

Real Analysis 1
Math 623

Notes by Patrick Lei, May 2020

Lectures by Robin Young, Fall 2018

DISCLAIMER

These notes are a May 2020 transcription of handwritten notes taken during lecture in Fall 2018. Any errors are mine and not the instructor's. In addition, my notes are picture-free (but will include commutative diagrams) and are a mix of my mathematical style (omit lengthy computations, use category theory) and that of the instructor. If you find any errors, please contact me at plei@umass.edu.

Contents

Contents • 2

- 1 Lebesgue Measure • 3
 - 1.1 MOTIVATION FOR THE COURSE • 3
 - 1.2 BASIC NOTIONS • 4
 - 1.3 RECTANGLES • 5
 - 1.4 OUTER MEASURE • 6
 - 1.5 MEASURABLE SETS • 7
 - 1.6 MEASURABLE FUNCTIONS • 9
- 2 Integration • 12
 - 2.1 DEFINING THE INTEGRAL • 12
 - 2.2 SOME CONVERGENCE RESULTS • 14
 - 2.3 VECTOR SPACE OF INTEGRABLE FUNCTIONS • 15
 - 2.4 SYMMETRIES • 17
- 3 Differentiation • 19
 - 3.1 INTEGRATION AND DIFFERENTIATION • 19
 - 3.2 KERNELS • 21
 - 3.3 DIFFERENTIATION • 25
 - 3.4 BOUNDED VARIATION • 27
 - 3.5 ABSOLUTE CONTINUITY • 29
- 4 General Measures • 31
 - 4.1 OUTER MEASURES • 32
 - 4.2 RESULTS AND APPLICATIONS • 33
 - 4.3 SIGNED MEASURES • 34

Lebesgue Measure

1.1 MOTIVATION FOR THE COURSE

We will consider some examples that show us some fundamental problems in analysis.

Example 1.1 (Fourier Series). Let f be periodic on $[0, 2\pi]$. Then if f is sufficiently nice, we can write

$$f(x) = \sum_n a_n e^{inx},$$

where $a_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx$. If this is possible, then we have Parseval's identity:

$$\sum |a_n|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx.$$

However, the function space we are working in on the RHS is not complete because there exist sequences a_n such that the associated function $|f|^2$ is not Riemann-integrable. This tells us that we need a better notion of integration.

Example 1.2 (Dirichlet Function). Enumerate the rationals in the interval $[0, 1]$ by r_n . Then define

$$f_N(x) = \begin{cases} 1 & x \in r_1, \dots, r_N \\ 0 & \text{otherwise} \end{cases}.$$

Then the limit of f_N is simply the indicator function $f = \chi_{\mathbb{Q}}$. Here, the f_N are Riemann integrable, but f is not.

Next, we can consider the problem of characterizing which functions are Riemann-integrable.

Theorem 1.3. *A function f is Riemann-integrable if and only if the set of points where f is not continuous has measure 0.*

Fourth, we may consider lengths of curves. Suppose we have a fractal (this will remain undefined) such as the Cantor set, a coastline, or broccoli. What is the correct dimension for such a set? When is the “length” of such a curve finite?

Finally, recall the fundamental theorem of calculus: For a differentiable function F ,

$$F(b) - F(a) = \int_a^b F'(t) dt$$

and for an integrable function f ,

$$\frac{d}{dx} \int_a^x f(t) dt = f(x).$$

When do these two results apply?

Question 1.4. *All of these examples tell us that the fundamental problem in analysis is the exchange of limits. We will see that many of the results in the course are of this form.*

1.2 BASIC NOTIONS

First recall the very basic notions of set theory: unions, intersections, complements, relations, and functions. We use the usual definitions of image, preimage, and partial and total orders.¹ We will also assume the axiom of choice, so in particular, we have the following results:

Hausdorff Maximal Principle: Every partially ordered set has a maximal totally ordered subset.

Zorn's Lemma: If X is a partially ordered set where every chain has an upper bound, then X has a maximal element.

Well Ordering Principle: Every set admits a well-ordering.

Next, we will recall some basic properties of Euclidean space.

