(\xi + h) - \widehat{f}(\xi)| \le \int |f(x)||e^{-2\pi i x\cdot h} - 1|\,dx.$$

This upper bound is independent of ξ and tends to zero by dominated convergence, using $2|f|$. Hence the continuity is uniform.

For decay, first let $\phi \in C_c^\infty$. If $\xi_i \ne 0$, integration by parts in x_i gives $2\pi i\xi_i\widehat{\phi}(\xi) = \widehat{\partial_i\phi}(\xi)$, with no boundary term because of compact support. For each nonzero ξ, choose i with $|\xi_i| \ge |\xi|/\sqrt{n}$. The transform is then bounded by $C_\phi/|\xi|$. Given general f and $\varepsilon > 0$, choose ϕ with $\|f - \phi\|_1 < \varepsilon$. The uniform Fourier bound gives $|\widehat{f}(\xi)| \le \varepsilon + |\widehat{\phi}(\xi)|$, proving decay after $|\xi| \to \infty$ and then $\varepsilon \downarrow 0$.

For convolution, the absolute product integral over (x, y) is $\|f\|_1\|g\|_1$, as in Theorem 4.15. Fubini and $z = x - y$ yield

$$\begin{aligned}\widehat{f * g}(\xi) &= \int\int f(x-y)g(y)e^{-2\pi i x\cdot\xi}\,dy\,dx\\ &= \Big(\int f(z)e^{-2\pi i z\cdot\xi}\,dz\Big)\Big(\int g(y)e^{-2\pi i y\cdot\xi}\,dy\Big).\end{aligned}$$

□

8.4 Gaussian kernels and inversion

Lemma 8.10. *For* $G(x) = e^{-\pi|x|^2}$*, one has* $\int_{\mathbb{R}^n} G = 1$ *and* $\widehat{G} = G$*. More generally,*

$$\widehat{e^{-\pi a|x|^2}}(\xi) = a^{-n/2}e^{-\pi|\xi|^2/a} \quad (a > 0).$$

Proof. We first justify the disc measure used in the normalization. Tonelli applied to its vertical sections gives

$$m_2(B(0,r)) = \int_{-r}^{r} 2\sqrt{r^2 - s^2}\, ds.$$

The one-variable substitution $s = r\sin\theta$, initially on inner intervals and then by convergence at the endpoints, turns this into $2r^2 \int_{-\pi/2}^{\pi/2} \cos^2\theta\, d\theta = \pi r^2$. This computation uses only the already established section theorem and one-variable integration, not a multivariable polar-coordinate formula.

In one dimension let $I = \int_{\mathbb{R}} e^{-\pi x^2}\, dx$, which is finite because the Gaussian is bounded by an integrable exponential outside a bounded interval. Tonelli gives $I^2 = \int_{\mathbb{R}^2} e^{-\pi(x^2+y^2)}\, d(x,y)$. By the layer-cake formula and the disc-measure computation, this equals

$$\int_0^1 m_2\Big(B\Big(0, \sqrt{-\log t/\pi}\Big)\Big)\, dt = \int_0^1 -\log t\, dt = 1.$$

Thus $I = 1$. The n-dimensional normalization is the product of the one-dimensional integrals, by Tonelli.

In one dimension put $F(\xi) = \widehat{G}(\xi)$. The derivative of the integrand is bounded in absolute value by the integrable function $2\pi|x|G(x)$. Differentiation under the integral therefore gives

$$F'(\xi) = -2\pi i \int xG(x)e^{-2\pi i x\xi}\, dx.$$

Since $G' = -2\pi x G$, integration by parts over $[-R, R]$ and then $R \to \infty$ gives

$$\int G'(x)e^{-2\pi i x\xi}\, dx = 2\pi i\xi F(\xi).$$

The endpoint terms tend to zero, and the derivative integrals converge absolutely. Therefore $F' = -2\pi\xi F$. The derivative of $e^{\pi\xi^2}F(\xi)$ is zero, and $F(0) = 1$, so $F(\xi) = e^{-\pi\xi^2}$. Fubini factors the n-dimensional transform into one-dimensional transforms. Finally, the substitution $y = \sqrt{a}\, x$ gives the scaled formula. □

Theorem 8.11 (Fourier inversion). *If $f \in L^1(\mathbb{R}^n)$ and $\widehat{f} \in L^1(\mathbb{R}^n)$, then*

$$f(x) = \int_{\mathbb{R}^n} \widehat{f}(\xi)e^{2\pi i x\cdot\xi}\, d\xi$$

at every Lebesgue point of f, and therefore almost everywhere. The integral on the right defines a continuous representative.

Proof. For $\varepsilon > 0$, define

$$I_\varepsilon(x) = \int \widehat{f}(\xi) e^{-\pi\varepsilon|\xi|^2} e^{2\pi i x\cdot\xi}\, d\xi.$$

Substitute the definition of $\widehat{f}$ and use Fubini. The absolute double integral is $\|f\|_1 \int e^{-\pi\varepsilon|\xi|^2}\, d\xi < \infty$, so the interchange is justified. The Gaussian formula gives

$$I_\varepsilon(x) = \int f(y)\varepsilon^{-n/2} e^{-\pi|x-y|^2/\varepsilon}\, dy = (K_{\sqrt{\varepsilon}} * f)(x),$$

where $K_\delta(z) = \delta^{-n} e^{-\pi|z|^2/\delta^2}$.

This kernel has mass one and satisfies (8.1): the first estimate is immediate, and the second follows from boundedness of $r^{n+1}e^{-\pi r^2}$ on $[0,\infty)$. Thus $I_\varepsilon(x) \to f(x)$ at every Lebesgue point. On the Fourier side, $\widehat{f} \in L^1$ and the Gaussian multiplier is bounded by one and tends to one. Dominated convergence gives $I_\varepsilon(x) \to \int \widehat{f}(\xi)e^{2\pi i x\cdot\xi}\, d\xi$ for every x. The two limits agree at every Lebesgue point. Continuity of the right side follows from the same dominated-convergence argument, or from the uniform continuity estimate for an L^1 Fourier transform with the sign reversed. □

Corollary 8.12. *If $f \in L^1$ and $\widehat{f} = 0$, then $f = 0$ a.e.*

Proof. The zero transform is integrable, so Theorem 8.11 applies. If the transform is only specified to be zero a.e., its continuity first makes it zero everywhere. □

8.5 An application to the heat equation

Define

$$H_t(x) = (4\pi t)^{-n/2} e^{-|x|^2/(4t)}, \quad t > 0.$$

The normalization and transform are $\int H_t = 1$ and $\widehat{H}_t(\xi) = e^{-4\pi^2 t|\xi|^2}$.

Proposition 8.13. *For $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$, the function $u(t,x) = H_t * f(x)$ is smooth for $t > 0$, satisfies $\partial_t u = \Delta u$, and tends to f in L^p as $t \downarrow 0$. Moreover $H_t * H_s = H_{t+s}$, and $\int u(t,x)\, dx = \int f$ when $f \in L^1$.*

Proof. The initial convergence follows from Theorem 8.2, with spatial scale proportional to $\sqrt{t}$. Direct differentiation gives

$$\partial_t H_t = \Big(-\frac{n}{2t} + \frac{|x|^2}{4t^2}\Big) H_t, \quad \Delta H_t = \Big(-\frac{n}{2t} + \frac{|x|^2}{4t^2}\Big) H_t.$$

Every mixed derivative in t and x is a polynomial in x with coefficients involving powers of t^{-1}, multiplied by H_t. On a compact positive time interval $a \le t \le b$, these derivatives are bounded by $C(1+|x|)^N e^{-c|x|^2}$ for suitable positive C, c and an integer N. This majorant belongs to every L^q, $1 \le q \le \infty$.

For x in a compact set, the translated kernel derivatives admit the same kind of majorant in the integration variable. Hölder makes their products with f integrable. Difference quotients are controlled by one additional kernel derivative through the mean value theorem. Differentiation under the convolution is therefore valid to every order. This proves smoothness and transfers the identity $\partial_t H_t = \Delta H_t$ to u.

The same justification can be made in L^p norm: the derivative kernels are continuous in L^1 on positive time intervals by dominated convergence, and their difference quotients tend in L^1 to the corresponding derivatives. Young's inequality transfers that convergence to L^p after convolution with f. Thus no inference of norm differentiability is made from pointwise kernel differentiation alone.

For the semigroup identity, the convolution formula gives $\widehat{H_t * H_s} = e^{-4\pi^2(t+s)|\xi|^2} = \widehat{H}_{t+s}$. Fourier uniqueness implies equality a.e.; both kernels are continuous, so equality holds everywhere. The mass identity for $f \in L^1$ follows from the integral formula for L^1 convolution and $\int H_t = 1$. □

The norm differentiation used for the heat kernel can also be checked directly. On $t \in [a,b] \subset (0,\infty)$, every fixed mixed derivative of H_t is bounded in absolute value by $C(1+|x|)^N e^{-c|x|^2}$ for constants depending on the interval and the derivative. Differentiating $(4\pi t)^{-n/2} e^{-|x|^2/(4t)}$ produces only powers of t^{-1} and polynomial factors in x; these are bounded in the stated form on $[a,b]$. The majorant is integrable, by its Gaussian decay and the measures of dyadic annuli. For small h, the fundamental theorem in the parameter gives

$$\frac{H_{t+h}(x) - H_t(x)}{h} = \int_0^1 \partial_t H_{t+sh}(x)\, ds.$$

The integrand is continuous in the parameter and has the preceding common integrable majorant. Dominated convergence in x therefore gives convergence of this difference quotient to $\partial_t H_t$ in L^1. Spatial difference quotients have the same property, using the majorant on a bounded set of translations. Young's convolution inequality then transfers these norm limits to $H_t * f$. This supplies the norm justification separately from the pointwise differentiation under the integral.

Exercises

Exercise 8.1. Verify the good-kernel conditions for $K_\delta(x) = \delta^{-n}K(x/\delta)$ when $K \in L^1$ and $\int K = 1$.

Exercise 8.2. Show that the uniform L^1 bound in the good-kernel definition is needed when kernels are allowed to change sign.

Exercise 8.3. Give a bounded function for which mollification does not converge in L^∞.

Exercise 8.4. Verify the pointwise kernel bounds for the one-dimensional Poisson kernel $P_\delta(x) = \delta/(\pi(x^2 + \delta^2))$.

Exercise 8.5. Prove local L^p convergence of mollifiers without extending the original function across the boundary of its domain.

Exercise 8.6. Show that the derivative of a mollification is the convolution with the kernel derivative, and explain why this alone does not identify it with a mollification of a weak derivative unless that derivative exists.

Exercise 8.7. Prove the Fourier identities for translation, modulation, dilation, and reflection with the convention used in these notes.

Exercise 8.8. If $x_i f \in L^1$, prove differentiability of $\widehat{f}$ in ξ_i and identify its derivative.

Exercise 8.9. Prove rapid decay of the Fourier transform of a compactly supported smooth function by repeated integration by parts.

Exercise 8.10. Derive $\int \widehat{f}\, g = \int f\, \widehat{g}$ for $f, g \in L^1$, and state why both expressions are absolutely integrable.

Exercise 8.11. Prove the Gaussian convolution identity directly by completing the square, independently of Fourier uniqueness.

Exercise 8.12. For $f \in L^1$, prove $\|H_t * f\|_1 \le \|f\|_1$ and preservation of the integral. Explain why equality of the norms need not hold for signed f.

Exercise 8.13. Show that convolution with a good kernel tends to the identity on $C_0(\mathbb{R}^n)$ in the uniform norm.

Exercise 8.14. Let $f \in L^1$ and suppose its Fourier transform has compact support and is integrable. Prove that its continuous representative is smooth and obtain bounds for its derivatives.

Chapter 9

Weak derivatives and Sobolev spaces

Integration by parts makes it possible to define derivatives without assuming pointwise differentiability. The test functions have compact support, so the definition is local and does not require a regular boundary. Approximation in Sobolev norms will then be proved by transferring derivatives to a mollifier and by a locally finite construction.

9.1 Test functions and weak differentiation

Let $\Omega \subset \mathbb{R}^n$ be open. A test function is an element of $C_c^\infty(\Omega)$. A multi-index is $\alpha = (\alpha_1, \ldots, \alpha_n) \in \{0, 1, \ldots\}^n$, with $|\alpha| = \sum_i \alpha_i$ and $D^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$.

Definition 9.1. For $u, v \in L^1_{\mathrm{loc}}(\Omega)$, say that $D^\alpha u = v$ weakly if

$$\int_\Omega u D^\alpha \phi = (-1)^{|\alpha|} \int_\Omega v\phi \quad \text{for every } \phi \in C_c^\infty(\Omega).$$

The integrals are finite because all test functions and their derivatives are bounded on one compact support and u, v are locally integrable. The sign is the one obtained by repeated classical integration by parts. When a classical $C^{|\alpha|}$ derivative exists, it satisfies this identity. Boundary terms vanish because the test function is zero near the boundary of the region where integration is performed.

It is useful to distinguish weak derivatives represented by functions from general distributional derivatives. A distribution T is a linear functional on test functions such that for each compact $K \Subset \Omega$ there are C and an integer m

with
$$|T(\phi)| \le C \max_{|\alpha|\le m} \sup |D^\alpha \phi| \quad (\operatorname{supp}\phi \subset K).$$
A locally integrable function gives $T_u(\phi) = \int u\phi$. Define $D^\alpha T(\phi) = (-1)^{|\alpha|} T(D^\alpha \phi)$. Thus every locally integrable function has distributional derivatives, but those derivatives need not be represented by locally integrable functions.

Lemma 9.2. *If $w \in L^1_{\mathrm{loc}}(\Omega)$ and $\int w\phi = 0$ for all test functions, then $w = 0$ a.e. Consequently weak derivatives, when they exist, are unique up to null sets.*

Proof. Fix a ball $B \Subset \Omega$ and a larger relatively compact open set W containing its closure. For sufficiently small δ and each $x \in B$, the function $y \mapsto \eta_\delta(x - y)$ is a test function in Ω. Hence $(\eta_\delta * w)(x) = 0$ on B. By Proposition 8.5, these mollifications tend to w in $L^1(B)$. Therefore $\|w\|_{L^1(B)} = 0$.

A countable family of relatively compact balls covers Ω, so the union of their exceptional null sets is null. This gives $w = 0$ a.e. on Ω. If v_1, v_2 are two weak derivatives, their difference has zero pairing with every test function, and the first assertion applies. □

Lemma 9.3. *If Ω is connected, $u \in L^1_{\mathrm{loc}}(\Omega)$, and every weak first derivative is zero, then u is constant a.e. on Ω.*

Proof. Fix $B \Subset \Omega$. For small δ, differentiate the mollification and use the weak identity with the translated kernel as test function:
$$\partial_i u_\delta(x) = \int u(y)\partial_{x_i}\eta_\delta(x-y)\,dy = -\int u(y)\partial_{y_i}\eta_\delta(x-y)\,dy = 0.$$
Thus the smooth function u_δ is constant, say c_δ, on the ball B. Its average on B is c_δ, and $L^1(B)$ convergence to u shows $c_\delta \to m(B)^{-1}\int_B u$. Therefore u equals this constant a.e. on B.

If two such balls overlap, their intersection is a nonempty open set and has positive measure, so their constants must agree. In a connected open set, any two points can be joined by a finite chain of overlapping relatively compact balls: the set of points reachable from one fixed ball is open, and its complement is also open. Connectedness makes the complement empty. Taking a countable ball cover makes the a.e. conclusions simultaneous, so a single constant works on all of Ω. □

Example 9.4. On $(0, 2)$, let $u(x) = x$ for $0 < x < 1$ and $u(x) = 1$ for $1 \le x < 2$. For a test function,
$$\int_0^2 u\phi' = \int_0^1 x\phi' + \int_1^2 \phi' = \phi(1) - \int_0^1 \phi - \phi(1) = -\int_0^1 \phi.$$

Thus the weak derivative is $\mathbf{1}_{(0,1)}$. The two interface terms cancel because u is continuous at one, although its classical derivative does not exist there.

Example 9.5. On $\mathbb{R}$, the weak derivative of $|x|$ is $\operatorname{sgn} x$, and its second distributional derivative is $2\delta_0$, where $\delta_0(\phi) = \phi(0)$. For $\phi \in C_c^\infty(\mathbb{R})$, integration by parts on the two half lines gives

$$\begin{aligned}\int_{\mathbb{R}} |x|\phi'(x)\,dx &= \int_{-\infty}^{0} (-x)\phi'(x)\,dx + \int_{0}^{\infty} x\phi'(x)\,dx \\ &= \int_{-\infty}^{0} \phi(x)\,dx - \int_{0}^{\infty} \phi(x)\,dx = -\int_{\mathbb{R}} \operatorname{sgn}(x)\phi(x)\,dx.\end{aligned}$$

The boundary terms vanish at infinity by compact support and at zero because of the factor x. Differentiating the representative $\operatorname{sgn} x$ distributionally then gives

$$-\int_{\mathbb{R}} \operatorname{sgn}(x)\phi'(x)\,dx = \int_{-\infty}^{0} \phi'(x)\,dx - \int_{0}^{\infty} \phi'(x)\,dx = \phi(0) + \phi(0) = 2\phi(0).$$

The Dirac distribution is not a locally integrable function. If $\delta_0 = T_g$ for $g \in L^1_{\mathrm{loc}}$, choose $0 \le \phi \le 1$ smooth, compactly supported, with $\phi(0) = 1$, and set $\phi_\varepsilon(x) = \phi(x/\varepsilon)$. Then $\int g\phi_\varepsilon \to 0$ by absolute continuity of the integral on shrinking supports, but $\delta_0(\phi_\varepsilon) = 1$. This is a contradiction. In particular, it is incorrect to describe the Dirac distribution as an ordinary function that is infinite at one point and zero elsewhere.

More generally, suppose u is C^1 on either side of zero, has finite one-sided limits, and its ordinary derivative is locally integrable across zero. Choose R with $\operatorname{supp}\phi \subset (-R, R)$ and integrate away from $(-\delta, \delta)$:

$$\begin{aligned}\int_{-R}^{-\delta} u\phi' + \int_{\delta}^{R} u\phi' = u(-\delta)\phi(-\delta) - u(\delta)\phi(\delta) \\ - \int_{-R}^{-\delta} u'\phi - \int_{\delta}^{R} u'\phi.\end{aligned}$$

Finite one-sided limits make u bounded near zero, so the omitted integral of $u\phi'$ tends to zero. Local integrability of u' permits passage to the limit in the other two integrals. Continuity of ϕ therefore gives

$$-\int_{\mathbb{R}} u\phi' = \int_{\mathbb{R}} u'_{\text{ordinary}}\phi + (u(0+) - u(0-))\phi(0).$$

This is the distributional identity $Du = u'_{\text{ordinary}} + (u(0+) - u(0-))\delta_0$. A nonzero jump therefore produces a term that cannot be represented by an L^1_{loc} weak derivative.

9.2 The Sobolev norm and elementary properties

Definition 9.6. For an integer $k \geq 0$ and $1 \leq p \leq \infty$, let $W^{k,p}(\Omega)$ be the space of $u \in L^p(\Omega)$ whose weak derivatives $D^\alpha u$ belong to $L^p(\Omega)$ for all $|\alpha| \leq k$. For $p < \infty$, use

$$\|u\|_{W^{k,p}} = \Big(\sum_{|\alpha|\leq k} \|D^\alpha u\|_p^p\Big)^{1/p},$$

and for $p = \infty$ use the maximum of the finitely many essential suprema. Write $H^k = W^{k,2}$. The local space requires these conditions on every relatively compact open subset.

The condition $|\alpha| = 0$ is included. Local integrability of u and its derivatives is not the same as membership in the global Sobolev space.

Proposition 9.7. *Weak derivatives are linear, restrict to open subsets, and commute as distributional derivatives. If $u \in W^{k,p}_{\mathrm{loc}}(\Omega)$ and $\psi \in C^\infty(\Omega)$, then $\psi u \in W^{k,p}_{\mathrm{loc}}(\Omega)$ and*

$$D^\alpha(\psi u) = \sum_{\beta\leq\alpha} \binom{\alpha}{\beta} D^\beta\psi\, D^{\alpha-\beta}u.$$

If all derivatives of ψ through order k are bounded on Ω, multiplication also preserves global $W^{k,p}(\Omega)$.

Proof. Linearity follows by taking linear combinations of the test identities. A test function supported in an open subset is also a test function in Ω, proving restriction. For distributions,

$$D^\alpha D^\beta T(\phi) = (-1)^{|\alpha|+|\beta|} T(D^\beta D^\alpha \phi) = D^{\alpha+\beta}T(\phi),$$

since ordinary derivatives of a smooth test function commute. Uniqueness then identifies any locally integrable representatives of these iterated derivatives.

For a first derivative, use the test function $\psi\phi$:

$$\int \psi u\, \partial_i \phi = \int u\, \partial_i(\psi\phi) - \int u(\partial_i\psi)\phi = -\int (\psi\partial_i u + u\partial_i\psi)\phi.$$

All terms are locally integrable because the smooth factors are bounded on the compact support under consideration. To obtain higher derivatives, write $\binom{\alpha}{\beta} = \prod_{l=1}^n \binom{\alpha_l}{\beta_l}$, with value zero if some β_l lies outside $[0, \alpha_l]$. Suppose the asserted formula holds for α with $|\alpha| < k$. Apply the first-derivative rule in coordinate i to each term. Differentiating the smooth factor contributes $\binom{\alpha}{\beta-e_i}$ to the coefficient of $D^\beta\psi D^{\alpha+e_i-\beta}u$; differentiating the Sobolev factor contributes $\binom{\alpha}{\beta}$. Their sum is $\binom{\alpha+e_i}{\beta}$ by the one-coordinate Pascal identity. This proves the formula for $\alpha + e_i$. Induction from order zero gives every $|\alpha| \le k$.

On a relatively compact open set, all smooth derivatives in this finite sum have finite bounds, so the displayed products are L^p there. If these bounds hold on all of Ω, each term satisfies $\|D^\beta \psi D^{\alpha-\beta} u\|_p \le \|D^\beta \psi\|_\infty \|D^{\alpha-\beta} u\|_p$. Summing proves global Sobolev membership, including the essential-supremum case. □

A general smooth multiplier need not preserve global Sobolev membership. On $(0, 1)$, the function $u = 1$ belongs to every $W^{k,p}$, while multiplication by the smooth function $\psi(x) = 1/x$ does not even preserve L^p for $p \ge 1$.

Theorem 9.8. *Every $W^{k,p}(\Omega)$ is a Banach space.*

Proof. The norm properties follow from those of the finite product of L^p spaces. Let (u_j) be Cauchy in the Sobolev norm. For every $|\alpha| \le k$, completeness of L^p gives $D^\alpha u_j \to v_\alpha$ in L^p. Put $u = v_0$. For a test function ϕ, both ϕ and $D^\alpha \phi$ belong to every $L^{p'}$ on their compact support. Thus

$$\left|\int (u_j - u) D^\alpha \phi\right| \le \|u_j - u\|_p \|D^\alpha \phi\|_{p'} \to 0,$$

and $|\int (D^\alpha u_j - v_\alpha)\phi| \le \|D^\alpha u_j - v_\alpha\|_p \|\phi\|_{p'} \to 0$. In particular,

$$\int u_j D^\alpha \phi \to \int u D^\alpha \phi, \quad \int D^\alpha u_j \,\phi \to \int v_\alpha \phi.$$

This also holds at $p = 1$ and $p = \infty$, since the test functions belong to the respective conjugate spaces. Passing to the limit in the weak identity shows $D^\alpha u = v_\alpha$. Thus $u \in W^{k,p}$, and convergence of the finitely many derivative components gives $u_j \to u$ in the Sobolev norm. □

9.3 Local and global smooth approximation

Theorem 9.9. *If $u \in W^{k,p}_{\mathrm{loc}}(\Omega)$, $1 \le p < \infty$, then its mollifications are smooth on Ω_δ, satisfy*

$$D^\alpha u_\delta = \eta_\delta * D^\alpha u \quad (|\alpha| \le k),$$

and tend to u in $W^{k,p}(V)$ for every $V \Subset \Omega$.

Proof. Smoothness and differentiation of the kernel were proved in Proposition 8.5. For $x \in \Omega_\delta$, the function $y \mapsto \eta_\delta(x - y)$ is compactly supported in Ω.

Therefore

$$\begin{aligned} D_x^\alpha u_\delta(x) &= \int u(y) D_x^\alpha \eta_\delta(x-y)\,dy \\ &= (-1)^{|\alpha|} \int u(y) D_y^\alpha \eta_\delta(x-y)\,dy \\ &= \int D^\alpha u(y) \eta_\delta(x-y)\,dy. \end{aligned}$$

The two signs cancel. For $V \Subset W \Subset \Omega$ and sufficiently small δ, all convolutions on V depend only on the derivative functions restricted to W. Extend these restrictions by zero and apply L^p approximate-identity convergence to each. There are finitely many multi-indices, so convergence of the components gives convergence in the Sobolev norm. □

The restriction $p < \infty$ is essential for norm approximation. For $u(x) = |x|$ on $(-1, 1)$, a smooth derivative cannot approximate $\operatorname{sgn} x$ in essential supremum with error below one: such a bound would force the derivative to be strictly negative immediately to the left and strictly positive immediately to the right of zero, contrary to its continuity.

Lemma 9.10. *If $u \in W^{k,p}(\Omega)$ vanishes a.e. outside a compact subset $K \Subset \Omega$, its zero extension belongs to $W^{k,p}(\mathbb{R}^n)$, with weak derivatives equal to the zero extensions of its derivatives.*

Proof. On $\Omega \setminus K$, the function is zero a.e., so the test identity and Lemma 9.2 show that all its weak derivatives are zero there. Choose a smooth cutoff $\chi \in C_c^\infty(\Omega)$ equal to one on a neighbourhood of K. For any test function ϕ on $\mathbb{R}^n$, the integrals involving u or $D^\alpha u$ are unchanged by replacing ϕ with $\chi\phi$; the additional derivatives of χ vanish near K. The weak identity for $\chi\phi$ in Ω is exactly the weak identity for the zero extensions on $\mathbb{R}^n$. Their L^p norms are unchanged. □

Theorem 9.11 (Global smooth approximation)**.** *For every open Ω, integer $k \geq 0$, and $1 \leq p < \infty$, the space $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$. No boundary regularity is assumed.*

Proof. Choose the exhaustion

$$\Omega_j = \{x \in \Omega : |x| < j,\ \operatorname{dist}(x, \Omega^c) > 1/j\}.$$

When $\Omega = \mathbb{R}^n$, interpret the distance to the empty complement as infinity. The sets increase to Ω and satisfy $\overline{\Omega_j} \Subset \Omega_{j+1}$. By Lemma 8.7, choose $0 \leq \chi_j \leq 1$ in

$C_c^\infty(\Omega_{j+1})$, equal to one on a neighbourhood of $\overline{\Omega_j}$. Put

$$\theta_1 = \chi_1, \quad \theta_j = \chi_j \prod_{i<j}(1-\chi_i) \quad (j \geq 2).$$

Then $\theta_j \geq 0$, and $\sum_{j=1}^N \theta_j = 1 - \prod_{j=1}^N (1-\chi_j)$. On Ω_N, every term with index greater than N vanishes and the sum through N is one. Hence $\sum_j \theta_j = 1$ locally finitely on Ω. Furthermore $\operatorname{supp}\theta_j$ is compact in Ω_{j+1} and, for $j \geq 2$, disjoint from Ω_{j-1}.

For $u \in W^{k,p}(\Omega)$, set $w_j = \theta_j u$ and extend it by zero. The multiplier formula and Lemma 9.10 place w_j in $W^{k,p}(\mathbb{R}^n)$. Fix $\varepsilon > 0$. Choose a separate radius $\delta_j > 0$ for each j so that

$$\|\eta_{\delta_j} * w_j - w_j\|_{W^{k,p}(\mathbb{R}^n)} < \varepsilon 2^{-j}.$$

This is possible by the derivative-transfer identity and L^p approximation on the whole space. Shrink δ_j further so that the enlarged support stays compactly inside Ω, and, for $j \geq 3$, stays disjoint from $\overline{\Omega_{j-2}}$. The required distances are positive because $\operatorname{supp}\theta_j$ is compact, lies in Ω, and avoids $\Omega_{j-1} \supset \overline{\Omega_{j-2}}$. Empty supports are omitted.

Define $v = \sum_j \eta_{\delta_j} * w_j$. This sum is locally finite: on a compact subset of some Ω_N, terms with $j \geq N+2$ vanish by the support condition. Thus $v \in C^\infty(\Omega)$.

To justify the global error, rather than assuming convergence of the partition sum in Sobolev norm, consider the error series $\sum_j(\eta_{\delta_j} * w_j - w_j)$. Its terms have summable $W^{k,p}$ norms. Completeness gives a global Sobolev sum e with $\|e\|_{W^{k,p}} \leq \varepsilon$. On each relatively compact open subset, both partition sums are locally finite and $\sum_j w_j = u$ a.e., so the series equals $v - u$ there. Consequently $e = v - u$ a.e. on Ω. Thus $v \in W^{k,p}(\Omega)$ and $\|v-u\|_{W^{k,p}} \leq \varepsilon$. □

The approximating function need not have compact support in Ω. The separate radii preserve local finiteness; using one global mollification radius near an arbitrary boundary would not achieve this conclusion.

Corollary 9.12. *For $p < \infty$, $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$.*

Proof. Choose a fixed smooth cutoff χ equal to one on $B(0,1)$ and supported in $B(0,2)$, and set $\chi_R(x) = \chi(x/R)$. The product formula gives a principal error $(\chi_R - 1)D^\alpha u$, whose L^p norm tends to zero by integrable tails, and terms bounded by $CR^{-|\beta|}|D^{\alpha-\beta}u|$ on $R \leq |x| \leq 2R$ for nonzero β. Their norms also tend to zero. Hence $\chi_R u \to u$ in $W^{k,p}$. For each fixed R, mollify this compactly supported function in the Sobolev norm, keeping compact support after a slight enlargement. Choose R and then the mollifier radius to control the two errors. $\square$

9.4 Zero boundary values and the interval case

Definition 9.13. The space $W_0^{k,p}(\Omega)$ is the closure of $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$. In particular, $H_0^1 = W_0^{1,2}$.

This is a norm-closure definition. A statement identifying it with zero traces requires a trace theorem or, on an interval, the direct argument below.

Proposition 9.14. *For $u \in W_0^{k,p}(\Omega)$, its zero extension belongs to $W^{k,p}(\mathbb{R}^n)$ and has the zero-extended weak derivatives. The extension preserves the Sobolev norm.*

Proof. Choose $u_j \in C_c^\infty(\Omega)$ converging to u in $W^{k,p}$. Their zero extensions are smooth on $\mathbb{R}^n$ and are Cauchy in its Sobolev norm, with identical differences of norms. Completeness supplies a limit. Its zeroth-order component is u on Ω and zero outside, and each derivative component is the corresponding zero-extended derivative. This proves both the identity and norm preservation, including $p = \infty$ for functions actually belonging to the norm closure. $\square$

Theorem 9.15. *For a bounded interval (a, b) and $1 \le p \le \infty$, a class belongs to $W^{1,p}(a, b)$ if and only if it has a representative*

$$U(x) = c + \int_a^x g(t)\,dt, \quad g \in L^p(a, b).$$

This representative is continuous on $[a, b]$, absolutely continuous, and unique among continuous representatives. Its a.e. derivative is g.

Proof. If $u \in W^{1,p}$ with weak derivative g, then $g \in L^1(a, b)$ by Hölder on the finite interval. Let $V(x) = \int_a^x g$. By Proposition 7.12, V is absolutely continuous and has derivative g a.e. Its weak derivative is also g: for a test function ϕ, Fubini gives

$$\int_a^b V(x)\phi'(x)\,dx = \int_a^b g(t) \int_t^b \phi'(x)\,dx\,dt = -\int_a^b g(t)\phi(t)\,dt.$$

The double integral is absolutely convergent, bounded by $(b - a)\|\phi'\|_\infty \|g\|_1$. Therefore $u - V$ has zero weak derivative and is a.e. constant by Lemma 9.3. This gives the asserted representative.

Conversely, a function of the stated form is bounded on the finite interval and thus belongs to L^p, including $p = \infty$. The same Fubini calculation gives weak derivative g. The continuity and absolute continuity follow from Proposition 7.12. If two continuous representatives agree a.e., their difference cannot be nonzero at any interior point, since continuity would make it nonzero on an interval of positive measure. Continuity at the endpoints then gives uniqueness there as well. □

For $p > 1$, Hölder gives $|U(x) - U(y)| \le \|g\|_p |x - y|^{1-1/p}$, with the Lipschitz interpretation for $p = \infty$. Also, for every x, y, $|U(x)| \le |U(y)| + \int_a^b |g|$. Averaging in y and applying Hölder yields

$$\|U\|_\infty \le (b-a)^{-1/p}\|u\|_p + (b-a)^{1-1/p}\|g\|_p. \tag{9.1}$$

Proposition 9.16. *For $1 \le p < \infty$, $u \in W_0^{1,p}(a,b)$ if and only if its continuous representative satisfies $U(a) = U(b) = 0$.*

Proof. Estimate (9.1) shows that Sobolev convergence implies uniform convergence of the continuous representatives. A limit of compactly supported smooth functions therefore has zero values at both endpoints.

Conversely, suppose these endpoint values vanish. Choose smooth cutoffs χ_δ equal to zero within distance δ of the endpoints, equal to one outside their 2δ neighbourhoods, and satisfying $|\chi_\delta'| \le C/\delta$. The terms $(1-\chi_\delta)U$ and $(1-\chi_\delta)g$ tend to zero in L^p by absolute continuity of their pth-power integrals. For the remaining derivative term, $U(x) = \int_a^x g$ near a gives

$$|U(x)|^p \le (x-a)^{p-1} \int_a^x |g(t)|^p \, dt,$$

which also holds for $p = 1$. Hence

$$\int_a^{a+2\delta} |\chi_\delta' U|^p \le C\delta^{-p} \int_a^{a+2\delta} (x-a)^{p-1}\, dx \int_a^{a+2\delta} |g|^p \le C_p \int_a^{a+2\delta} |g|^p \to 0.$$

The analogous estimate near b uses $U(x) = -\int_x^b g$. Thus $\chi_\delta U \to U$ in $W^{1,p}$. Each product has compact support inside (a,b) and can be mollified there in the Sobolev norm, producing approximation by $C_c^\infty(a,b)$. □

For $p = \infty$ the converse fails with the norm-closure definition. The function $U(x) = x(1-x)$ has zero endpoint values on $[0,1]$, but any compactly supported smooth v has $v' = 0$ near zero, whereas $U'(x) = 1 - 2x$. Thus $\|U' - v'\|_\infty \ge 1$, so $U \notin W_0^{1,\infty}(0,1)$.

Exercises

Exercise 9.1. Compute the first and second distributional derivatives of $|x|$ directly from the test identities.

Exercise 9.2. Prove the jump formula for a piecewise C^1 function with finite one-sided limits and a locally integrable ordinary derivative across the interface.

Exercise 9.3. Prove the first-order multiplier rule by one test-function calculation, and give a smooth multiplier that fails to preserve global $W^{1,p}$.

Exercise 9.4. Verify commutation of weak derivatives whenever the corresponding iterated derivatives are locally integrable.

Exercise 9.5. Prove separability of $W^{k,p}(\Omega)$ for $p < \infty$ by its isometric embedding into a finite product of L^p spaces.

Exercise 9.6. Show that no smooth sequence tends to $|x|$ in $W^{1,\infty}(-1, 1)$.

Exercise 9.7. Prove that $x(1 - x)$ is not in $W_0^{1,\infty}(0, 1)$ despite its zero endpoint values.

Exercise 9.8. Derive (9.1) with the stated constants.

Exercise 9.9. On the union of two disjoint bounded balls, construct a Sobolev function with zero weak gradient that is not a.e. constant on the whole set.

Exercise 9.10. Prove that every weak derivative of a Sobolev function vanishes a.e. on any open subset where the function vanishes a.e.

Exercise 9.11. Verify the support and local finiteness assertions for θ_j and its mollified products in Theorem 9.11.

Exercise 9.12. Show that the zero extension of the constant one on $(0, 1)$ is not in $W^{1,1}(\mathbb{R})$.

Exercise 9.13. After the absolute-continuity characterization in Chapter 12, prove that a locally Lipschitz function has coordinate weak derivatives equal to its a.e. ordinary coordinate derivatives.

Exercise 9.14. For $u \in W^{1,p}(\mathbb{R}^n)$, $p < \infty$, prove $\|\tau_h u - u\|_p \leq |h| \|\nabla u\|_p$ by smooth approximation.

Chapter 10

Sobolev inequalities and compactness

The Sobolev exponent is determined by scaling, but scaling does not prove the inequality. The proof begins at $p = 1$, where one-dimensional integration in each coordinate is combined with a product estimate. The other exponents follow by applying that endpoint inequality to a power of the function.

10.1 Scaling and the Sobolev inequality

Suppose $\|u\|_q \le C\|\nabla u\|_p$ is to hold on $\mathbb{R}^n$ for all compactly supported smooth u, with one constant independent of the size of the support. Put $u_\lambda(x) = u(\lambda x)$. The change-of-variables formula gives

$$\|u_\lambda\|_q = \lambda^{-n/q}\|u\|_q, \quad \|\nabla u_\lambda\|_p = \lambda^{1-n/p}\|\nabla u\|_p.$$

For a fixed nonzero u, sending λ to zero or infinity forces $1 - n/p + n/q = 0$. Thus, when $p < n$, the only possible finite exponent is $p^* = np/(n-p)$.

Lemma 10.1. *Let $n \ge 2$. For $i = 1, \dots, n$, let $F_i \ge 0$ be an integrable function of all coordinates except x_i. Then*

$$\int_{\mathbb{R}^n} \prod_{i=1}^{n} F_i(\widehat{x}_i)^{1/(n-1)}\, dx \le \prod_{i=1}^{n} \Big(\int_{\mathbb{R}^{n-1}} F_i\Big)^{1/(n-1)},$$

where $\widehat{x}_i$ denotes omission of the ith coordinate.

Proof. First use bounded functions with bounded support in their own variable spaces; increasing truncation and monotone convergence then give the general case. For $n = 2$, the assertion is $\int F_1(x_2)F_2(x_1)\,dx_1\,dx_2 = (\int F_1)(\int F_2)$, which is Tonelli.