Theorem 1.5 (Bolzano-Weierstrass). *Every bounded sequence in \mathbb{R}^n has a limit point. Note that the lim sup is an accumulation point.*

Next recall that \mathbb{R}^n with the Euclidean distance ρ is a metric space. In fact, \mathbb{R}^n is a complete normed metric space, and the triangle inequality is implied by Cauchy-Schwarz. Recall that open balls are given by $B_r(x) = \{y \mid \rho(x, y) < r\}$ and closed balls are given by $\bar{B}_r(x) = \{y \mid \rho(x, y) \leq r\}$. Then a set G is *open* if around any $x \in G$ we can find an open ball $B_r(x) \subset G$ and a set F is *closed* if F^c is open. We can then define the interior, closure, and boundary of a set in the usual way.

Example 1.6. Consider $\mathbb{Q} \subset \mathbb{R}$. Then $\mathbb{Q}^\circ = \emptyset$ and $\bar{\mathbb{Q}} = \mathbb{R}$, so \mathbb{Q} is a countable dense subset of \mathbb{R} . Therefore, \mathbb{R} is *separable*.

Convergence is defined in the usual way.

Lemma 1.7. *Let X be a metric space and $E \subset X$. Then the following are equivalent:*

1. $x \in \bar{E}$;
2. For all $\varepsilon > 0$, $B_\varepsilon(x) \cap E \neq \emptyset$;
3. There exists a sequence $\{x_n\} \subset E$ such that $x_n \rightarrow x$.

¹The reader not familiar with these can refer back to their Math 300 materials.

1.3 RECTANGLES

We now build towards defining the Lebesgue measure. First, we call any set of the form

$$R = \prod_{i=1}^n [a_i, b_i]$$

a *closed rectangle*. We define the volume $|R|$ of any rectangle by

$$|R| = \prod_{i=1}^n (b_i - a_i).$$

It should be obvious what a cube and an open rectangle are.

Definition 1.8. A collection of rectangles is *almost disjoint* if their interiors are disjoint.

Lemma 1.9. If a rectangle R is an almost disjoint union $R = \bigcup R_j$, then $|R| = \sum |R_j|$.

Now let $\varepsilon > 0$ small enough. Then $R^\varepsilon = \prod [a_i + \varepsilon, b_i - \varepsilon] \subset R^\circ \subset R$. As $\varepsilon \rightarrow 0$, we obtain $|R^\circ| = |R|$.

Corollary 1.10. Let $R \subset \bigcup_{i=1}^n R_i$, where the union is not necessarily almost disjoint. Then $|R| \leq \sum |R_i|$.

Our goal is to extend this notion of volume to open sets in terms of rectangles.

Lemma 1.11. In \mathbb{R} , any open set G is a disjoint union of countable many open intervals.

Proof. Let $x \in G$. Then there is a maximal interval $I_x \subseteq G$. Then if $y \in I_x$, note that $I_x = I_y$. Thus $G = \bigcup_{x \in G} I_x$. However, each interval must contain a rational number, so there are only countably many of them. \square

Therefore, if $G \subseteq \mathbb{R}$ is open, we can declare its measure to be $\mu(G) = \sum |I_j|$, where G is a disjoint union of the intervals I_j . Note that this is not convenient in higher dimension, so we use closed rectangles instead. Note that any open set in \mathbb{R} is a countable union of closed rectangles $G = \bigcup [a_j + 1/n, b_j - 1/n]$, but this union is not even almost disjoint.

Theorem 1.12. In \mathbb{R}^d , any open set G can be written as an almost disjoint union $G = \bigcup_{j=1}^{\infty} F_j$ of closed rectangles. Given this, we expect $m(G) = \sum_{j=1}^{\infty} |F_j|$.

Proof. Grid \mathbb{R}^d by the points in \mathbb{Z}^d . Then each cube Q of length 1 satisfies either $Q \subseteq G$, $Q \subseteq \overline{G}^c$, or neither. At the first step, include all $Q \subseteq G$ and exclude all $Q \subseteq \overline{G}^c$. Then we reduce \mathbb{Z}^d to $\frac{1}{2}\mathbb{Z}^d$ and proceed by induction. \square

Example 1.13 (Cantor Set). This is not really an example of rectangles, but is an important example for this course. It is a closed, compact, totally disconnected set where every point is a limit point. Define the sets

- $C_0 = [0, 1]$;
- $C_1 = [0, 1/3] \cup [2/3, 1]$;
- $C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$,

and then continue the process. Then we define the *Cantor set* to be $C = \bigcap_{k=1}^{\infty} C_k$.