Assume the statement in dimension $n-1$, with $n \geq 3$. Integrate first in x_n and apply Hölder to the $n-1$ factors indexed by $i < n$, each with exponent $n-1$. Put $G_i = \int_{\mathbb{R}} F_i\,dx_n$; this is a function of $n-2$ variables, omitting x_i among $x_1, \ldots, x_{n-1}$. The original integral is bounded by

$$\int_{\mathbb{R}^{n-1}} F_n^{1/(n-1)} \prod_{i<n} G_i^{1/(n-1)}.$$

Set $P = \prod_{i<n} G_i^{1/(n-2)}$. Another application of Hölder, with exponents $n-1$ and $(n-1)/(n-2)$, bounds this by

$$\Big(\int F_n\Big)^{1/(n-1)} \Big(\int P\Big)^{(n-2)/(n-1)}.$$

The induction hypothesis gives $\int P \leq \prod_{i<n}(\int G_i)^{1/(n-2)}$. By Tonelli, $\int G_i = \int F_i$. Substitution yields exactly the desired product and completes the induction. □

Theorem 10.2 (Sobolev inequality). *If $n \geq 2$ and $1 \leq p < n$, then*

$$\|u\|_{p^*} \leq C_{n,p}\|\nabla u\|_p, \quad p^* = \frac{np}{n-p}, \tag{10.1}$$

for every $u \in C_c^1(\mathbb{R}^n)$. It extends to $W^{1,p}(\mathbb{R}^n)$ and to zero extensions of $W_0^{1,p}(\Omega)$.

Proof. First take $p = 1$. Fix all coordinates except x_i. Compact support and the one-dimensional fundamental theorem give

$$|u(x)| \leq \int_{\mathbb{R}} |\partial_i u(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)|\,dt =: F_i(\widehat{x}_i).$$

Multiply these n inequalities after raising each to the power $1/(n-1)$. Then $|u|^{n/(n-1)} \leq \prod_i F_i^{1/(n-1)}$. Lemma 10.1 gives

$$\int |u|^{n/(n-1)} \leq \prod_i \|\partial_i u\|_1^{1/(n-1)}.$$

Taking the power $(n-1)/n$ gives

$$\|u\|_{n/(n-1)} \leq \prod_{i=1}^n \|\partial_i u\|_1^{1/n} \leq \|\nabla u\|_1,$$

because $|\partial_i u| \le |\nabla u|$ pointwise for every i. A larger dimension-dependent constant is harmless in the stated inequality.

Let $1 < p < n$ and put $\gamma = p(n-1)/(n-p) > 1$. Apply the $p = 1$ result to $v = |u|^\gamma$. This is a compactly supported C^1 function even at zeros of u, because $\gamma > 1$, and $|\nabla v| \le \gamma |u|^{\gamma-1}|\nabla u|$. Thus

$$\left(\int |u|^{\gamma n/(n-1)}\right)^{(n-1)/n} \le C_n \gamma \int |u|^{\gamma-1}|\nabla u|.$$

The exponent calculations are

$$\frac{\gamma n}{n-1} = \frac{np}{n-p} = p^*, \quad \gamma - 1 = \frac{n(p-1)}{n-p}, \quad (\gamma-1)p' = \frac{np}{n-p} = p^*.$$

Consequently $\| |u|^{\gamma-1} \|_{p'} = \|u\|_{p^*}^{\gamma-1}$, and the left side is $\|u\|_{p^*}^{\gamma}$. Hölder therefore rewrites the inequality as

$$\|u\|_{p^*}^{\gamma} \le C_n \gamma \|u\|_{p^*}^{\gamma-1} \|\nabla u\|_p.$$

If $u = 0$, the result is trivial. Otherwise its L^{p^*} norm is positive and finite because of smoothness and compact support. Divide by its $(\gamma-1)$st power to obtain (10.1).

For $u \in W^{1,p}(\mathbb{R}^n)$, choose compactly supported smooth $u_j \to u$ in that norm by Corollary 9.12. A subsequence converges a.e. Fatou and convergence of the gradient norms give (10.1) for u. Applying the inequality to differences also shows convergence in L^{p^*}. For $u \in W_0^{1,p}(\Omega)$, first extend by zero using Proposition 9.14 and apply the whole-space result. □

In one dimension the endpoint statement is $\|u\|_\infty \le \|u'\|_1$ for $u \in C_c^1(\mathbb{R})$, directly from the fundamental theorem. This differs from the critical behaviour in dimensions $n \ge 2$ discussed below.

10.2 Poincaré inequalities and ball averages

Proposition 10.3. *If a bounded open set Ω lies in a strip $a < x_1 < b$, then for $1 \le p < \infty$ and $u \in W_0^{1,p}(\Omega)$,*

$$\|u\|_p \le (b-a)\|\nabla u\|_p.$$

Proof. For $u \in C_c^\infty(\Omega)$, extend by zero. For fixed $x' = (x_2, \dots, x_n)$ and $a < x_1 < b$,

$$|u(x_1, x')| \le \int_a^b |\partial_1 u(t, x')| \, dt.$$

Hölder on this interval gives $|u(x_1,x')|^p \le (b-a)^{p-1}\int_a^b |\partial_1 u(t,x')|^p\,dt$, including $p=1$. Integrate in x_1 and x' using Tonelli. The resulting estimate is $\|u\|_p^p \le (b-a)^p\|\partial_1 u\|_p^p$. Pass to the norm closure to obtain the assertion for $W_0^{1,p}$. □

This proof does not assume that an arbitrary Sobolev function can be extended by zero. That extension is justified here by the zero-boundary norm-closure hypothesis.

Lemma 10.4. *For a ball $B = B(c,r)$, $u \in C^1(\overline{B})$, and $x \in B$,*

$$\frac{1}{m(B)}\int_B |u(x)-u(y)|\,dy \le C_n \int_B \frac{|\nabla u(z)|}{|x-z|^{n-1}}\,dz.$$

Consequently $|u(x)-u_B|$ has the same upper bound, where $u_B = m(B)^{-1}\int_B u$.

Proof. The line segment from x to y lies in the ball. The fundamental theorem gives

$$|u(x)-u(y)| \le \int_0^1 |\nabla u(x+t(y-x))||y-x|\,dt.$$

Integrate in y and use Tonelli. For fixed $t>0$, set $z = x+t(y-x)$; the Jacobian is t^n and $|y-x| = |z-x|/t$. Thus the double integral is

$$\int_0^1 t^{-n-1}\int_{x+t(B-x)} |\nabla u(z)||z-x|\,dz\,dt.$$

The inner domain lies in B by convexity. If z belongs to it, then $|z-x| \le 2rt$, so $t \ge |z-x|/(2r)$. Enlarge the range of integration in t to this interval and use

$$\int_{|z-x|/(2r)}^1 t^{-n-1}\,dt \le \frac{(2r)^n}{n|z-x|^n}.$$

After division by $m(B) = \omega_n r^n$, this proves the stated kernel bound. The diagonal $z=x$ is null and does not affect the calculation. Finally, $|u(x)-u_B| \le m(B)^{-1}\int_B |u(x)-u(y)|\,dy$ by the integral triangle inequality. □

Theorem 10.5. *For $1 \le p < \infty$ and $u \in W^{1,p}(B(c,r))$,*

$$\|u-u_B\|_{L^p(B)} \le C_n r\|\nabla u\|_{L^p(B)}.$$

Proof. For $u \in C^1(\overline{B})$, the preceding lemma bounds $|u-u_B|$ by a convolution of the zero extension of $|\nabla u|$ with $K_r(z) = C_n|z|^{1-n}\mathbf{1}_{\{|z|<2r\}}$. The kernel has L^1 norm at most $C_n r$. This can be seen without a polar-coordinate formula by summing over annuli $2r\,2^{-k-1} \le |z| < 2r\,2^{-k}$: their contributions are bounded by $C_n r 2^{-k}$. Young's inequality gives the result.

For a smooth function inside B that belongs to $W^{1,p}(B)$ but need not be smooth on the boundary, apply the result to concentric closed inner balls of radii $\rho < r$. Their averages tend to u_B by integrability, and Fatou as $\rho \uparrow r$ gives the same estimate on B. Finally, use Theorem 9.11 to approximate arbitrary $u \in W^{1,p}(B)$ by smooth Sobolev functions on B. Their averages converge by Hölder, and the two sides of the inequality pass to the limit in norm. □

10.3 Morrey's inequality and the critical case

We record the singular-kernel calculation used in the next proof. If $0 \le \alpha < n$ and $R > 0$, split $B(0,R) \setminus \{0\}$ into $A_j = \{2^{-j-1}R \le |z| < 2^{-j}R\}$, $j \ge 0$. On A_j, $|z|^{-\alpha} \le (2^{j+1}/R)^\alpha$, while $m(A_j) \le \omega_n (2^{-j}R)^n$. The null singleton at the origin does not affect the integral. Therefore

$$\int_{|z|<R} |z|^{-\alpha}\, dz \le 2^\alpha \omega_n R^{n-\alpha} \sum_{j=0}^{\infty} 2^{-j(n-\alpha)} \le C_{n,\alpha} R^{n-\alpha}.$$

For $n < p < \infty$, take $\alpha = (n-1)p'$. The inequality $\alpha < n$ is equivalent to $p > n$, and $(n-\alpha)/p' = n/p' - (n-1) = 1 - n/p$. Hence

$$\big\| |z|^{1-n} \mathbf{1}_{B(0,R)} \big\|_{p'} \le C_{n,p} R^{1-n/p}.$$

This supplies both the integrability and the exact radius power without assuming a polar-coordinate formula.

Theorem 10.6 (Morrey). *For $p > n$, $p < \infty$, every $u \in W^{1,p}(\mathbb{R}^n)$ has a continuous representative U satisfying*

$$|U(x) - U(y)| \le C_{n,p} |x-y|^{1-n/p} \|\nabla u\|_p, \quad \|U\|_\infty \le C_{n,p} \|u\|_{W^{1,p}}.$$

For $p = \infty$, there is a Lipschitz representative with Lipschitz constant bounded by $\|\nabla u\|_\infty$ when the gradient is measured in Euclidean norm.

Proof. Start with smooth u. Hölder in Lemma 10.4 gives

$$\frac{1}{m(B)} \int_B |u(x) - u(z)|\, dz \le C_{n,p} r^{1-n/p} \|\nabla u\|_{L^p(B)} \quad (x \in B).$$

The kernel power is integrable because $(n-1)p' < n$, which is equivalent to $p > n$; its $L^{p'}$ norm on a ball of radius $2r$ scales as $r^{1-n/p}$. At $n = 1$ the kernel is constant and the same scaling follows directly.

For distinct x, y, put $r = |x - y|$ and $W = B(x, r) \cap B(y, r)$. This intersection contains the ball with centre $(x + y)/2$ and radius $r/2$, so $m(W) \geq \omega_n (r/2)^n$. Average

$$|u(x) - u(y)| \leq |u(x) - u(z)| + |u(y) - u(z)|$$

over $z \in W$. Each average is bounded by a dimension-dependent multiple of the corresponding average over $B(x, r)$ or $B(y, r)$. Applying the preceding estimate proves the Hölder bound. To control the supremum, use a ball of radius one centred at x:

$$|u(x)| \leq \frac{1}{m(B(x,1))} \int_{B(x,1)} |u(z)|\, dz + \frac{1}{m(B(x,1))} \int_{B(x,1)} |u(x) - u(z)|\, dz.$$

Hölder bounds the first term by $C_n \|u\|_p$ and the kernel estimate bounds the second by $C_{n,p} \|\nabla u\|_p$.

For general $u \in W^{1,p}(\mathbb{R}^n)$, choose smooth compactly supported u_j converging in Sobolev norm. The supremum estimate applied to differences shows that u_j is uniformly Cauchy on $\mathbb{R}^n$. Let U be its continuous uniform limit. A subsequence also tends to u a.e. by L^p convergence, so U represents u. Passing to the limit for fixed x, y gives the Hölder estimate, and the supremum bound passes as well.

For $p = \infty$, use mollifications u_δ. They satisfy $\|u_\delta\|_\infty \leq \|u\|_\infty$ and $|\nabla u_\delta| \leq \|\nabla u\|_\infty$ by the derivative-transfer identity and positivity of the kernel. Integrating on a segment shows that they have a common Lipschitz constant. Choose a sequence $\delta_j \downarrow 0$ and extract a diagonal subsequence converging at every point of a countable dense set. The common Lipschitz bound extends this limit uniquely and continuously to $\mathbb{R}^n$; finite nets show that convergence is uniform on compact sets. Local L^1 convergence of mollifiers identifies the limit with u a.e. The Lipschitz bound passes to the limit. This argument does not assume smooth density in $W^{1,\infty}$. $\square$

For later radial calculations, the identity $m(B(0, r)) = \omega_n r^n$ gives

$$\int_{\mathbb{R}^n} h(|x|)\, dx = n\omega_n \int_0^\infty h(r) r^{n-1}\, dr$$

for nonnegative continuous functions supported in annuli, by radial step approximation and the measures of shells. Monotone exhaustion of the annuli extends it to the nonnegative singular radial functions used below.

Example 10.7 (Failure of boundedness at the critical exponent)**.** For $n \geq 2$, let $\chi \in C_c^\infty(B(0, 1/2))$ equal one near zero, and set

$$u(x) = \chi(x) \log\log(e/|x|), \quad x \neq 0,$$

with any finite value assigned at zero. This function is unbounded near zero but belongs to $W^{1,n}(\mathbb{R}^n)$. Indeed, near zero the magnitude of its ordinary gradient is $1/(|x|\log(e/|x|))$. The radial formula and the substitution $t = \log(e/r)$ give

$$\int_0^{\varepsilon} \frac{r^{n-1}}{r^n \log(e/r)^n}\, dr = \int_{\log(e/\varepsilon)}^{\infty} t^{-n}\, dt < \infty.$$

The function itself is in L^n, because its radial integral is bounded by a constant times $\int^{\infty} e^{-nt}(\log t)^n\, dt$.

To verify that the ordinary gradient away from zero is the weak gradient, use a smooth radial cutoff ρ_ε vanishing on $B(0,\varepsilon)$ and equal to one outside $B(0,2\varepsilon)$, with $|\nabla\rho_\varepsilon| \le C/\varepsilon$. Integrate by parts for u against $\rho_\varepsilon\phi$, whose support avoids zero. The extra term is bounded by

$$C\varepsilon^{-1}\int_{\varepsilon<|x|<2\varepsilon} |u(x)|\, dx \le C\varepsilon^{n-1}\log\log(e/\varepsilon) \longrightarrow 0.$$

The remaining terms converge by local integrability of u and its candidate gradient. Thus the weak identity holds. This example shows that the $p > n$ condition in Morrey's boundedness assertion cannot simply be replaced by $p = n$.

10.4 Compactness from translations

Lemma 10.8. *For $u \in W^{1,p}(\mathbb{R}^n)$, $1 \le p < \infty$,*

$$\|\tau_h u - u\|_p \le |h|\|\nabla u\|_p.$$

Proof. For smooth u, write $u(x-h) - u(x) = -\int_0^1 \nabla u(x-th)\cdot h\, dt$. Hölder on the unit t interval gives

$$|u(x-h)-u(x)|^p \le |h|^p \int_0^1 |\nabla u(x-th)|^p\, dt.$$

Integrate in x by Tonelli and use translation invariance. For general u, approximate by C_c^∞ in $W^{1,p}$ and pass to the limit, using translation isometry. □

Lemma 10.9. *A uniformly bounded equicontinuous sequence of continuous functions on a compact subset of $\mathbb{R}^n$ has a uniformly convergent subsequence.*

Proof. Choose a countable dense subset by taking the union of finite $1/k$ nets. Successive subsequence extraction at these points, followed by a diagonal choice, gives a subsequence whose values converge at every point of the dense subset. Given $\varepsilon > 0$, equicontinuity gives one radius δ controlling the oscillation of every function by $\varepsilon/3$. A finite δ net from the dense set covers the compact set. For sufficiently late indices, the values at all finitely many net points differ by less than $\varepsilon/3$. The triangle inequality then bounds the difference everywhere by ε. Thus the subsequence is uniformly Cauchy, and its uniform limit exists and is continuous. □

Theorem 10.10. *If Ω is bounded and open and $1 \le p < \infty$, every bounded sequence in $W_0^{1,p}(\Omega)$ has a subsequence converging in $L^p(\Omega)$. If $p < n$, convergence can be obtained in every $L^q(\Omega)$ with $1 \le q < p^*$.*

Proof. Extend the sequence u_j by zero to $\mathbb{R}^n$. Its Sobolev norms remain bounded, and all extensions vanish outside the fixed compact set $\overline{\Omega}$. Positivity of the mollifier, the weighted Hölder argument from Theorem 8.2, and Lemma 10.8 give the uniform estimate

$$\|\eta_\varepsilon * u_j - u_j\|_p^p \le \int \eta_\varepsilon(y)\|\tau_y u_j - u_j\|_p^p \, dy \le \varepsilon^p \|\nabla u_j\|_p^p.$$

For each fixed $\varepsilon > 0$, Hölder bounds $\eta_\varepsilon * u_j$ and its first derivatives uniformly in x, j by $\|\eta_\varepsilon\|_{p'}\|u_j\|_p$ and $\|\nabla\eta_\varepsilon\|_{p'}\|u_j\|_p$. Their supports lie in one compact ε-neighbourhood of $\overline{\Omega}$. Lemma 10.9 therefore gives a uniformly convergent, hence L^p convergent, subsequence of these mollifications.

Choose $\varepsilon_k \downarrow 0$ and use a diagonal extraction so that the mollifications at each fixed scale converge along one subsequence. To see that the original subsequence is L^p Cauchy, first choose k making both approximation errors less than $\eta/3$ uniformly in the sequence. Then choose two indices sufficiently late that their mollifications at this fixed scale differ in L^p by less than $\eta/3$. The triangle inequality gives an original difference less than η. Completeness of L^p proves convergence.

If $p < n$, the Sobolev inequality bounds the sequence uniformly in L^{p^*}. Interpolation between its strong L^p convergence and the uniform endpoint bound gives strong L^q convergence for $p < q < p^*$. For $1 \le q \le p$, use the finite-measure inclusion. The same subsequence works for each such exponent. □

For $p \ge n$ in dimensions $n \ge 2$, finite-measure Hölder gives a uniform $W^{1,p_0}(\Omega)$ bound for every $1 \le p_0 < n$. Approximation of each member by

$C_c^\infty(\Omega)$ in $W^{1,p}$ also gives approximation in W^{1,p_0}, so the members lie in $W_0^{1,p_0}(\Omega)$. For any specified finite q, choose $p_0 < n$ sufficiently close to n that $q < np_0/(n-p_0)$. The theorem then supplies a subsequence converging in L^q.

In dimension one, choose a bounded interval $[a,b]$ containing $\overline{\Omega}$ in its interior. The zero extensions satisfy, first for compactly supported smooth functions and then by Sobolev approximation,

$$\|u_j\|_\infty \le \|u_j'\|_{L^1(\mathbb{R})} \le (b-a)^{1-1/p}\|u_j'\|_{L^p(\mathbb{R})}.$$

The first inequality follows from $u_j(x) = \int_a^x u_j'(t)\,dt$ for the continuous representative. Thus the sequence is uniformly bounded in L^∞. The subsequence already converging in L^p has a further subsequence converging almost everywhere, so its limit obeys the same L^∞ bound. If $q > p$, the estimate $\|u_j - u\|_q^q \le \|u_j - u\|_\infty^{q-p}\|u_j - u\|_p^p$ proves convergence in L^q; if $q \le p$, use the finite-measure inclusion. No regularity of the boundary of Ω is needed for these zero-boundary statements.

For $n \ge 2$ and $1 \le p < n$, compactness at the critical exponent generally fails. For a fixed nonzero $\phi \in C_c^\infty(B(0,1))$ and a point $x_0 \in \Omega$, the functions

$$u_\lambda(x) = \lambda^{(n-p)/p}\phi(\lambda(x-x_0))$$

belong to $W_0^{1,p}(\Omega)$ for large λ. Their gradient L^p norms and their L^{p^*} norms are constant, while their L^p norms tend to zero. An L^{p^*} convergent subsequence would have limit zero by local convergence in measure, contradicting the fixed positive L^{p^*} norm.

Exercises

Exercise 10.1. Derive the Sobolev exponent solely from dilation, and explain why this does not establish the inequality.

Exercise 10.2. Write out the product-integral lemma in dimensions two and three, identifying both Hölder applications.

Exercise 10.3. Check the two exponent identities used when the endpoint Sobolev inequality is applied to $|u|^\gamma$.

Exercise 10.4. Prove the one-dimensional inequality $\|u\|_\infty \le \|u'\|_1$ for compactly supported Sobolev functions.

Exercise 10.5. Prove Poincaré's inequality for zero-boundary functions in a bounded strip, including the L^∞ case when membership in the norm closure is assumed.

Exercise 10.6. Compute the L^1 bound for the truncated kernel $|x|^{1-n}\mathbf{1}_{\{|x|<2r\}}$ by dyadic annuli.

Exercise 10.7. Extend the mean-zero Poincaré inequality on a ball to $p = \infty$ by local mollification and passage to a.e. limits.

Exercise 10.8. Verify the geometric lower bound for the intersection of two equal-radius balls whose centres are one radius apart.

Exercise 10.9. Prove the weak-derivative assertion in Example 10.7 using the explicit radial cutoffs.

Exercise 10.10. Show that the critical example is not essentially bounded, not merely unbounded at its assigned value at zero.

Exercise 10.11. Prove uniform mollification error bounds for a family with a common Sobolev norm bound.

Exercise 10.12. In the compactness proof, explain why fixing the mollification scale before extracting a subsequence is necessary.

Exercise 10.13. Verify the norm scalings of the concentrating family u_λ and its failure of critical compactness.

Exercise 10.14. Show that compact support in a fixed bounded set is needed for the compactness argument by considering translations of one nonzero smooth function on $\mathbb{R}^n$.

Chapter 11

Hilbert spaces and Fourier series

The completeness of L^2 allows orthogonal expansions to be treated as convergent series in a function space. We first prove the geometric results that make this possible. On the circle, the remaining question is whether the exponentials form a complete orthonormal system. Approximation by Fejér kernels will answer this question without assuming pointwise convergence of ordinary Fourier series.

11.1 Inner products and orthogonal projection

Definition 11.1. An inner product on a complex vector space H is a function $\langle u, v\rangle$ that is linear in u, satisfies $\langle u, v\rangle = \overline{\langle v, u\rangle}$, and has $\langle u, u\rangle > 0$ for $u \neq 0$. Its associated norm is $\|u\| = \langle u, u\rangle^{1/2}$. A Hilbert space is an inner product space complete in this norm. In the real case the conjugates are omitted. Separability is not included in the definition.

Proposition 11.2 (Cauchy–Schwarz and the parallelogram identity)**.** *For $u, v \in H$,*

$$|\langle u, v\rangle| \leq \|u\|\|v\|, \quad \|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

In particular, the associated norm satisfies the triangle inequality.

Proof. If $v = 0$, the first assertion is immediate. Otherwise, put $a = \langle u, v\rangle/\|v\|^2$. Expanding the nonnegative quantity $\|u - av\|^2$ gives

$$0 \leq \|u\|^2 - \frac{|\langle u, v\rangle|^2}{\|v\|^2}.$$

This proves Cauchy–Schwarz. Expansion of the two squares gives the parallelogram identity, since their mixed terms cancel. Finally,

$$\|u+v\|^2 = \|u\|^2 + 2\operatorname{Re}\langle u,v\rangle + \|v\|^2 \le (\|u\| + \|v\|)^2.$$

Taking square roots proves the triangle inequality. □

For a subset $V \subset H$, write $V^\perp = \{u \in H : \langle u,v\rangle = 0 \text{ for all } v \in V\}$. The continuity of each map $u \mapsto \langle u,v\rangle$ shows that $V^\perp$ is a closed linear subspace.

Theorem 11.3 (Orthogonal projection)**.** *Let V be a closed linear subspace of a Hilbert space H. For every $x \in H$ there is a unique $v \in V$ minimizing $\|x-v\|$. It is characterized by $x - v \in V^\perp$. Consequently,*

$$H = V \oplus V^\perp.$$

The map $P_V x = v$ is linear and satisfies $\|P_V x\| \le \|x\|$.

Proof. Set $d = \inf_{w\in V}\|x-w\|$ and choose $v_j \in V$ with $\|x-v_j\| \to d$. Applying the parallelogram identity to $x - v_j$ and $x - v_k$, and using $(v_j+v_k)/2 \in V$, gives

$$\begin{aligned}\|v_j - v_k\|^2 &= 2\|x-v_j\|^2 + 2\|x-v_k\|^2 - 4\left\|x - \frac{v_j+v_k}{2}\right\|^2 \\ &\le 2\|x-v_j\|^2 + 2\|x-v_k\|^2 - 4d^2.\end{aligned}$$

The right side tends to zero as $j,k \to \infty$. Thus (v_j) is Cauchy. Completeness gives a limit $v \in H$, and closedness of V gives $v \in V$. Continuity of the norm gives $\|x-v\| = d$.

For $w \in V$ and $t \in \mathbb{R}$, minimality implies

$$0 \le \|x-v-tw\|^2 - \|x-v\|^2 = -2t\operatorname{Re}\langle x-v,w\rangle + t^2\|w\|^2.$$

Divide by positive t and let $t \downarrow 0$, and then use negative t and let $t \uparrow 0$. The two inequalities give $\operatorname{Re}\langle x-v,w\rangle = 0$. Applying the same conclusion with iw in place of w gives the vanishing of the imaginary part. Hence $x - v \in V^\perp$.

Conversely, if $v \in V$ and $x - v \in V^\perp$, then for every $w \in V$,

$$\|x-w\|^2 = \|x-v\|^2 + \|v-w\|^2.$$

This proves both minimality and uniqueness. Adding or multiplying the unique orthogonal decompositions proves linearity of P_V. The identity $\|x\|^2 = \|P_V x\|^2 + \|x - P_V x\|^2$ proves the norm bound. □

Theorem 11.4 (Riesz representation). *If $L : H \to \mathbb{C}$ is a bounded linear functional, there is a unique $y \in H$ such that $L(x) = \langle x, y\rangle$ for all $x \in H$. Moreover, $\|L\| = \|y\|$.*

Proof. For $L = 0$, take $y = 0$. Otherwise, $V = \ker L$ is a proper closed subspace. Choose z with $L(z) \neq 0$ and put $w = z - P_V z$. Then $w \in V^\perp$ and $L(w) = L(z) \neq 0$. In particular, $w \neq 0$. Set $e = w/\|w\|$. For any $x \in H$,

$$x - \frac{L(x)}{L(e)} e \in V.$$

Taking the inner product with e gives $\langle x, e\rangle = L(x)/L(e)$. Thus

$$L(x) = L(e)\langle x, e\rangle = \langle x, \overline{L(e)} e\rangle.$$

The conjugate in the representing vector is required because the inner product is linear in its first argument. If two vectors represent L, their difference is orthogonal to every x, and taking x equal to that difference proves uniqueness. Cauchy–Schwarz gives $\|L\| \leq \|y\|$. When $y \neq 0$, evaluation at $x = y/\|y\|$ gives the reverse inequality. □

11.2 Orthonormal systems and expansions

A sequence (e_j) is orthonormal if $\langle e_j, e_k\rangle = \delta_{jk}$. For $x \in H$ define $a_j = \langle x, e_j\rangle$ and $S_N x = \sum_{j=1}^N a_j e_j$. Direct computation gives $x - S_N x \perp e_k$ for $k \leq N$, and therefore

$$\|x\|^2 = \|x - S_N x\|^2 + \sum_{j=1}^{N} |a_j|^2. \tag{11.1}$$

In particular, $\sum_j |a_j|^2 \leq \|x\|^2$. This is Bessel's inequality.

Theorem 11.5 (Completeness and Parseval). *For an orthonormal sequence (e_j) in a Hilbert space, the following are equivalent: its finite linear combinations are dense; the only vector orthogonal to every e_j is zero; $S_N x \to x$ for every $x \in H$; and*

$$\|x\|^2 = \sum_{j=1}^{\infty} |\langle x, e_j\rangle|^2 \quad \textit{for every } x \in H.$$

Under these conditions, (e_j) is called an orthonormal basis.

Proof. Suppose finite linear combinations are dense and $x \perp e_j$ for all j. Choose such combinations $v_k \to x$. Then $\langle x, v_k\rangle = 0$ and

$$\|x\|^2 = \langle x, x - v_k\rangle \leq \|x\|\|x - v_k\| \longrightarrow 0.$$

Thus $x = 0$.

Now suppose the only common orthogonal vector is zero. Bessel's inequality implies, for $N > M$,

$$\|S_N x - S_M x\|^2 = \sum_{j=M+1}^{N} |a_j|^2 \longrightarrow 0.$$

Completeness gives $S_N x \to y$ for some $y \in H$. Fix j. For $N \geq j$ one has $\langle x - S_N x, e_j\rangle = 0$, so passage to the limit gives $\langle x - y, e_j\rangle = 0$. The assumption implies $y = x$. Formula (11.1) now gives Parseval's identity. Conversely, Parseval and the same formula imply $\|x - S_N x\| \to 0$, which gives density. □

The sequence space ℓ^2 consists of complex sequences $a = (a_j)_{j\geq 1}$ with $\sum_j |a_j|^2 < \infty$, and its inner product is $\langle a, b\rangle = \sum_j a_j \overline{b_j}$. The series is absolutely convergent: Cauchy–Schwarz for finite sums bounds each partial sum of $\sum_j |a_j b_j|$ by $\|a\|_{\ell^2}\|b\|_{\ell^2}$. Its norm is consequently $\|a\|_{\ell^2} = (\sum_j |a_j|^2)^{1/2}$.

Here is a direct verification of completeness. If $(a^{(k)})$ is Cauchy in ℓ^2, each coordinate is Cauchy since $|a_j^{(k)} - a_j^{(l)}| \leq \|a^{(k)} - a^{(l)}\|_{\ell^2}$. Let a_j be its limit. The norms of the sequence are bounded by some M. For every finite N, passage to the limit in the finite sum gives $\sum_{j=1}^{N} |a_j|^2 \leq M^2$. Taking the supremum over N shows $a \in \ell^2$. Given $\varepsilon > 0$, choose K so that $\|a^{(k)} - a^{(l)}\|_{\ell^2} < \varepsilon$ for $k, l \geq K$. For fixed $k \geq K$, let $l \to \infty$ in the first N coordinates and then $N \to \infty$. This gives $\|a^{(k)} - a\|_{\ell^2} \leq \varepsilon$. Thus ℓ^2 is a Hilbert space.

Proposition 11.6. *Every separable Hilbert space has a finite or countable orthonormal basis. Every sequence $(a_j) \in \ell^2$ is the coefficient sequence of a unique vector relative to a countable orthonormal basis.*

Proof. Take a countable dense sequence (x_j). Suppose that $e_1, \ldots, e_N$ have already been constructed, with $N = 0$ allowed at the beginning. Define

$$r_j = x_j - \sum_{k=1}^{N} \langle x_j, e_k\rangle e_k.$$

For $1 \leq l \leq N$, orthonormality gives $\langle r_j, e_l\rangle = \langle x_j, e_l\rangle - \langle x_j, e_l\rangle = 0$. If $r_j = 0$, discard x_j. Otherwise append $r_j/\|r_j\|$ to the orthonormal list. In either case, after this stage x_j belongs to the span of the vectors constructed so far. The closed span of the resulting finite or countable family therefore contains the dense set $\{x_j : j \geq 1\}$ and equals H.

If only finitely many vectors are obtained, their span is closed. To verify this rather than assume it, let $v_l = \sum_{k=1}^{N} c_{l,k} e_k$ converge in H to v. Continuity of the inner product gives $c_{l,k} = \langle v_l, e_k\rangle \to \langle v, e_k\rangle$. A finite sum may be passed to the limit, so $v = \sum_{k=1}^N \langle v, e_k\rangle e_k$ belongs to the same span.

Given $(a_j) \in \ell^2$ and a countable orthonormal basis, put $v_N = \sum_{j=1}^N a_j e_j$. For $N > M$, orthonormality gives $\|v_N - v_M\|^2 = \sum_{j=M+1}^N |a_j|^2$, which tends to zero with M, N. Completeness gives $v_N \to v \in H$. For each fixed j, the equality $\langle v_N, e_j\rangle = a_j$ holds for all $N \geq j$, and continuity gives $\langle v, e_j\rangle = a_j$. If w has the same coefficients, $v - w$ is orthogonal to every basis vector and is zero by Theorem 11.5. This proves uniqueness. □

It follows that every infinite dimensional separable Hilbert space is unitarily isomorphic to ℓ^2: the isomorphism sends x to $(\langle x, e_j\rangle)_j$. The separability qualification must not be omitted. Here a unitary map means a surjective linear map preserving the inner product.

For a complex inner product linear in the first argument, polarization reads

$$\langle x, y\rangle = \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\big).$$

Expansion of the four squares proves the identity. Thus a linear norm-preserving map also preserves the inner product. The coefficient map above is linear, norm preserving by Parseval, and surjective by the proposition, so it is unitary. Applying two such coefficient maps also proves that any two infinite dimensional separable Hilbert spaces are unitarily equivalent.

An inner product space not assumed complete is also called a pre-Hilbert space. The following construction supplies the completeness needed for orthogonal series.

Auxiliary lemma (Completion of an inner product space). Every inner product space H_0 admits an isometric linear embedding into a Hilbert space H whose image is dense. Two such completions are related by a unique unitary map that agrees with the identity on the embedded copy of H_0.

Proof. Consider all Cauchy sequences in H_0. Declare $(x_j) \sim (y_j)$ when $\|x_j - y_j\| \to 0$, and let H be the set of equivalence classes. The triangle inequality proves transitivity; reflexivity and symmetry follow directly. Addition and scalar multiplication term by term preserve Cauchy sequences and respect the equivalence relation.

Define an inner product on these classes by

$$\langle [(x_j)], [(y_j)]\rangle_H = \lim_{j\to\infty} \langle x_j, y_j\rangle_{H_0}.$$

The limit exists. Both Cauchy sequences are bounded, and

$$|\langle x_j, y_j\rangle - \langle x_k, y_k\rangle| \leq \|x_j - x_k\|\,\|y_j\| + \|x_k\|\,\|y_j - y_k\| \longrightarrow 0.$$

The same estimate, with equivalent representatives in place of x_k, y_k, shows independence of representatives. Linearity and conjugate symmetry pass to the limit. Moreover, $\|[(x_j)]\|_H = \lim_j \|x_j\|$, so zero norm means equivalence to the zero sequence. This proves positive definiteness.

Map $x \in H_0$ to the constant sequence, denoted by Jx. This is a linear isometry. If $X = [(x_j)]$, then

$$\|Jx_k - X\|_H = \lim_{j\to\infty} \|x_k - x_j\| \longrightarrow 0 \quad \text{as } k \to \infty,$$

by the Cauchy property of (x_j). Hence $J(H_0)$ is dense.

To prove completeness, let (X_k) be Cauchy in H. By density choose $z_k \in H_0$ with $\|Jz_k - X_k\|_H < 1/k$. For $k, l \geq 1$,

$$\|z_k - z_l\|_{H_0} \leq 1/k + \|X_k - X_l\|_H + 1/l.$$

Thus (z_k) is Cauchy in H_0 and defines $Z = [(z_k)] \in H$. The density calculation gives $Jz_k \to Z$, and $\|X_k - Z\|_H \leq 1/k + \|Jz_k - Z\|_H \to 0$. Therefore H is complete.

Finally, for another completion $(\widetilde{H}, \widetilde{J})$, send $[(x_j)]$ to $\lim_j \widetilde{J}x_j$. Completeness of $\widetilde{H}$ gives the limit; equivalence of sequences proves that it is well-defined. It is a linear isometry and agrees with the prescribed map on H_0. Its image is complete, hence closed: a limit of image points has Cauchy preimages, which converge in H. The image also contains the dense set $\widetilde{J}(H_0)$, so it is all of $\widetilde{H}$. Density and continuity give uniqueness. □

11.3 Fourier series on the circle

Identify $\mathbb{T}$ with $[-\pi, \pi)$ with endpoints joined, and use the normalized measure $d\mu(t) = dt/(2\pi)$. For $f \in L^1(\mathbb{T})$ set

$$\widehat{f}(k) = \int_{\mathbb{T}} f(t)e^{-ikt}\, d\mu(t), \quad k \in \mathbb{Z}.$$

Convolution is $(f * g)(t) = \int_{\mathbb{T}} f(t-s)g(s)\, d\mu(s)$. Periodic changes of variable preserve this integral. The functions $e_k(t) = e^{ikt}$ form an orthonormal system in $L^2(\mathbb{T})$, since the integral of $e^{i(k-l)t}$ is zero when $k \neq l$ and one when $k = l$.

Lemma 11.7. *Continuous periodic functions are dense in $L^p(\mathbb{T})$ for $1 \leq p < \infty$. Translations are continuous in these norms.*

Proof. Represent f on $(-\pi, \pi)$. Truncation to $[-\pi + \delta, \pi - \delta]$ converges to f in L^p as $\delta \downarrow 0$, by absolute continuity of the integral of $|f|^p$. Extend the truncated function by zero to $\mathbb{R}$. It can be approximated in $L^p(\mathbb{R})$ by compactly supported smooth functions whose support stays inside $(-\pi, \pi)$: first approximate by a compactly supported smooth function and multiply by a fixed smooth cutoff equal to one on the truncated support. Such functions have smooth periodic extensions. This proves density. Alternatively, the local mollification construction gives the same approximants directly.

For continuous periodic g, uniform continuity on the compact circle gives $\|g(\cdot - h) - g\|_p \to 0$. If $\|f - g\|_p < \varepsilon$, translation invariance and the triangle inequality give

$$\|f(\cdot - h) - f\|_p \leq 2\varepsilon + \|g(\cdot - h) - g\|_p.$$

First let $h \to 0$ and then $\varepsilon \downarrow 0$. □

Theorem 11.8 (Fejér approximation)**.** *For $N \geq 1$, define*

$$F_N(t) = \frac{1}{N}\Big|\sum_{j=0}^{N-1} e^{ijt}\Big|^2.$$

Then $F_N \geq 0$, $\int_{\mathbb{T}} F_N \, d\mu = 1$, and

$$F_N(t) = \sum_{|k|<N} \big(1 - \frac{|k|}{N}\big) e^{ikt}.$$

For $1 \leq p < \infty$, $F_N * f \to f$ in $L^p(\mathbb{T})$. For continuous periodic f, the convergence is uniform.

Proof. Expanding the square gives $N^{-1} \sum_{j,l=0}^{N-1} e^{i(j-l)t}$. There are $N - |k|$ pairs with $j - l = k$ when $|k| < N$, which proves the displayed expansion. Integrating it leaves only the constant coefficient, equal to one. For $t \notin 2\pi\mathbb{Z}$, the finite geometric sum also gives

$$F_N(t) = \frac{1}{N} \frac{\sin^2(Nt/2)}{\sin^2(t/2)}.$$

For each fixed $0 < \delta < \pi$, this yields

$$\sup_{\delta \leq |t| \leq \pi} F_N(t) \leq \frac{1}{N \sin^2(\delta/2)} \longrightarrow 0.$$

Thus the mass outside $(-\delta, \delta)$ tends to zero.