First, note that C is an intersection of closed sets, so it is closed. Then given $x, y \in C$, note that there exists k such that $x - y > \frac{2}{3^k}$, so x, y are in different intervals in C_k . Clearly there are no isolated points. Finally, note that C is the set of all real numbers with no 1 in their ternary expansions, so it is uncountable. However, note that C^c is a disjoint union of open intervals with measure

$$m(C^c) = \frac{1}{3} + 2\frac{1}{9} + 4\frac{1}{27} + \cdots = \frac{1}{3} \sum_{k=0}^{\infty} \left(\frac{2}{3}\right)^k = 1.$$

Therefore, we expect $m(C) = 0$.

1.4 OUTER MEASURE

It is reasonable to expect that if G is an almost disjoint union $G = \bigcup_j Q_j$ then $m(G) = \sum_j |Q_j|$. Instead of imposing this inequality for general E , we will make it an inequality, so we drop the condition of being almost disjoint. Define for any $E \subset \mathbb{R}^d$ the *outer measure*

$$m_*(E) = \inf_{E \subseteq \bigcup_j Q_j} \sum_j |Q_j|.$$

We can show that $m_*(C) = 0$ directly. Note that C_k is a union of 2^k closed subintervals of length 3^{-k} , so we see that $m_*(C) \leq \left(\frac{2}{3}\right)^k$ for all k . Therefore, $m_*(C) = 0$. For another example, it is easy to show that if P is countable, then $m_*(P) = 0$. To see this, let $\varepsilon > 0$ and then take the rectangle of width $\frac{\varepsilon}{2^i}$ centered at p . Then $m_*(P) \leq 2\varepsilon$.

Properties of outer measure. Here are some properties of the outer measure:

1. (Monotonicity) If $E_1 \subseteq E_2$, then $m_*(E_1) \leq m_*(E_2)$. As a corollary, bounded sets have finite outer measure.
2. (Countable subadditivity) If $E \subseteq \bigcup_j E_j$, then $m_*(E) \leq \sum_j m_*(E_j)$.

Proof. Assume that each $m_*(E_j) < \infty$. Then for all j , there exist countably many cubes $Q_{j,k}$ such that $E_j \subseteq \bigcup_k Q_{j,k}$, so we have

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

Then any $E \subseteq \bigcup_j E_j \subseteq \bigcup_{j,k} Q_{j,k}$, so

$$m_*(E) \leq \sum_{j,k} |Q_{j,k}| \leq \sum_j \sum_k |Q_{j,k}| \leq \sum_j m_*(E_j) + \frac{\varepsilon}{2^j} \leq \sum_j m_*(E_j) + \varepsilon. \quad \square$$

3. For any $E \subseteq \mathbb{R}^d$, we can write

$$m_*(E) = \inf_{\substack{G \text{ open} \\ E \subseteq G}} m_*(G).$$

Proof. It is easy to see that $m_*(E) \leq \inf m_*(G)$. In the other direction, assume $m_*(E)$ is finite. Then $E \subseteq \bigcup_j Q_j$, so $\sum |Q_j| - \varepsilon \leq m_*(E_j) \leq \sum |Q_j|$. Choose G_j open such that $\overline{Q_j} \subseteq G_j$ and $|G_j| \leq |Q_j| + \frac{\varepsilon}{2^k}$.

Now $E \subseteq \bigcup G_j$ and $G = \bigcup G_j$, so

$$m_*(G) \leq \sum \left(|Q_j| + \frac{\varepsilon}{2^j} \right) \leq \sum |Q_j| + \varepsilon \leq m_*(E) + 2\varepsilon. \quad \square$$

4. (Wannabe additivity) If $E = E_1 + E_2$ and $d(E_1, E_2) > 0$, then $m_*(E) = m_*(E_1) + m_*(E_2)$. Here $d(A, B) = \inf\{d(x, y) \mid x \in A, y \in B\}$.

Proof. Choose $0 < \delta < d(E_1, E_2)$ and choose Q_j such that $E \subseteq \bigcup Q_j$, $m_*(E) > \sum |Q_j| - \varepsilon$, and $\text{diam}(Q_j) < \delta$. If $E_1 \cap Q_j \neq \emptyset$, then $E_2 \cap Q_j = \emptyset$. Therefore, there exist J_1, J_2 such that $E_1 \subseteq \bigcup_{j \in J_1} Q_j$ and $E_2 \subseteq \bigcup_{i \in J_2} Q_i$, so

$$m_*(E_1) + m_*(E_2) \leq \sum_{j \in J_1} |Q_j| + \sum_{i \in J_2} |Q_i| = \sum |Q_j| \leq m_*(E) + \varepsilon. \quad \square$$