Since $F_N\, d\mu$ is a probability measure, Hölder's inequality, including its direct $p = 1$ form, gives

$$\|F_N * f - f\|_p^p \le \int_{\mathbb{T}} F_N(s)\|f(\cdot - s) - f\|_p^p \, d\mu(s).$$

The application of Fubini here is justified by Tonelli applied to the nonnegative integrand. Given $\varepsilon > 0$, Lemma 11.7 supplies δ so that the translation norm is below ε for $|s| < \delta$. The remaining translation norms are at most $2\|f\|_p$. Consequently,

$$\|F_N * f - f\|_p^p \le \varepsilon^p + 2^p\|f\|_p^p \int_{|s|\ge\delta} F_N(s)\, d\mu(s).$$

Let $N \to \infty$ and then $\varepsilon \downarrow 0$. For continuous f, the same argument is used before integrating in t: uniform continuity controls the near part uniformly in t, and $2\|f\|_\infty$ controls the far part. This proves uniform convergence. □

Corollary 11.9 (Uniqueness and the L^2 Fourier series)**.** *If all Fourier coefficients of $f \in L^1(\mathbb{T})$ vanish, then $f = 0$ almost everywhere. The exponentials form an orthonormal basis of $L^2(\mathbb{T})$. In particular,*

$$\|f\|_2^2 = \sum_{k\in\mathbb{Z}} |\widehat{f}(k)|^2, \qquad \sum_{|k|\le N} \widehat{f}(k)e^{ikt} \longrightarrow f \quad \text{in } L^2(\mathbb{T}).$$

Every square summable sequence indexed by $\mathbb{Z}$ occurs as the Fourier coefficients of a unique L^2 function.

Proof. The finite expansion of F_N allows direct integration term by term and gives

$$(F_N * f)(t) = \sum_{|k|<N} \big(1 - \frac{|k|}{N}\big)\widehat{f}(k)e^{ikt}.$$

If all coefficients vanish, every such convolution is zero. Its L^1 limit is f, so $f = 0$ almost everywhere. Since the circle has measure one, $L^2(\mathbb{T}) \subset L^1(\mathbb{T})$ by Hölder. An L^2 function orthogonal to all exponentials therefore vanishes. Theorem 11.5 and Proposition 11.6 give all remaining assertions. Symmetric partial sums are precisely the orthogonal projections onto the spans of e_k, $|k| \le N$. □

The last result is convergence in norm. It does not assert pointwise convergence of the ordinary partial sums. Fejér means and ordinary partial sums are different summation procedures.

11.4 Poisson integrals and radial limits

Proposition 11.10 (The Poisson kernel). *For $0 \le r < 1$, put*

$$P_r(t) = \frac{1-r^2}{1-2r\cos t + r^2}.$$

It is nonnegative, has integral one on $\mathbb{T}$, and has the absolutely and uniformly convergent expansion

$$P_r(t) = \sum_{k\in\mathbb{Z}} r^{|k|} e^{ikt}.$$

If $f \in L^p(\mathbb{T})$, $1 \le p < \infty$, then $P_r * f \longrightarrow f$ in L^p as $r \uparrow 1$. If t is a Lebesgue point of $f \in L^1(\mathbb{T})$, then $(P_r * f)(t) \longrightarrow f(t)$ as $r \uparrow 1$.

Proof. Sum the geometric series $\sum_{k\ge1} r^k e^{ikt}$ and its conjugate, and add one. The denominator is $|1 - re^{it}|^2$ and the resulting numerator is $1 - r^2$. Absolute uniform convergence follows from $\sum_{k\in\mathbb{Z}} r^{|k|} < \infty$. Integration gives mass one. For each fixed $\eta > 0$, the denominator is bounded away from zero on $\eta \le |t| \le \pi$ as $r \uparrow 1$, while the numerator tends to zero. The same near and far argument used for Fejér kernels proves norm convergence.

For the pointwise assertion fix a Lebesgue point t_0 and extend f periodically. Write $\delta = 1 - r$ and restrict to $r \ge 1/2$, which does not affect the limit. Since $1 - \cos s \ge cs^2$ for $|s| \le \pi$, there is an absolute constant C such that

$$0 \le P_r(s) \le \frac{C\delta}{\delta^2 + s^2}.$$

Here the trigonometric lower bound follows, for example, from concavity of $\sin u$ on $[0, \pi/2]$ and $1 - \cos s = 2\sin^2(s/2)$.

Set $H(s) = |f(t_0 - s) - f(t_0)|$ on $[-\pi, \pi]$ and extend H by zero outside this interval. Then $H \in L^1(\mathbb{R})$ and

$$A(a) = \frac{1}{a}\int_{|s|<a} H(s)\,ds \longrightarrow 0 \quad (a \downarrow 0).$$

Moreover A is bounded for $a > 0$: it is bounded near zero by the displayed limit, and for $a \ge a_0 > 0$ it is at most $\|H\|_1/a_0$. Split the integral against P_r into $|s| < \delta$ and the annuli $2^j\delta \le |s| < 2^{j+1}\delta$. The kernel bound gives

$$\int_{-\pi}^{\pi} P_r(s)H(s)\,ds \le CA(\delta) + C\sum_{j=0}^{\infty} 2^{-j} A(2^{j+1}\delta).$$

The zero extension permits writing an infinite sum even after the annuli leave $[-\pi, \pi]$. Let $M = \sup_{a>0} A(a)$. Given $\varepsilon > 0$, first choose J so that

$CM \sum_{j \geq J} 2^{-j} < \varepsilon$. For the remaining finitely many terms, choose δ sufficiently small that each $A(\delta)$ and $A(2^{j+1}\delta)$ is as small as needed. The right side tends to zero. Finally,

$$|(P_r * f)(t_0) - f(t_0)| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} P_r(s) H(s)\, ds,$$

which proves convergence for the full limit $r \uparrow 1$, not merely along a subsequence. □

Theorem 11.11 (Fatou's radial limit theorem for bounded holomorphic functions)**.** *Let F be holomorphic in the unit disc and satisfy $|F(z)| \leq M$ there. There is a function $g \in L^\infty(\mathbb{T})$, with $\|g\|_\infty \leq M$, such that*

$$\lim_{r \uparrow 1} F(re^{it}) = g(t) \quad \textit{for almost every } t.$$

Proof. Write the power series $F(z) = \sum_{k=0}^{\infty} a_k z^k$. For each $r < 1$ it converges uniformly on the circle $|z| = r$, so the Fourier coefficients of $t \mapsto F(re^{it})$ are $a_k r^k$ for $k \geq 0$ and zero for $k < 0$. Parseval gives

$$\sum_{k=0}^{\infty} |a_k|^2 r^{2k} = \int_{\mathbb{T}} |F(re^{it})|^2\, d\mu(t) \leq M^2.$$

Letting $r \uparrow 1$ and applying monotone convergence to this nonnegative series gives $\sum_{k \geq 0} |a_k|^2 \leq M^2$. Corollary 11.9 therefore supplies $g \in L^2(\mathbb{T})$ whose coefficients are a_k for $k \geq 0$ and zero for $k < 0$.

For fixed r, the coefficients of $P_r * g$ are $r^{|k|}\widehat{g}(k)$, by Fubini; absolute integrability follows from $P_r \in L^1$ and $g \in L^1$. These agree with the coefficients of $F(re^{it})$. Uniqueness gives equality almost everywhere in t. In fact equality holds everywhere: $F(re^{it})$ is continuous, and $P_r * g$ is continuous because P_r is uniformly continuous and

$$|(P_r * g)(t+h) - (P_r * g)(t)| \leq \|g\|_1 \sup_s |P_r(s+h) - P_r(s)|.$$

Proposition 11.10 now gives the radial limit at every Lebesgue point of g, hence almost everywhere. The bound $|F(re^{it})| \leq M$ passes to this pointwise limit, showing $g \in L^\infty$ and $\|g\|_\infty \leq M$. □

11.5 The Fourier transform on $L^2(\mathbb{R}^n)$

Let $\mathcal{S}(\mathbb{R}^n)$ denote the Schwartz space: $f \in C^\infty(\mathbb{R}^n)$ belongs to $\mathcal{S}$ when $\sup_x |x^\alpha D^\beta f(x)| < \infty$ for every pair of multiindices. These functions and all their derivatives decay faster than any fixed negative power of $1+|x|$. In particular they belong to every L^p, $1 \le p \le \infty$.

Lemma 11.12. *The Fourier transform with phase $e^{-2\pi i x\cdot\xi}$ maps $\mathcal{S}$ into itself. Its inverse on $\mathcal{S}$ is $h \mapsto \check{h}$, where $\check{h}(x) = \int h(\xi)e^{2\pi i x\cdot\xi}\,d\xi$.*

Proof. Differentiation under the integral is justified by the integrable majorant $(2\pi)^{|\beta|}|x^\beta f(x)|$. Thus

$$D_\xi^\beta \widehat{f} = \widehat{(-2\pi i x)^\beta f}.$$

Integration by parts in each coordinate gives

$$\widehat{D^\alpha h}(\xi) = (2\pi i\xi)^\alpha \widehat{h}(\xi) \quad (h \in \mathcal{S}).$$

One may first integrate over bounded coordinate intervals; all boundary terms tend to zero by rapid decay, and the derivative integrals converge absolutely. Applying this identity to $h = (-2\pi i x)^\beta f$ bounds every $\xi^\alpha D_\xi^\beta \widehat{f}(\xi)$ by a finite linear combination of L^1 norms of derivatives of $x^\beta f$. Hence $\widehat{f} \in \mathcal{S}$.

In particular $f, \widehat{f} \in L^1$. The inversion theorem in Chapter 8 applies and gives $f = \check{\widehat{f}}$ at every point, because f is continuous. Applying the same argument to the transform with the opposite sign gives $\widehat{\check{h}} = h$ for $h \in \mathcal{S}$. □

Theorem 11.13 (Plancherel)**.** *The Fourier transform on $\mathcal{S}(\mathbb{R}^n)$ extends uniquely to a unitary operator on $L^2(\mathbb{R}^n)$. If $f \in L^1 \cap L^2$, this extension agrees almost everywhere with the integral definition of $\widehat{f}$.*

Proof. For $f, g \in \mathcal{S}$, use $\widehat{f}(\xi) = \int f(x)e^{-2\pi i x\cdot\xi}\,dx$ and Fubini to obtain

$$\begin{aligned}\int \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\,d\xi &= \int f(x)\Big(\int e^{-2\pi i x\cdot\xi}\,\overline{\widehat{g}(\xi)}\,d\xi\Big)dx \\ &= \int f(x)\overline{g(x)}\,dx.\end{aligned}$$

The double integral is absolutely convergent since its absolute integral is $\|f\|_1\|\widehat{g}\|_1$. Inversion for g identifies the inner integral. Taking $g = f$ proves the isometry on $\mathcal{S}$.

Since $C_c^\infty \subset \mathcal{S}$ is dense in L^2, for each $f \in L^2$ choose $f_j \in \mathcal{S}$ with $f_j \to f$ in L^2. The isometry makes $(\widehat{f_j})$ Cauchy in L^2; define the extended transform to be its limit. If a different approximating sequence is used, the norm of the difference of the transforms equals that of the original difference and tends to zero. Thus the definition is independent of the sequence and preserves the norm and inner product.

The range of an isometry from a complete space is closed: a Cauchy sequence of images comes from a Cauchy sequence of preimages. The range contains $\mathcal{S}$, since every $h \in \mathcal{S}$ is the transform of $\check{h} \in \mathcal{S}$. It is therefore both dense and closed in L^2, so the extension is surjective.

Finally, if $f \in L^1 \cap L^2$, first truncate its values and spatial support, and then mollify, to obtain $f_j \in C_c^\infty$ converging to f in both norms. Spatial and value truncation converge by dominated convergence separately for $|f|$ and $|f|^2$; after each truncation choose the mollification scale to make both errors smaller than $1/j$. The integral transforms then converge uniformly to the integral transform of f, because their uniform difference is bounded by $\|f_j - f\|_1$. They also converge in L^2 to the extended transform. An almost everywhere convergent subsequence of the latter convergence identifies the two limits almost everywhere. □

11.6 Weak convergence and a Dirichlet problem

A sequence x_j converges weakly to x in a Hilbert space if $\langle x_j, v\rangle \to \langle x, v\rangle$ for every fixed $v \in H$. Norm convergence implies weak convergence by Cauchy–Schwarz. The converse fails: an infinite orthonormal sequence converges weakly to zero, by Bessel's inequality, while every term has norm one.

Theorem 11.14. *Every bounded sequence in a Hilbert space has a weakly convergent subsequence. If $x_j \rightharpoonup x$, then $\|x\| \le \liminf_j \|x_j\|$.*

Proof. Let $\|x_j\| \le C$ and let V be the closed linear span of the sequence. Finite linear combinations with rational real and imaginary coefficients form a countable dense subset of V. Thus V has a finite or countable orthonormal basis. In the countably infinite case write it as (e_k). Each scalar sequence $\langle x_j, e_k\rangle$ is bounded. Repeated subsequence selection followed by the diagonal choice gives indices j_l for which every coordinate converges, say to a_k. For each N,

$$\sum_{k=1}^{N} |a_k|^2 = \lim_{l\to\infty} \sum_{k=1}^{N} |\langle x_{j_l}, e_k\rangle|^2 \le C^2.$$

Letting $N \to \infty$ gives $(a_k) \in \ell^2$. Hence $x = \sum_k a_k e_k \in V$ and $\|x\| \le C$. Convergence against finite linear combinations follows from convergence of coordinates. For $v \in V$, choose a finite combination v_N with $\|v - v_N\| \to 0$ and use

$$|\langle x_{j_l} - x, v - v_N\rangle| \le 2C\|v - v_N\|.$$

First choose N and then l, which proves convergence against v. For arbitrary $v \in H$, replace it by $P_V v$, since all the vectors x_{j_l} and x lie in V. The finite dimensional case uses the same proof with finitely many coordinates.

For the lower bound, if $x = 0$ there is nothing to prove. Otherwise test weak convergence against $v = x/\|x\|$. Then $\langle x_j, v\rangle \to \|x\|$ and $|\langle x_j, v\rangle| \leq \|x_j\|$. Taking the lower limit proves the assertion. □

Theorem 11.15 (Weak Dirichlet problem). *Let $\Omega \subset \mathbb{R}^n$ be bounded and open, and let $F \in L^2(\Omega)$. There is a unique $u \in H_0^1(\Omega) = W_0^{1,2}(\Omega)$ such that*

$$\int_\Omega \nabla u \cdot \overline{\nabla v}\, dx = \int_\Omega F\overline{v}\, dx \quad \text{for every } v \in H_0^1(\Omega).$$

Moreover, $\|u\|_{H^1} \leq C_\Omega \|F\|_2$. The function u uniquely minimizes

$$J(w) = \frac{1}{2}\int_\Omega |\nabla w|^2\, dx - \operatorname{Re}\int_\Omega F\overline{w}\, dx \quad \text{over } H_0^1(\Omega).$$

Proof. The Poincaré inequality for $W_0^{1,2}$ shows that $\|v\|_2 \leq C_\Omega \|\nabla v\|_2$. Consequently, $\langle v, w\rangle_E = \int \nabla v \cdot \overline{\nabla w}$ is an inner product on H_0^1, and its norm is equivalent to the usual H^1 norm. This space is complete in the energy norm: an energy Cauchy sequence is H^1 Cauchy by Poincaré, its limit lies in the closed space H_0^1, and its gradients converge in L^2.

The linear functional $L(v) = \int_\Omega v\overline{F}$ satisfies

$$|L(v)| \leq \|v\|_2 \|F\|_2 \leq C_\Omega \|\nabla v\|_2 \|F\|_2.$$

Riesz representation in the energy space gives u such that $L(v) = \langle v, u\rangle_E$. Taking complex conjugates gives the asserted weak equation. Its representation norm, or the equation with $v = u$, gives $\|\nabla u\|_2 \leq C_\Omega \|F\|_2$, and Poincaré controls the remaining part of the H^1 norm. The difference of two solutions has zero energy by testing against itself, so uniqueness follows.

For $w = u + v$, expand the square in J. The mixed term $\operatorname{Re}\int \nabla u \cdot \overline{\nabla v}$ cancels $\operatorname{Re}\int F\overline{v}$ by the weak equation. Hence $J(u + v) = J(u) + \frac{1}{2}\|\nabla v\|_2^2$. This is minimized only when $v = 0$, again by Poincaré. No assertion about pointwise boundary traces is needed: the zero boundary condition is the closure definition of H_0^1. □

Notes. The Hilbert space, Fourier series, and radial limit arguments follow the topics of the handwritten course notes. General background is provided by [1, 2]. The order above separates completeness of the Fourier system from any pointwise claim about its ordinary partial sums.

Exercises

Exercise 11.1. Prove that a finite dimensional subspace of a normed space is closed. Identify where this fact is used in the Gram–Schmidt construction.

Exercise 11.2. For a closed subspace V of a Hilbert space, prove $(V^\perp)^\perp = V$, $P_V^2 = P_V$, and $\langle P_V x, y\rangle = \langle x, P_V y\rangle$.

Exercise 11.3. Show that every continuous linear functional on ℓ^2 is given by pairing with a unique ℓ^2 sequence. Derive the formula directly from the coordinate vectors and compare it with Riesz representation.

Exercise 11.4. Let (e_j) be orthonormal and $(a_j) \in \ell^2$. Prove that the sum $\sum a_j e_j$ is unchanged by any permutation, with convergence in Hilbert norm.

Exercise 11.5. Construct a nonseparable Hilbert space as $\ell^2(I)$ for an uncountable set I. Prove that every element has at most countably many nonzero coordinates, although the space has no countable dense subset.

Exercise 11.6. Show that $F_N = N^{-1} \sum_{j=0}^{N-1} D_j$, where $D_j(t) = \sum_{|k|\le j} e^{ikt}$. Explain why averaging partial sums produces a nonnegative kernel.

Exercise 11.7. Prove that trigonometric polynomials are dense in $C(\mathbb{T})$ in the uniform norm and in $L^p(\mathbb{T})$ for $1 \le p < \infty$. Do not assume Fourier series convergence in advance.

Exercise 11.8. Show that the symmetric Fourier partial sum is the best L^2 approximation to f among trigonometric polynomials of degree at most N. Express the squared error in terms of the coefficients.

Exercise 11.9. Prove $P_r * P_s = P_{rs}$ on $\mathbb{T}$ for $0 \le r, s < 1$. Justify the coefficient calculation and the passage from equality almost everywhere to equality everywhere.

Exercise 11.10. If $g \in L^2(\mathbb{T})$ has zero negative Fourier coefficients, show that $\sum_{k\ge 0} \widehat{g}(k) z^k$ defines a holomorphic function in the disc and has radial limit g almost everywhere.

Exercise 11.11. Using Plancherel, prove the L^2 identities for translations and modulations. Then prove that $\widehat{\partial_j f} = 2\pi i \xi_j \widehat{f}$ for $f \in H^1(\mathbb{R}^n)$, as an L^2 identity.

Exercise 11.12. Suppose $x_j \rightharpoonup x$ in a Hilbert space and $\|x_j\| \to \|x\|$. Prove $x_j \to x$ in norm. Give a weakly convergent sequence for which the norms do not converge to the norm of its limit.

Exercise 11.13. In the weak Dirichlet problem, replace $F \in L^2$ by a bounded conjugate linear functional on the energy space. Prove existence, uniqueness, and the corresponding norm estimate.

Exercise 11.14. Let (u_j) be bounded in $H_0^1(\Omega)$ for bounded open Ω. Combine the weak subsequence theorem with the compactness result of Chapter 10 to find a subsequence converging weakly in H_0^1 and strongly in L^2, with the same limit.

Chapter 12

Further measure theory

The preceding chapters began with Lebesgue measure on Euclidean space. We now distinguish the arguments that use countable additivity alone from those that use Euclidean approximation by open and compact sets. This leads to the Radon–Nikodym theorem and the representation of bounded functionals on L^p. On the line, measures associated with increasing functions then give a proof of the converse part of the fundamental theorem for absolutely continuous functions.

12.1 Measure spaces and abstract integration

Definition 12.1. A σ-algebra $\mathcal{M}$ on a set X contains X and is closed under complements and countable unions. A positive measure on $(X, \mathcal{M})$ is a map $\mu : \mathcal{M} \to [0, \infty]$ such that $\mu(\varnothing) = 0$ and

$$\mu\Big(\bigcup_{j=1}^{\infty} E_j\Big) = \sum_{j=1}^{\infty} \mu(E_j)$$

for every pairwise disjoint sequence in $\mathcal{M}$. The measure is finite if $\mu(X) < \infty$, and σ-finite if $X = \bigcup_j X_j$ for measurable sets of finite measure. The space is complete if every subset of a measurable null set belongs to $\mathcal{M}$.

Counting measure is given by the number of points in a finite set and by ∞ on an infinite set. It is σ-finite precisely when X is countable. The Dirac measure at x_0 is $\delta_{x_0}(A) = \mathbf{1}_A(x_0)$. A nonnegative measurable density w on $\mathbb{R}^n$ gives a measure $\mu(A) = \int_A w\, dx$ on Lebesgue measurable sets. Countable additivity of this last example follows by monotone convergence of the partial sums of the disjoint indicators.

Proposition 12.2 (Continuity and completion). *If $E_j \uparrow E$, then $\mu(E_j) \uparrow \mu(E)$. If $E_j \downarrow E$ and one E_j has finite measure, then $\mu(E_j) \downarrow \mu(E)$. Every measure space has a completion on the same underlying set.*

Proof. For increasing sets, decompose E into the disjoint sets E_1 and $E_j \setminus E_{j-1}$, $j \geq 2$. Countable additivity identifies $\mu(E_j)$ with the partial sums. For decreasing sets, discard a finite initial segment so that $\mu(E_1) < \infty$, and apply the increasing assertion to $E_1 \setminus E_j$. Subtraction is permitted because $\mu(E_1)$ is finite.

For the completion, let $\overline{\mathcal{M}}$ consist of sets E for which there exist $A, B \in \mathcal{M}$ with $A \subset E \subset B$ and $\mu(B \setminus A) = 0$. Define $\overline{\mu}(E) = \mu(A)$. This does not depend on the choices. Indeed, if also $A' \subset E \subset B'$ is such a sandwich, then $A \setminus A' \subset B' \setminus A'$ and $A' \setminus A \subset B \setminus A$ have measure zero; additivity gives $\mu(A) = \mu(A')$, including the infinite case without subtracting infinities.

Complements have the sandwiches $B^c \subset E^c \subset A^c$. Countable unions have the sandwiches $\bigcup A_j \subset \bigcup E_j \subset \bigcup B_j$, and the excess lies in $\bigcup(B_j \setminus A_j)$, a measurable null set. Thus $\overline{\mathcal{M}}$ is a σ-algebra. For disjoint E_j, their measurable lower sets A_j are disjoint, so countable additivity of $\overline{\mu}$ follows from that of μ and the union sandwich. Finally, if $\overline{\mu}(E) = 0$, the upper set B in a sandwich is null, so every subset of E has the sandwich $\varnothing \subset F \subset B$. The extension is complete. □

A real function is $\mathcal{M}$-measurable when all its upper level sets belong to $\mathcal{M}$. Simple functions and nonnegative integrals are defined by

$$\int \sum_{j=1}^{N} a_j \mathbf{1}_{E_j}\, d\mu = \sum_{j=1}^{N} a_j \mu(E_j), \qquad \int f\, d\mu = \sup_{0 \leq \varphi \leq f,\ \varphi \text{ simple}} \int \varphi\, d\mu,$$

using disjoint level sets in the first formula and the convention $0 \cdot \infty = 0$. In this abstract definition, simple functions may have nonzero level sets of infinite measure. Their nonnegative integrals are allowed to be infinite.

Theorem 12.3 (The convergence theorems on a measure space). *Monotone convergence, Fatou's lemma, and dominated convergence hold on every measure space. The spaces $L^p(X, \mu)$ are complete for $1 \leq p \leq \infty$.*

Proof. We give the part of the integral construction that requires attention on a possibly non-σ-finite space. Let $0 \leq f_j \uparrow f$, let $0 < c < 1$, and let $\varphi = \sum_{i=1}^{N} a_i \mathbf{1}_{A_i} \leq f$ be nonnegative simple with $a_i > 0$ and disjoint A_i. Set $A_{i,j} = A_i \cap \{f_j \geq c a_i\}$. For each i, these sets increase to A_i. Monotonicity of the simple integral gives

$$\int f_j\,d\mu \geq c\sum_{i=1}^{N} a_i\mu(A_{i,j}).$$

Continuity from below of the measure gives, after $j \to \infty$, a lower bound $c\int \varphi\,d\mu$, even if one of the measures is infinite. Let $c \uparrow 1$ and then take the supremum over φ. The opposite bound follows from $f_j \leq f$. This proves monotone convergence.

Dyadic simple approximants, using both truncation of values and refinement of the value mesh, exist on any measurable space exactly as in Chapter 2. Applying monotone convergence to these approximants gives additivity of nonnegative integrals. Fatou follows by applying monotone convergence to $\inf_{j\geq N} f_j$ and then bounding its integral by $\inf_{j\geq N}\int f_j$. These proofs do not use compactness or regularity.

For dominated convergence suppose $f_j \to f$ almost everywhere and $|f_j| \leq g \in L^1$. After removing one common measurable null set, $|f| \leq g$. Fatou applied to the nonnegative functions $2g - |f_j - f|$ gives

$$2\int g\,d\mu \leq \liminf_j\Big(2\int g\,d\mu - \int |f_j - f|\,d\mu\Big).$$

The integral of g is finite, so rearrangement yields $\limsup_j \int |f_j - f| \leq 0$. This proves L^1 convergence and hence convergence of the integrals. If the space is not complete, an almost everywhere limit is understood as a measurable representative; arbitrary redefinitions on nonmeasurable subsets of null sets are not made.

Hölder and Minkowski use only the integral properties just proved, so their proofs in Chapter 6 apply unchanged. For completeness when $p < \infty$, select from a Cauchy sequence a subsequence with successive L^p distances at most 2^{-j}. Minkowski bounds the norms of the finite sums of the absolute successive differences by $\sum 2^{-j}$. Fatou then shows that the infinite sum lies in L^p and is finite almost everywhere. Thus the subsequence has a measurable pointwise limit outside a measurable null set. Its norm distance to that limit is bounded by the corresponding series tail, again by Fatou. The Cauchy property brings the full sequence to the same limit. For $p = \infty$, remove the countable union of the exceptional sets in the essential bounds for the successive differences. The subsequence converges uniformly on the complement, with the same summable tail bound. This gives an L^∞ limit and then convergence of the full sequence. □

12.2 Outer measures and extension

Definition 12.4. An outer measure on X is a function $\mu_* : \mathcal{P}(X) \to [0, \infty]$ that vanishes on the empty set, is monotone, and is countably subadditive. A set $E \subset X$ is Carathéodory measurable for μ_* if

$$\mu_*(A) = \mu_*(A \cap E) + \mu_*(A \setminus E) \quad \text{for every } A \subset X. \tag{12.1}$$

Theorem 12.5 (Carathéodory's construction)**.** *The sets satisfying* (12.1) *form a σ-algebra. Restricting μ_* to this σ-algebra gives a complete measure.*

Proof. Subadditivity always gives the inequality from left to right in the sense $\mu_*(A) \le \mu_*(A \cap E) + \mu_*(A \setminus E)$. Thus only the reverse inequality needs proof when verifying membership. The empty set and X belong, and the condition is unchanged on replacing E by E^c.

If E, F satisfy the condition, split first by E and then split $A \setminus E$ by F. This gives

$$\mu_*(A) = \mu_*(A \cap E) + \mu_*((A \setminus E) \cap F) + \mu_*(A \setminus (E \cup F)).$$

Subadditivity on the first two pieces gives the needed lower bound for splitting by $E \cup F$. Thus finite unions belong. Complements then give finite intersections and differences.

Let E_j be disjoint members and set $E = \bigcup_j E_j$. Repeated finite splitting gives, for every N,

$$\begin{aligned} \mu_*(A) &= \sum_{j=1}^{N} \mu_*(A \cap E_j) + \mu_*\big(A \setminus \bigcup_{j=1}^{N} E_j\big) \\ &\ge \sum_{j=1}^{N} \mu_*(A \cap E_j) + \mu_*(A \setminus E). \end{aligned}$$

Let $N \to \infty$. Since all terms are nonnegative and $A \cap E = \bigcup_j (A \cap E_j)$, subadditivity gives

$$\mu_*(A) \ge \sum_j \mu_*(A \cap E_j) + \mu_*(A \setminus E) \ge \mu_*(A \cap E) + \mu_*(A \setminus E).$$

This proves that E belongs. An arbitrary countable union is first disjointified by replacing E_j with $E_j \setminus \bigcup_{i<j} E_i$, already known to belong. Hence the class is a σ-algebra.

Take $A = E$ in the finite splitting inequality and let $N \to \infty$ to obtain countable superadditivity on disjoint measurable sets. Outer subadditivity gives equality. Finally, if $N \subset Z$ and $\mu_*(Z) = 0$, then $\mu_*(A \cap N) = 0$ and monotonicity gives

$$\mu_*(A) \geq \mu_*(A \setminus N) = \mu_*(A \cap N) + \mu_*(A \setminus N).$$

Thus every subset of an outer null set is measurable and null. This proves completeness. No step subtracts two infinite quantities. □

A ring $\mathcal{A}$ of sets is closed under finite unions and differences. An algebra is a ring containing X. A premeasure μ_0 on a ring vanishes on the empty set and is countably additive whenever a disjoint sequence in the ring has its union in the ring. Finite additivity is part of this condition.

Theorem 12.6 (Extension from a premeasure)**.** *Suppose $\mathcal{A}$ is a ring covering X by countably many of its sets and μ_0 is a premeasure on $\mathcal{A}$. Define*

$$\mu_*(E) = \inf\Big\{\sum_{j=1}^{\infty} \mu_0(A_j) : E \subset \bigcup_j A_j,\ A_j \in \mathcal{A}\Big\}.$$

Then μ_* is an outer measure, every member of $\mathcal{A}$ is Carathéodory measurable, and $\mu_*(A) = \mu_0(A)$ for $A \in \mathcal{A}$. Thus μ_0 extends to a measure on $\sigma(\mathcal{A})$. If X has a countable cover by members of $\mathcal{A}$ of finite premeasure, the extension on $\sigma(\mathcal{A})$ is unique.

Proof. The empty covering gives zero for the empty set, and every covering of a larger set covers a smaller one. For countable subadditivity, the assertion is automatic when $\sum_j \mu_*(E_j) = \infty$. Otherwise, for each j choose a covering of E_j with cost at most $\mu_*(E_j) + \varepsilon 2^{-j}$. Combining these countably many countable coverings gives a covering of the union with total cost at most $\sum_j \mu_*(E_j) + \varepsilon$. Let $\varepsilon \downarrow 0$.

Fix $A \in \mathcal{A}$. The single-set cover gives $\mu_*(A) \leq \mu_0(A)$. For the reverse inequality, take any cover $A \subset \bigcup_j A_j$ by ring sets and define

$$B_j = A \cap \big(A_j \setminus \bigcup_{i<j} A_i\big).$$

These sets belong to the ring, are disjoint, and have union A. The premeasure property and finite monotonicity give

$$\mu_0(A) = \sum_j \mu_0(B_j) \leq \sum_j \mu_0(A_j).$$

Taking the infimum proves equality, with no finiteness requirement on $\mu_0(A)$.

For an arbitrary $E \subset X$, split each set in a ring cover into $A_j \cap A$ and $A_j \setminus A$. Finite additivity gives

$$\mu_*(E \cap A) + \mu_*(E \setminus A) \leq \sum_j \bigl(\mu_0(A_j \cap A) + \mu_0(A_j \setminus A)\bigr) = \sum_j \mu_0(A_j).$$

Take the infimum over covers. Theorem 12.5 now gives the extension, since its measurable class contains the ring and therefore $\sigma(\mathcal{A})$.

For uniqueness we use the elementary π–λ principle, with its proof given below. First suppose two finite measures on X agree on an algebra containing X. The class on which they agree contains X, is closed under complements, and is closed under countable disjoint unions. The complement assertion uses finiteness to subtract from the common total mass. Thus this class is a λ-system containing the algebra, a π-system, and consequently contains its generated σ-algebra.

In the σ-finite ring case, disjointify a countable finite-premeasure cover to obtain $X = \bigcup_j C_j$ with disjoint $C_j \in \mathcal{A}$ and $\mu_0(C_j) < \infty$. On each C_j the sets $A \cap C_j$, $A \in \mathcal{A}$, form an algebra with universe C_j. The two candidate extensions restrict to finite measures agreeing on this algebra. The finite uniqueness assertion gives equality on $\sigma(\mathcal{A})$ restricted to C_j. Summing over the disjoint C_j proves global uniqueness. □

Lemma 12.7 (π–λ principle)**.** *A π-system is a collection closed under finite intersections. A λ-system contains X, is closed under complements, and is closed under countable disjoint unions. Every λ-system containing a π-system $\mathcal{P}$ contains $\sigma(\mathcal{P})$.*

Proof. Let $\mathcal{D}$ be the smallest λ-system containing $\mathcal{P}$. In a λ-system, if $A \subset B$ are members, then $B \setminus A$ is a member: A and B^c are disjoint, so take the complement of their union.

Fix $A \in \mathcal{P}$. The class of $B \in \mathcal{D}$ for which $A \cap B \in \mathcal{D}$ is a λ-system. It contains X because $A \in \mathcal{D}$; complement closure follows from $A \cap B^c = A \setminus (A \cap B)$; disjoint union closure follows by distributing intersection over the union. This class contains $\mathcal{P}$ by its intersection property, and thus equals $\mathcal{D}$. We have proved $A \cap B \in \mathcal{D}$ for $A \in \mathcal{P}$, $B \in \mathcal{D}$.

Now fix $B \in \mathcal{D}$ and repeat the argument with the class of $A \in \mathcal{D}$ such that $A \cap B \in \mathcal{D}$. It contains $\mathcal{P}$ by the preceding paragraph and therefore all of $\mathcal{D}$. Hence $\mathcal{D}$ is closed under finite intersections. Complements then give finite unions. Every countable union can be disjointified using finite differences, so it too belongs to $\mathcal{D}$. Thus $\mathcal{D}$ is a σ-algebra containing $\mathcal{P}$, as required. □

12.3 Absolute continuity and the Radon–Nikodym theorem

For positive measures μ, ν on the same σ-algebra, write $\nu \ll \mu$ if $\mu(A) = 0$ implies $\nu(A) = 0$. Write $\nu \perp \mu$ if there is a measurable set S with $\mu(S) = 0$ and $\nu(X \setminus S) = 0$. The latter statement means that ν is concentrated on a μ-null set; it is not a statement about the topological support.

Theorem 12.8 (Lebesgue decomposition and Radon–Nikodym)**.** *Let μ and ν be σ-finite positive measures on $(X, \mathcal{M})$. There is a unique decomposition*

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

Moreover, $\nu_a(A) = \int_A h\, d\mu$ for a nonnegative measurable h, unique up to μ-null sets and finite μ-almost everywhere. If $\nu \ll \mu$, then $\nu = h\mu$. If ν is finite, then $h \in L^1(\mu)$.

Proof. Assume first that both measures are finite and put $\lambda = \mu + \nu$. Work in the real Hilbert space $L^2(X, \lambda)$. The map $L(\varphi) = \int \varphi\, d\nu$ is well-defined on λ-almost everywhere classes, since $\nu \le \lambda$. It is bounded because

$$|L(\varphi)| \le \int |\varphi|\, d\lambda \le \lambda(X)^{1/2} \|\varphi\|_{L^2(\lambda)}.$$

Theorem 11.4 gives a real $g \in L^2(\lambda)$ such that

$$\int \varphi\, d\nu = \int \varphi g\, d\lambda \quad \text{for every } \varphi \in L^2(\lambda).$$

In particular, all indicators are permitted, since λ is finite. Thus $\nu(A) = \int_A g\, d\lambda$.

For every measurable A, $0 \le \int_A g\, d\lambda \le \lambda(A)$. If $\lambda(\{g \le -1/k\}) > 0$ for some k, integration over this set contradicts the first inequality. If $\lambda(\{g \ge 1 + 1/k\}) > 0$, it contradicts the second. Taking the union over k and redefining on the resulting measurable null set, we may assume $0 \le g \le 1$ everywhere. Subtracting finite measures on indicators gives $\mu(A) = \int_A (1 - g)\, d\lambda$.

An identity of measures on indicators extends to integration of every nonnegative measurable function: prove it for simple functions by addition and then pass to increasing simple approximants by monotone convergence. Consequently,

$$\int q\, d\mu = \int q(1 - g)\, d\lambda \quad (q \ge 0 \text{ measurable}). \tag{12.2}$$

Let $S = \{g = 1\}$ and set $h = g/(1-g)$ on S^c, $h = 0$ on S. Then $\mu(S) = 0$ by (12.2). On S^c the denominator is strictly positive, so h is finite and measurable. Another use of (12.2) gives

$$\int_A h\, d\mu = \int_{A\setminus S} g\, d\lambda = \nu(A \setminus S).$$

Thus $\nu_a = \nu|_{S^c} = h\mu$ and $\nu_s = \nu|_S$ have the required properties. Their masses are at most $\nu(X)$, so $h \in L^1(\mu)$.

For uniqueness in the finite case, suppose another decomposition has singular support S'. On the complement of $S \cup S'$, each absolutely continuous part equals ν. On the μ-null set $S \cup S'$, both absolutely continuous parts vanish. They therefore agree on every measurable set, and subtraction from the finite measure ν gives equality of the singular parts. To identify densities, suppose h and $\widetilde{h}$ give the same integrals. On the measurable set $\{h \geq \widetilde{h} + 1/k\}$ their integrals can be equal only if that set is null, since both densities are integrable. The reverse inequality is treated by interchanging the two densities. Countable union gives equality almost everywhere.

For the σ-finite case, take finite-measure covers for μ and for ν. Their pairwise intersections give a countable cover on which both measures are finite; disjointify it to obtain measurable C_j with $X = \bigcup_j C_j$ and $\mu(C_j), \nu(C_j) < \infty$. Apply the finite construction on each C_j. Define $h = h_j$ there and let ν_s be the sum of the singular pieces. The union of their measurable μ-null concentration sets is again μ-null. Countable additivity on the disjoint pieces gives $\nu = h\mu + \nu_s$. The densities can be chosen finite on all pieces outside a single μ-null set. Any candidate decomposition restricts to a decomposition of the finite measure $\nu|_{C_j}$; finite uniqueness on each piece proves global uniqueness. If $\nu \ll \mu$, the singular part is bounded above by ν and concentrated on a μ-null set, hence is zero. □

Proposition 12.9 (Uniform absolute continuity for a finite measure)**.** *If ν is finite and $\nu \ll \mu$, then for every $\varepsilon > 0$ there is $\delta > 0$ such that $\mu(A) < \delta$ implies $\nu(A) < \varepsilon$. This assertion does not require σ-finiteness of μ.*

Proof. If it failed, there would be $\varepsilon_0 > 0$ and measurable A_j with $\mu(A_j) < 2^{-j}$ and $\nu(A_j) \geq \varepsilon_0$. Let $B_N = \bigcup_{j \geq N} A_j$. Then B_N decrease, $\mu(B_N) \leq 2^{1-N}$, and $\nu(B_N) \geq \varepsilon_0$. Their intersection B has $\mu(B) = 0$ by monotonicity and the displayed bound. Since ν is finite, continuity from above gives $\nu(B) = \lim_N \nu(B_N) \geq \varepsilon_0$, contradicting $\nu \ll \mu$. □

A finite signed measure is a countably additive real valued set function whose total variation is finite. Its variation is defined by

$$|\nu|(A) = \sup\Big\{\sum_{k=1}^{N} |\nu(A_k)| : A = \bigcup_{k=1}^{N} A_k \text{ a measurable disjoint partition}\Big\}.$$

The finiteness requirement is $|\nu|(X) < \infty$.