5. If E is a countable union of almost disjoint cubes, then $m_*(E) = \sum |Q_j|$.

Proof. For each j , pick $\tilde{Q}_j \subseteq Q_j^\circ$ such that $|\tilde{Q}_j| > |Q_j| - \frac{\varepsilon}{2^j}$. Then $\bigcup \tilde{Q}_j \subseteq E$ and are disjoint and a finite distance apart, so for each N , we have $\sum_{j=1}^N m_*(\tilde{Q}_j) \leq m_*(E)$, and therefore $\sum_{j=1}^N |Q_j| - \varepsilon \leq m_*(E)$, and thus $\sum_{j=1}^\infty |Q_j| \leq m_*(E)$. \square

Note we cannot conclude finite additivity for general disjoint sets E_1, E_2 .

1.5 MEASURABLE SETS

Definition 1.14. A subset $E \subseteq \mathbb{R}^n$ is *measurable* if for all $\varepsilon > 0$, there exists $G \supseteq E$ such that $m_*(G \setminus E) < \varepsilon$.

If E is measurable, define its *Lebesgue measure* as $m(E) = m_*(E)$. Here are some properties of measurable sets:

1. Open sets are measurable;
2. If $m_*(E) = 0$, then E is measurable;
3. Countable unions of measurable sets are measurable;
4. Closed sets are measurable;
5. If E is measurable, then so is E^c ;
6. Countable intersections of measurable sets are measurable.

This says that measurable sets form a σ -algebra. In addition, we should note that if F is closed and K is compact such that $F \cap K = \emptyset$, then $d(F, K) > 0$.²

Theorem 1.15. If E_j are measurable and disjoint, then $m(\bigcup E_j) = \sum m(E_j)$.

Proof. Assume that the E_j are bounded. Then there exist F_j closed with $F_j \subseteq E_j$ such that $m(E_j \setminus F_j) < \frac{\varepsilon}{2^j}$ for all j . Then for finite N , $\bigcup_{j=1}^N F_j$ is compact, so $m(\bigcup F_j) = \sum m(F_j)$. Therefore $m(E) \geq \sum_{j=1}^N m(F_j) \geq \sum_{j=1}^N m(E_j) - \varepsilon$ uniformly in N . Therefore, $m(E) \geq \sum m(E_j)$. The other inequality is from subadditivity.

Next, if some of the E_j are unbounded, set Q_j to be the cube with sides $[-k, k]$. Then set $S_1 = Q_1$ and $S_{j+1} = Q_{j+1} \setminus Q_j$. Finally, set $E_{j,k} = E_j \cap S_k$. Then each $E_{j,k}$ is bounded, so we have

$$m(E) = \sum_{j,k} m(E_{j,k}) = \sum_j \sum_k m(E_{j,k}) = \sum_j m\left(\bigcup_k E_{j,k}\right) = \sum_j m(E_j). \quad \square$$

²This was proved in complex analysis.

Corollary 1.16. 1. If $E_k \nearrow E = \bigcup E_k$ with E_k measurable, then $m(E) = \lim m(E_k)$;

2. If $E_k \searrow F = \bigcap E_k$ and some E_k has finite measure, then $\lim m(E_k) = m(F)$.

Proof. 1. Set $D_k = E_{k+1} \setminus E_k$. Then the D_k are disjoint and measurable, so $E = \bigcup D_k$. Therefore,

$$m(E) = \sum_{k=1}^{\infty} m(D_k) = \lim_{N \rightarrow \infty} \sum_{k=1}^N m(D_k) = \lim m(E_k).$$

2. Set $B_k = E_k \setminus E_{k+1}$. Then $E_{k_0} = (\bigcap E_k) \cup (\bigcup_{k \geq k_0} B_k)$. Then $m(E_{k_0}) < \infty$ by assumption, so

$$\begin{aligned} m(E_{k_0}) &= m\left(\bigcap E_k\right) + \sum_{k \geq k_0} m(B_k) \\ &= m\left(\bigcap E_k\right) + \lim_{N \rightarrow \infty} \sum_{k=k_0}^N m(B_k) \\ &= m\left(\bigcap E_k\right) + m(E_{k_0}) - \lim_{N \rightarrow \infty} m(E_N). \end{aligned}$$

Therefore $m(\bigcap E_k) = \lim m(E_N)$. □

Theorem 1.17. If E is measurable, then for all $\varepsilon > 0$:

1. There exists an open $G \supseteq E$ such that $m(G \setminus E) < \varepsilon$;
2. There exists closed $F \subseteq E$ such that $m(E \setminus F) < \varepsilon$;
3. If $m(E) < \infty$, then there exists compact $K \subseteq E$ such that $m(E \setminus K) < \varepsilon$;
4. If $m(E) < \infty$, then there exists $F = \bigcup_{j=1}^N Q_j$ such that $m(E \Delta F) < \varepsilon$.