Proposition 12.10 (Variation and Jordan decomposition)**.** *The variation of a finite signed measure is a finite positive measure. The positive measures $\nu^+ = (|\nu| + \nu)/2$ and $\nu^- = (|\nu| - \nu)/2$ are mutually singular and satisfy*

$$\nu = \nu^+ - \nu^-, \quad |\nu| = \nu^+ + \nu^-.$$

If $\nu \ll \mu$ and μ is σ-finite, then $\nu = h\mu$ for an integrable real h, and $|\nu| = |h|\mu$.

Proof. Monotonicity and nonnegativity of $|\nu|$ follow from the definition. Let E_j be disjoint with union E. For each of the first N sets, choose a finite partition whose variation sum is within ε/N of $|\nu|(E_j)$. Combine the partitions, adding the remaining part of E as one more set. It follows that $|\nu|(E) \geq \sum_{j=1}^{N} |\nu|(E_j) - \varepsilon$. Let $\varepsilon \downarrow 0$ and then $N \to \infty$.

For the reverse inequality, take a finite partition $(A_k)_{k=1}^{K}$ of E. For fixed k, the series $\sum_j \nu(A_k \cap E_j)$ is absolutely convergent: every finite partial sum of its absolute values is bounded by $|\nu|(X)$ by completing the finite family to a partition of X. Countable additivity and the triangle inequality give

$$\sum_{k=1}^{K} |\nu(A_k)| \leq \sum_{k=1}^{K} \sum_{j} |\nu(A_k \cap E_j)| \leq \sum_{j} |\nu|(E_j).$$

Taking the supremum over the partition proves countable additivity of $|\nu|$. Also $|\nu|(A) \geq |\nu(A)|$, so the formulas for $\nu^{\pm}$ give positive measures and the two algebraic identities.

Put $\lambda = |\nu|$. Since $\nu^{\pm} \ll \lambda$ and all these measures are finite, Radon–Nikodym applied separately to the two positive measures gives $\nu = h\lambda$ for an integrable real h. For any such density, triangle inequalities on partitions show $|\nu|(A) \leq \int_A |h| \, d\lambda$. The partition of A into $A \cap \{h \geq 0\}$ and $A \cap \{h < 0\}$ gives the reverse inequality exactly. Thus $|\nu| = |h|\lambda$. But $|\nu| = \lambda$, so testing the sets where $|h|$ differs from one gives $|h| = 1$ almost everywhere for λ. The formulas for $\nu^{\pm}$ then show that they are concentrated on $\{h = 1\}$ and $\{h = -1\}$, respectively. These disjoint sets prove mutual singularity.

If $\nu \ll \mu$, then $|\nu| \ll \mu$: on a μ-null set every piece in a finite measurable partition is μ-null, so every signed value is zero. Therefore $\nu^{\pm} \ll \mu$. Apply Radon–Nikodym to them and subtract their integrable densities. The partition by the sign of this density proves the variation formula as before. $\square$

A complex measure of finite variation has real and imaginary parts that are finite signed measures. Applying the last proposition to these parts gives an integrable complex density whenever the complex measure is absolutely continuous with respect to a σ-finite positive measure. Only this consequence is needed below.

12.4 The dual of L^p

Theorem 12.11. *Let $(X, \mathcal{M}, \mu)$ be σ-finite and $1 \le p < \infty$. For every bounded linear functional L on the complex space $L^p(X, \mu)$, there is a unique $g \in L^{p'}(X, \mu)$ such that*

$$L(f) = \int_X f\overline{g}\, d\mu \quad (f \in L^p), \quad \|L\| = \|g\|_{p'}.$$

Here $p' = \infty$ when $p = 1$. The assertion does not identify the dual of L^∞.

Proof. Choose a disjoint measurable partition $X = \bigcup_j C_j$ with $\mu(C_j) < \infty$. On measurable subsets of C_j, define $\nu_j(A) = L(\mathbf{1}_A)$. If $A_k \subset C_j$ are disjoint and $A = \bigcup_k A_k$, then

$$\Big\|\mathbf{1}_A - \sum_{k=1}^{N} \mathbf{1}_{A_k}\Big\|_p^p = \mu\Big(A \setminus \bigcup_{k=1}^{N} A_k\Big) \longrightarrow 0.$$

The convergence follows from finiteness of $\mu(C_j)$ and continuity of measure. Continuity of L proves countable additivity of ν_j.

For a finite partition (A_k) of C_j, choose complex c_k with $|c_k| \le 1$ and $c_k\nu_j(A_k) = |\nu_j(A_k)|$. For a zero signed value choose $c_k = 0$. The function $v = \sum_k c_k \mathbf{1}_{A_k}$ has $\|v\|_p \le \mu(C_j)^{1/p}$, and

$$\sum_k |\nu_j(A_k)| = L(v) \le |L(v)| \le \|L\|\mu(C_j)^{1/p}.$$

The equality makes $L(v)$ a nonnegative real number. Thus ν_j has finite variation. It vanishes on μ-null sets, since their indicators are zero in L^p. The signed and complex forms of Radon–Nikodym give $h_j \in L^1(C_j, \mu)$ with $\nu_j = h_j\mu$. Set $g = \overline{h_j}$ on C_j.

The representation now holds for simple functions supported in a finite union of the C_j. It also holds for every bounded measurable function on such a union: approximate it uniformly by simple functions. Uniform convergence gives L^p convergence because the union has finite measure, and convergence of the density integrals because g is integrable there.

Let $C = \|L\|$. Suppose first that $1 < p < \infty$, and put

$$A_N = \Big(\bigcup_{j=1}^{N} C_j\Big) \cap \{|g| \le N\}, \quad f_N = \mathbf{1}_{A_N} g |g|^{p'-2},$$

where the last product is defined to be zero at $g = 0$. The magnitude of f_N is $|g|^{p'-1}$, so it is bounded and supported in a finite-measure union of pieces. Therefore the representation already proved applies. If $I_N = \int_{A_N} |g|^{p'} \, d\mu$, then

$$I_N = L(f_N) \le C\|f_N\|_p = C I_N^{1/p},$$

since $(p'-1)p = p'$. The quantity I_N is finite. If it is nonzero, division gives $I_N^{1/p'} \le C$; the same bound holds when it is zero. The sets A_N increase to X outside a measurable null set because each h_j is finite almost everywhere. Monotone convergence gives $\|g\|_{p'} \le C$.

For $p = 1$, suppose $\mu(\{|g| > C + \eta\}) > 0$ for some $\eta > 0$. Some finite union of the C_j meets this set in a set A of positive finite measure. The bounded test $f = \mathbf{1}_A g/|g|$ then gives

$$(C + \eta)\mu(A) < \int_A |g| \, d\mu = L(f) \le C\mu(A),$$

a contradiction. Hence $\|g\|_\infty \le C$.

For arbitrary $f \in L^p$, truncate its values and restrict it to the first N finite pieces; the resulting bounded functions converge in L^p by dominated convergence of the pth powers of the errors. Boundedness of L and Hölder with the now established $g \in L^{p'}$ allow passage to the limit in the representation. Hölder also gives $\|L\| \le \|g\|_{p'}$, so equality follows from the reverse bound above. For uniqueness, put $h = g - \widetilde{g}$ and let $D_N = (\bigcup_{j=1}^N C_j) \cap \{|h| \le N\}$. The function $v_N = \mathbf{1}_{D_N} h/|h|$, defined to be zero where $h = 0$, is bounded and has finite measure support, so $v_N \in L^p$. Equality of the two representations gives $0 = \int v_N \overline{h} \, d\mu = \int_{D_N} |h| \, d\mu$. The sets D_N increase to X apart from a measurable null set, because both densities are finite almost everywhere. Monotone convergence gives $\int |h| \, d\mu = 0$, and hence $h = 0$ almost everywhere. $\square$

12.5 Stieltjes measures and differentiation on the line

When comparing a Borel measure on $\mathbb{R}$ with Lebesgue measure, both are initially restricted to the Borel σ-algebra. Radon–Nikodym densities and singular supports can therefore be chosen Borel. The resulting density is also

Lebesgue measurable, so the Euclidean differentiation theorem applies to it. This distinguishes the Borel measure from its completion without changing the integrals under consideration.

Theorem 12.12 (Lebesgue–Stieltjes measure). *Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and right continuous. There is a unique locally finite Borel measure μ_F such that*

$$\mu_F((a,b]) = F(b) - F(a) \quad (a < b).$$

It satisfies $\mu_F(\{x\}) = F(x) - F(x-)$, where $F(x-) = \lim_{t\uparrow x} F(t)$.

Proof. Let $\mathcal{A}$ be the ring of finite disjoint unions of bounded intervals $(a,b]$. Define the value on each interval by its increment and add over the disjoint components. This is well-defined: two finite representations have a common refinement obtained by arranging all their endpoints, and the sums of increments telescope on each original interval. The same refinement proves finite additivity and monotonicity.

To prove the premeasure property, we first prove the covering inequality for one interval. Suppose $(a,b] \subset \bigcup_j (a_j, b_j]$. Fix $0 < \delta < b - a$ and $\varepsilon > 0$. By right continuity at b_j, choose $\eta_j > 0$ with

$$F(b_j + \eta_j) - F(b_j) < \varepsilon 2^{-j}.$$

The open intervals $(a_j, b_j + \eta_j)$ cover the compact set $[a+\delta, b]$, so finitely many suffice. Their corresponding half open intervals also cover $(a+\delta, b]$. Finite subadditivity on the ring therefore gives

$$\begin{aligned} F(b) - F(a+\delta) &\le \sum_{j \text{ in the finite subcover}} \big(F(b_j + \eta_j) - F(a_j)\big) \\ &\le \sum_{j=1}^{\infty} \big(F(b_j) - F(a_j)\big) + \varepsilon. \end{aligned}$$

Let $\delta \downarrow 0$ and use right continuity at a, and then let $\varepsilon \downarrow 0$. This proves the desired countable covering inequality for one interval.

If a ring set A is covered by ring sets A_j, intersect each A_j with each of the finitely many disjoint interval components of A. Apply the interval covering inequality to a component and then sum over the components. For a fixed j, the intersections are disjoint ring sets contained in A_j, so their total premeasure is at most that of A_j. This proves the covering inequality for arbitrary ring sets. For a disjoint sequence with union in the ring, this inequality gives countable subadditivity; finite additivity and monotonicity give the reverse inequality by taking finite partial unions. Hence the set function is a premeasure.

The ring generates the Borel sets: open intervals are countable unions of bounded half open intervals, and every open subset of $\mathbb{R}$ is a countable union of open intervals. The ring covers $\mathbb{R}$ by bounded intervals with finite increments. Theorem 12.6 gives existence and uniqueness. Local finiteness follows by bounding the measure of a compact interval by a slightly larger half open interval. Finally, $(x-1/j, x]$ decrease to $\{x\}$ and their measures are finite, so continuity from above gives

$$\mu_F(\{x\}) = \lim_{j\to\infty}\big(F(x) - F(x-1/j)\big) = F(x) - F(x-).$$

□

The use of $(a, b]$ agrees with right continuity. The increment of F at the right endpoint counts an atom there. For a continuous F, all singletons have zero μ_F measure, and the choice of open or closed endpoints does not affect interval measures.

Lemma 12.13 (Regularity of finite Borel measures on the line)**.** *If ν is a finite Borel measure on $\mathbb{R}$, then for every Borel A and $\varepsilon > 0$ there exist a compact K and an open O with*

$$K \subset A \subset O, \quad \nu(O \setminus K) < \varepsilon.$$

In particular, if ν is concentrated on a Borel set S, compact subsets of S can contain all but an arbitrarily small amount of its mass.

Proof. First allow the inner set to be closed instead of compact. A closed set A has open enlargements $O_j = \{x : \operatorname{dist}(x, A) < 1/j\}$ decreasing to A. Since ν is finite, $\nu(O_j \setminus A) \to 0$. The empty set is treated separately with its empty enlargement.

Let $\mathcal{C}$ be the Borel sets admitting, for every positive error, a closed inner set and an open outer set with excess measure below that error. Complements preserve this property by interchanging the two complements. Let $A_i \in \mathcal{C}$ and $A = \bigcup_i A_i$. Choose closed $K_i \subset A_i \subset O_i$ with O_i open and $\nu(O_i \setminus K_i) < \varepsilon 2^{-i-3}$. Then $O = \bigcup_i O_i$ is open and

$$\nu(O \setminus A) \le \sum_i \nu(O_i \setminus A_i) < \varepsilon/4.$$

By continuity from below and finiteness of $\nu(A)$, choose N so that $\nu(A \setminus \bigcup_{i=1}^N A_i) < \varepsilon/4$. The closed set $K = \bigcup_{i=1}^N K_i$ is contained in A, and

$$\nu(A \setminus K) \le \nu\big(A \setminus \bigcup_{i=1}^N A_i\big) + \sum_{i=1}^N \nu(A_i \setminus K_i) < \varepsilon/2.$$

Thus $A \in \mathcal{C}$. The class is a σ-algebra containing all closed sets, hence all Borel sets. Finally, replace the closed K by $K \cap [-R, R]$. Since ν is finite, the lost mass tends to zero as $R \to \infty$. Starting with a smaller error gives the required compact set. □

For a locally finite Borel measure, applying this lemma to its restrictions to bounded intervals gives compact inner approximation on each bounded piece. We will apply it below only to finite measures or such finite restrictions.

Lemma 12.14 (A maximal estimate for measures). *For a finite positive Borel measure ν on $\mathbb{R}$, define*

$$M\nu(x) = \sup_{I \ni x} \frac{\nu(I)}{m(I)},$$

where the supremum is over bounded open intervals. Then $M\nu$ is measurable and

$$m(\{M\nu > \lambda\}) \le \frac{3\nu(\mathbb{R})}{\lambda} \quad (\lambda > 0).$$

Proof. If one interval witnesses $M\nu(x) > \lambda$, it witnesses the same inequality at every point of that interval. Thus each strict upper level set is open. Given a compact subset K of this level set, choose for each of its points a witnessing interval and take a finite subcover. The finite covering lemma of Chapter 7 supplies disjoint intervals $I_1, \ldots, I_N$ whose triples cover K. Consequently,

$$m(K) \le 3\sum_{j=1}^{N} m(I_j) < \frac{3}{\lambda}\sum_{j=1}^{N} \nu(I_j) \le \frac{3\nu(\mathbb{R})}{\lambda}.$$

The last inequality uses disjointness of the selected open intervals. Inner regularity of Lebesgue measure on the open level set gives the assertion by taking the supremum over compact K. □

Theorem 12.15 (Differentiation of a measure). *Let ν be a finite positive Borel measure on $\mathbb{R}$, with decomposition $\nu = h\,dx + \sigma$ relative to Lebesgue measure. Then*

$$\lim_{r \downarrow 0} \frac{\nu((x-r, x+r))}{2r} = h(x) \quad \textit{for almost every } x.$$

The analogous density limit holds on the one sided intervals $(x, x+t]$ and $(x-t, x]$, divided by t, as $t \downarrow 0$.

Proof. For the absolutely continuous part, the centered assertion is the Lebesgue differentiation theorem applied to $h \in L^1$. It remains to prove zero density for σ. Choose a Borel null set S on which σ is concentrated. For each $\varepsilon > 0$, Lemma 12.13 gives compact $K \subset S$ with $\sigma(\mathbb{R} \setminus K) < \varepsilon$. If $x \notin K$, its distance from K is positive. Every sufficiently small interval centered at x therefore avoids K, so

$$\limsup_{r \downarrow 0} \frac{\sigma((x-r, x+r))}{2r} \leq M(\sigma|_{\mathbb{R}\setminus K})(x).$$

For fixed $\lambda > 0$, the set where this limsup exceeds λ is contained in $K \cup \{M(\sigma|_{\mathbb{R}\setminus K}) > \lambda\}$. Its Lebesgue outer measure is at most $3\varepsilon/\lambda$. Let $\varepsilon \downarrow 0$. The bad set is outer null and hence Lebesgue measurable and null. Taking the countable union over $\lambda = 1/k$ proves zero centered density almost everywhere. This argument does not assume measurability of the limsup before establishing its nullity.

For one sided intervals, work at a point where the singular centered density is zero and x is a Lebesgue point of h. For $t > 0$,

$$\frac{\sigma((x, x+t])}{t} \leq 4\frac{\sigma((x-2t, x+2t))}{4t} \longrightarrow 0.$$

The larger open interval includes the possibly closed right endpoint. Also,

$$\Big|\frac{1}{t}\int_x^{x+t} h(s)\,ds - h(x)\Big| \leq 4\frac{1}{4t}\int_{x-2t}^{x+2t} |h(s) - h(x)|\,ds \longrightarrow 0.$$

The interval to the left has the same estimates. This proves both assertions. □

Corollary 12.16 (Differentiability of increasing functions)**.** *An increasing real function on a compact interval $[a, b]$ has a finite derivative almost everywhere in (a, b). This derivative is nonnegative, integrable, and satisfies*

$$\int_a^b F'(x)\,dx \leq F(b) - F(a).$$

Proof. First suppose F is right continuous. On any compact subinterval of (a, b) extend it, if necessary, to an increasing right continuous function on $\mathbb{R}$ that is constant beyond a slightly larger interval. Its Stieltjes measure is finite. The right and left difference quotients are respectively the measures of $(x, x+t]$ and $(x-t, x]$, divided by t. Theorem 12.15 shows that both converge almost everywhere to the same finite density. A countable exhaustion of (a, b) proves existence of the derivative there. On every inner interval its integral is bounded by the corresponding Stieltjes mass, and hence by $F(b) - F(a)$. Let the inner intervals increase to (a, b) and use monotone convergence for the nonnegative derivative.

For a general increasing F, put $G(x) = F(x+)$ at interior points. This is increasing and right continuous: if $y > x$ is chosen with $F(y) < G(x) + \varepsilon$, then $G(x + h) \leq F(y)$ for $0 < h < y - x$, while $G(x + h) \geq G(x)$. An increasing function has at most countably many discontinuities. Indeed, for each discontinuity the open interval between its left and right limits has positive length; these intervals for distinct points are disjoint, so assigning a rational point in each gives countability. Equivalently one can group the jumps by a positive lower bound on their sizes.

At a continuity point x where G has finite derivative d, one has $F(x) = G(x)$. For $h > 0$ and $0 < \eta < 1$,

$$G(x + (1 - \eta)h) \leq F(x + h) \leq G(x + h).$$

Subtract $G(x)$, divide by h, and let $h \downarrow 0$. The lower and upper limits lie between $(1 - \eta)d$ and d. Let $\eta \downarrow 0$. For $h < 0$, use

$$G(x + (1 + \eta)h) \leq F(x + h) \leq G(x + h)$$

and remember that division by h reverses the inequalities. The resulting limits lie between d and $(1 + \eta)d$, and again $\eta \downarrow 0$ gives derivative d. Thus $F' = G'$ almost everywhere. The integral bound follows by applying the right continuous case on inner intervals: their increments of G are at most $F(b) - F(a)$, and exhaustion finishes the proof. □

12.6 Bounded variation and absolute continuity

The almost everywhere derivatives just obtained are measurable. On an inner interval, the measurable difference quotients $k\big(F(x + 1/k) - F(x)\big)$ converge to $F'(x)$ wherever the derivative exists. Define the derivative to be zero on the null exceptional set and apply the measurability of almost everywhere limits. A countable exhaustion by inner intervals gives a measurable derivative throughout (a, b).

Definition 12.17. For a real function $F : [a, b] \to \mathbb{R}$, define its total variation by

$$\operatorname{Var}_{[a,b]} F = \sup_{a=x_0<\cdots<x_N=b} \sum_{j=1}^{N} |F(x_j) - F(x_{j-1})|.$$

It belongs to $BV[a, b]$ if this is finite. Set $V(x) = \operatorname{Var}_{[a,x]} F$, with $V(a) = 0$.

Proposition 12.18 (Variation and increasing functions)**.** *Variation is additive over adjacent intervals. If $F \in BV[a,b]$, the functions*

$$P(x) = \frac{V(x) + F(x) - F(a)}{2}, \quad N(x) = \frac{V(x) - F(x) + F(a)}{2}$$

are increasing, $F = F(a) + P - N$, and $V = P + N$. Every BV function has a derivative almost everywhere, with

$$\int_a^b |F'(x)|\, dx \le \operatorname{Var}_{[a,b]} F.$$

Proof. For $a < c < b$, refine any partition to contain c. The triangle inequality ensures that refinement cannot decrease the variation sum. The refined sum is bounded above by the sum of the two variations on $[a,c]$ and $[c,b]$. Conversely, choose partitions on those two intervals whose sums are within $\varepsilon/2$ of their variations and join them. Letting $\varepsilon \downarrow 0$ proves additivity.

Thus, for $x < y$, $V(y) - V(x) = \operatorname{Var}_{[x,y]} F \ge |F(y) - F(x)|$. This proves that both P and N are increasing. Their algebraic identities follow directly. By Corollary 12.16, both are differentiable almost everywhere and their nonnegative derivatives are integrable. Outside the union of their exceptional null sets, $F' = P' - N'$. Hence

$$\int_a^b |F'| \le \int_a^b (P' + N') \le P(b) - P(a) + N(b) - N(a) = V(b).$$

□

Recall that F is absolutely continuous on $[a,b]$ if, for every $\varepsilon > 0$, there is $\delta > 0$ such that

$$\sum_{j=1}^{N} |F(b_j) - F(a_j)| < \varepsilon$$

whenever the intervals (a_j, b_j) are pairwise disjoint in $[a,b]$ and $\sum_j (b_j - a_j) < \delta$. Finite families suffice in this definition. The corresponding conclusion for a countable family follows by applying it to finite partial families and passing to the limit, with a smaller initial error if a strict inequality is desired.

Lemma 12.19. *An absolutely continuous function belongs to BV. Its variation function V, and hence the increasing functions P, N above, are absolutely continuous.*

Proof. Choose $\delta > 0$ corresponding to error one in absolute continuity of F. Divide $[a,b]$ into finitely many intervals of lengths smaller than δ. On each such interval every partition sum is at most one, since the total length of the partition intervals is smaller than δ. The variation of F on $[a,b]$ is therefore at most the number of these intervals, by additivity of variation.

For absolute continuity of V, fix $\varepsilon > 0$ and choose δ for F with error $\varepsilon/2$. Let (a_j, b_j) be a finite disjoint family of total length less than δ. Within each $[a_j, b_j]$ take an arbitrary finite partition. All the resulting small intervals have disjoint interiors and their total length is still below δ. Thus the sum of their absolute increments of F is less than $\varepsilon/2$. Taking suprema over the finitely many partitions gives

$$\sum_j \bigl(V(b_j) - V(a_j)\bigr) = \sum_j \operatorname{Var}_{[a_j,b_j]} F \le \varepsilon/2 < \varepsilon.$$

The formulas for P, N express them as linear combinations of absolutely continuous functions, so they too are absolutely continuous. □

Lemma 12.20. *If F is increasing and absolutely continuous on $[a,b]$, its Stieltjes measure, obtained by extending F constantly outside $[a,b]$, is absolutely continuous with respect to Lebesgue measure.*

Proof. Absolute continuity implies continuity by applying the definition to a single short interval. Thus the constant extension is increasing and continuous, and its Stieltjes measure ν has no atoms, is concentrated on $[a,b]$, and has total mass $F(b) - F(a)$.

Let $A \subset \mathbb{R}$ be Borel with $m(A) = 0$, and fix $\varepsilon > 0$. Choose $\delta > 0$ in the definition of absolute continuity of F, with error $\varepsilon/2$. Cover $A \cap (a,b)$ by an open set of Lebesgue measure less than δ, and intersect the open set with (a,b). It is a countable disjoint union of open intervals (a_j, b_j), where endpoints may equal a or b. Since F is continuous, each such interval has ν measure $F(b_j) - F(a_j)$. For every finite partial family, absolute continuity bounds the sum of these increments by $\varepsilon/2$, because its total length is smaller than δ. Countable additivity then gives $\nu(A \cap (a,b)) \le \varepsilon/2$. The endpoints are atomless and the measure is zero outside $[a,b]$. Hence $\nu(A) \le \varepsilon/2$. Letting $\varepsilon \downarrow 0$ proves $\nu \ll m$. □

Theorem 12.21 (The fundamental theorem for absolutely continuous functions)**.** *For a real or complex function F on $[a,b]$, the following are equivalent:*

(i) *F is absolutely continuous.*

(ii) *There exists $h \in L^1(a,b)$ such that $F(x) = F(a) + \int_a^x h(t)\,dt$ for every $x \in [a,b]$.*

(iii) *F is differentiable almost everywhere, $F' \in L^1(a,b)$, and $F(x) = F(a) + \int_a^x F'(t)\,dt$ for every $x \in [a,b]$.*

In the second assertion, $h = F'$ almost everywhere.

Proof. First suppose F is real, increasing, and absolutely continuous. By Lemma 12.20, its finite Stieltjes measure ν satisfies $\nu \ll m$. Radon–Nikodym gives $h \in L^1(a,b)$ with $\nu = h\,dx$, taking $h = 0$ outside the interval. For every $x \in [a,b]$,

$$F(x) - F(a) = \nu((a,x]) = \int_a^x h(t)\,dt.$$

This is an identity at every point, not merely an almost everywhere identity.

For a general real absolutely continuous F, Lemma 12.19 supplies increasing absolutely continuous functions P, N with $F = F(a) + P - N$ and $P(a) = N(a) = 0$. Apply the increasing case to each and subtract their integrable densities. This proves (i)⇒(ii). For a complex function apply the argument to the real and imaginary parts. Their absolute continuity follows because the absolute increments of each component are bounded by those of F.

Assume (ii). For any disjoint finite interval family,

$$\sum_j |F(b_j) - F(a_j)| \le \sum_j \int_{a_j}^{b_j} |h(t)|\,dt = \int_{\bigcup_j (a_j,b_j)} |h(t)|\,dt.$$

The absolute continuity of the L^1 integral makes this less than ε when the total length is sufficiently small. Thus F is absolutely continuous. The differentiation theorem for indefinite integrals in Chapter 7 gives $F' = h$ at every Lebesgue point of h, hence almost everywhere, proving (iii). Assertion (iii) is a special case of (ii). This proves all implications without using the desired representation in the proof of its converse. □

Corollary 12.22 (Variation and integration by parts)**.** *For absolutely continuous $F : [a,b] \to \mathbb{R}$,*

$$\mathrm{Var}_{[a,b]} F = \int_a^b |F'(t)|\,dt.$$

If F, G are real or complex absolutely continuous functions, then their product is absolutely continuous and

$$\int_a^b F(t)G'(t)\,dt = F(b)G(b) - F(a)G(a) - \int_a^b F'(t)G(t)\,dt.$$

Proof. Put $h = F'$. For each partition, the fundamental theorem and the triangle inequality bound its variation sum by $\int |h|$. This gives the upper bound. For the reverse bound, let h_k equal the average of h on each interval of the uniform dyadic partition of $[a,b]$. Averaging is a contraction on L^1: on every interval I, $m(I)|h_I| \le \int_I |h|$, and summation proves the norm bound. Moreover $h_k \to h$ in L^1. To verify this, approximate h in L^1 by a continuous function q. The averages of q converge uniformly to q by uniform continuity, and

$$\|h_k - h\|_1 \le 2\|h - q\|_1 + \|q_k - q\|_1.$$

First let $k \to \infty$ and then make the approximation error arbitrarily small. On the dyadic partition,

$$\sum_I |F(\sup I) - F(\inf I)| = \sum_I \Big|\int_I h\Big| = \int_a^b |h_k|.$$

The right side tends to $\int |h|$, because $h_k \to h$ in L^1. Every left side is bounded by the variation. This proves equality.

Absolutely continuous functions are continuous and bounded on $[a, b]$. The identity

$$F(y)G(y) - F(x)G(x) = F(y)(G(y) - G(x)) + G(x)(F(y) - F(x))$$

shows absolute continuity of the product after summing over a disjoint interval family and using the two uniform bounds. Outside the union of the null sets where either derivative fails, the ordinary product rule gives $(FG)' = F'G + FG'$. This derivative is integrable because F, G are bounded. Apply the fundamental theorem to FG and rearrange the finite integrals. □

This completes the interval characterization used in Chapter 9: the elements of $W^{1,p}(a, b)$, $1 \le p \le \infty$, are precisely the almost everywhere classes with an absolutely continuous representative whose derivative is in L^p. On a bounded interval such a representative is bounded, hence itself in L^p. Its weak derivative agrees with the classical derivative almost everywhere, by the integral representation and the test function identity.

Proposition 12.23 (Signed Stieltjes measures for BV functions)**.** *Let $F \in BV[a, b]$ be right continuous on $[a, b)$, and extend it constantly outside $[a, b]$. There is a finite signed Borel measure ν_F concentrated on $(a, b]$ such that $\nu_F((x, y]) = F(y) - F(x)$ for $a \le x < y \le b$. If $V(x) = \mathrm{Var}_{[a,x]} F$, extended constantly as well, then $|\nu_F| = \mu_V$.*

Proof. We first check right continuity of V, which is needed to use the Stieltjes construction. Fix $x < b$. Since V is bounded and increasing, $V(x+)$ exists. For a partition $x = t_0 < t_1 < \cdots < t_N = x + h$, its sum is bounded by

$$|F(t_1) - F(x)| + \mathrm{Var}_{[t_1,x+h]} F \le \sup_{x<t\le x+h} |F(t) - F(x)| + V(x + h) - V(x+).$$

Taking the supremum gives $V(x+h)-V(x)$ on the left. Rearranging the finite quantities yields

$$0 \le V(x+) - V(x) \le \sup_{x<t\le x+h} |F(t)-F(x)|.$$

Right continuity of F makes the last quantity tend to zero. Thus V is right continuous. The increasing functions P, N in Proposition 12.18 are right continuous too.

Let $\nu_F = \mu_P - \mu_N$. The increment formula follows by subtraction, and the measures are finite since P, N are constant outside $[a, b]$. They have no atom at a, since their left extensions have the same values as at a. Their sum is μ_V: the two positive measures have the same interval increments, and uniqueness of the Stieltjes measure gives equality. Triangle inequalities on finite partitions therefore imply $|\nu_F| \le \mu_V$.

For any partition $a = x_0 < \cdots < x_N = b$, the variation of the signed measure on $(a, b]$ is at least

$$\sum_{j=1}^{N} |\nu_F((x_{j-1}, x_j])| = \sum_{j=1}^{N} |F(x_j) - F(x_{j-1})|.$$

Take the supremum. This gives $|\nu_F|((a,b]) \ge V(b) = \mu_V((a,b])$. Since both measures are concentrated on $(a, b]$, the nonnegative finite measure $\mu_V - |\nu_F|$ has total mass zero, hence vanishes on every Borel set. This proves the claimed identity. □

Corollary 12.24 (Absolutely continuous and singular parts of a BV function)**.** *Under the hypotheses of Proposition 12.23, there is a decomposition $F = F_a + F_s$ on $[a, b]$ such that F_a is absolutely continuous, $F_s \in BV[a, b]$, and $F_s' = 0$ almost everywhere. It is uniquely determined by $F_a(a) = F(a)$ and $F_s(a) = 0$. In particular,*

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

Proof. Apply Lebesgue decomposition to the positive and negative parts of ν_F. This gives $\nu_F = h\,dx + \sigma$, with $h \in L^1(a, b)$ and a finite signed σ concentrated on a Borel Lebesgue null set. Put

$$F_a(x) = F(a) + \int_a^x h(t)\,dt, \quad F_s(x) = \sigma((a, x]).$$

The increment formula gives $F = F_a + F_s$ at every point. The variation of F_s is at most $|\sigma|((a, b])$, by summing the absolute signed increments on any partition, so $F_s \in BV$. The measure $|\sigma|$ is singular. The proof of Theorem 12.15, applied to this positive singular measure, gives zero one sided densities almost everywhere. Since the absolute value of every increment of F_s is bounded by the corresponding $|\sigma|$ measure, both one sided difference quotients tend to zero almost everywhere. Thus $F_s' = 0$ and $F' = h$ almost everywhere.

For uniqueness, subtract two such decompositions. The difference of their absolutely continuous parts is absolutely continuous and has derivative zero almost everywhere, since it also equals the difference of two singular functions whose derivatives are zero almost everywhere. Theorem 12.21 makes it constant. The prescribed values at a make that constant zero. □

Example 12.25 (The Cantor function)**.** There is a continuous increasing function $F : [0,1] \to [0,1]$ with $F(0) = 0$, $F(1) = 1$, and $F' = 0$ almost everywhere. It is not absolutely continuous, and its Stieltjes measure is a singular probability measure.

Proof. Let C_k be the kth stage of the middle thirds Cantor construction. It has 2^k closed interval components. Define F_k to be constant on the gaps already removed and affine on each remaining component, with the image of each component an interval of length 2^{-k}. Start with $F_0(x) = x$. In passing to the next stage, keep the endpoint values of each component, replace the middle third of its graph by a constant segment at the midpoint of its range, and use affine increasing pieces on the two remaining thirds. This gives a continuous increasing F_{k+1} with the same endpoint values on every old component.

On an old component both F_k and F_{k+1} lie in the same range interval of length 2^{-k}, and on old gaps they agree. Thus $\|F_{k+1} - F_k\|_\infty \le 2^{-k}$. The geometric series of these bounds shows that (F_k) is uniformly Cauchy. Its limit F is continuous, increasing, and has endpoint values zero and one. Every removed open gap is a region on which all subsequent F_k are the same constant, so F is locally constant there. The complement of the Cantor set is the union of these gaps, and the Cantor set has Lebesgue measure zero. Hence $F' = 0$ almost everywhere.

If F were absolutely continuous, its endpoint difference would equal $\int_0^1 F' = 0$, contradicting $F(1) - F(0) = 1$. Its Stieltjes measure has total mass one. Every complementary gap has zero measure by the increment formula and continuity at its endpoints. The constant extension outside $[0,1]$ contributes no mass. Thus the measure is concentrated on the null Cantor set, and is singular.

Finally F does not represent an element of $W^{1,1}(0,1)$. Such membership would provide an absolutely continuous representative U equal to F almost everywhere. Since both U and F are continuous, they would agree everywhere: a nonzero value of their continuous difference would persist on a nonempty interval of positive measure. This would make F absolutely continuous, a contradiction. □

The distinction is therefore between existence of an almost everywhere derivative and recovery from that derivative. Bounded variation ensures an integrable derivative but allows a singular part. Absolute continuity removes that singular part and gives the pointwise integral representation.

Exercises

Exercise 12.1. Prove that a finite positive measure is σ-finite, and give a complete measure space that is not σ-finite. Show that continuity from above can fail without a finite-measure term.

Exercise 12.2. Describe the completion of a Borel measure explicitly as sets differing from Borel sets by subsets of Borel null sets. Prove that the integral of a completed-measurable integrable function agrees with that of a Borel measurable representative.

Exercise 12.3. Let μ_* be an outer measure. Show that an arbitrary subset of an outer null set satisfies the Carathéodory criterion. Explain why countable unions of outer null sets are outer null.

Exercise 12.4. Verify every closure assertion in the proof of the π–λ principle. Give an example of a λ-system that is not a σ-algebra.

Exercise 12.5. Let μ be counting measure on all subsets of $[0, 1]$ and consider instead the common Borel σ-algebra for μ and Lebesgue measure ν. Prove $\nu \ll \mu$, but prove that no nonnegative Borel h satisfies $\nu = h\mu$. Identify the failed hypothesis of Radon–Nikodym.

Exercise 12.6. If $\nu \ll \mu$ and $\mu \ll \lambda$ are σ-finite positive measures, prove the chain rule $d\nu/d\lambda = (d\nu/d\mu)(d\mu/d\lambda)$ almost everywhere, with an explicit convention on zero-density sets.

Exercise 12.7. For a finite signed measure $\nu = h\mu$, prove directly from finite partitions that $|\nu| = |h|\mu$. Use this to identify its positive and negative parts.

Exercise 12.8. Prove uniqueness of the decomposition into mutually singular positive measures in the Jordan theorem. Explain why an arbitrary expression as the difference of two positive measures need not be unique.

Exercise 12.9. For $1 < p < \infty$, determine a unit-norm $f \in L^p$ that attains the norm of the functional represented by a nonzero $g \in L^{p'}$. Compare the situation with $p = 1$.

Exercise 12.10. Compute the Stieltjes measure of an increasing right continuous step function with finitely many jumps. Then treat a countable sum of positive jumps whose total size is finite.

Exercise 12.11. For a locally finite Stieltjes measure, derive the formulas for $\mu_F([a,b])$, $\mu_F((a,b))$, and $\mu_F([a,b))$ in terms of values and left limits of F.

Exercise 12.12. Give a continuous function on $[0,1]$ that is not of bounded variation. Prove the claim by constructing partitions with unbounded variation sums.

Exercise 12.13. Prove that Lipschitz functions on a compact interval are absolutely continuous. Deduce their almost everywhere differentiability and prove that the derivative is essentially bounded by the Lipschitz constant.

Exercise 12.14. If $F : [a,b] \to \mathbb{R}$ is absolutely continuous and $\Phi \in C^1(\mathbb{R})$, prove that $\Phi \circ F$ is absolutely continuous and $(\Phi \circ F)' = \Phi'(F)F'$ almost everywhere. Use boundedness of Φ' on the compact range of F, not a global bound.

Exercise 12.15. For the Cantor function, prove explicitly that the sum of its increments over the 2^k components of C_k is one while their total length tends to zero. Deduce failure of absolute continuity directly from its definition.

Exercise 12.16. Let $F \in BV[a,b]$ be right continuous. Show that F is absolutely continuous if and only if its signed Stieltjes measure is absolutely continuous with respect to Lebesgue measure. Include the endpoint convention in your argument.

Notes. The measure construction and representation arguments are standard parts of abstract integration theory; see [1, 2]. The last section supplies the full converse statement for indefinite integrals and the interval Sobolev interpretation used earlier. The distinctions between null-set modifications, representatives, ordinary derivatives, and weak derivatives remain essential even in one dimension.

Bibliography

[1] E. M. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton Lectures in Analysis III, Princeton University Press, Princeton, 2005.

[2] R. L. Wheeden and A. Zygmund, *Measure and Integral: An Introduction to Real Analysis*, second edition, CRC Press, Boca Raton, 2015.

[3] T. Tao, *Analysis I*, Hindustan Book Agency. Course reference edition supplied with these notes.

[4] T. Tao, *Analysis II*, Hindustan Book Agency. Course reference edition supplied with these notes.