Proof. We only need to prove the last two, the first two are by definition.

3. There exist a closed F with $m(E \setminus F) < \varepsilon/2$. Then for all n , set $K_n = \overline{B_n} \cap F$. Then the $E \setminus K_n$ are measurable, and converge down to $E \setminus F$, so there exists n such that $m(E \setminus K_n) < \varepsilon/2 + m(E \setminus F) < \varepsilon$.
4. Note there exists $G = \bigcup_{j=1}^{\infty} Q_j$ containing E such that $m(G) < m(E) + \varepsilon/2$. Then there exists N such that $\sum_{j > N} |Q_j| < \varepsilon/2$, so we can set $F = \bigcup_{i=1}^N Q_i$. Then

$$\begin{aligned} m(E \Delta F) &= m(E \setminus F) + m(F \setminus E) \\ &\leq m\left(\bigcup_{j > N} Q_j\right) + m\left(\bigcup_{j=1}^{\infty} Q_j \setminus E\right) \\ &\leq \sum_{j=N}^{\infty} |Q_j| + m(G \setminus E) \\ &< \varepsilon. \end{aligned} \quad \square$$

Corollary 1.18. E is measurable if and only if

1. E differs from a G_{δ} set by a null set;
2. E differs from an F_{σ} set by a null set.

- Proof.* 1. For all n there exists G_n open such that $m(G_n \setminus E) < 1/n$ and $G_n \supseteq E$. Therefore, set $S = \bigcap G_n$. Clearly $E \subseteq S$ and $m(S \setminus E) = 0$ by Corollary 1.16.
2. The argument is the same. □

Other Properties of Lebesgue Measure

1. m is translation invariant;
2. If E is measurable, then $m(\lambda E) = \lambda^d m(E)$;
3. m is reflection invariant.

Definition 1.19. A set $\Sigma \subset \mathcal{P}(X)$ is a σ -algebra if it is closed under complements and countable unions.

The *Borel Sets* are the smallest σ -algebra containing the open balls.

Lemma 1.20. If $E \subseteq \mathbb{R}$ is measurable and $m(E) > 0$, then the set of differences $E \ominus E = \{x - y \mid x, y \in E\}$ contains an open interval around 0.

Proof. There exists an open set $G \supseteq E$ such that $m(G) \leq (1 + \varepsilon)m(E)$. Then G is a disjoint union $G = \bigcup_{k=1}^{\infty} I_k$ of open intervals. Set $E_k = I_k \cap E$. Then we know that $\sum |I_k| \leq (1 + \varepsilon) \sum m_*(E_k)$, so there exists k_0 such that $|I_{k_0}| \leq m_*(E_{k_0})$. Then for small δ if E_{k_0} and $E_{k_0} + \delta$ are disjoint, then $m(E_{k_0} \cup E_{k_0} + \delta) \geq \frac{2}{1+\varepsilon}|I_{k_0}|$. However, if $\delta < |I_{k_0}|$, then $I_{k_0} \cup I_{k_0} + \delta$ is an interval. Thus $m(E_{k_0} \cup E_{k_0} + \delta) \leq |I_{k_0}| + \delta$. Finally, we see that $\delta \geq \frac{1-\varepsilon}{1+\varepsilon}|I_{k_0}|$, so $E_{k_0} \cap E_{k_0} + \delta \neq \emptyset$ for small enough δ . □

Theorem 1.21 (Vitali). There is a non-measurable set $S \subset [0, 1]$.

Proof. Enumerate the rationals in $[0, 1]$ by r_k . Now consider the set \mathbb{R}/\mathbb{Q} and choose one representative of each equivalence class using the axiom of choice.³ Call this set P . Next, note that $P \ominus P$ contains no interval because it contains no rationals. If $m_*(P) = 0$, then $[0, 1] \subseteq \bigcup_{r \in \mathbb{Q}} P + r$, which would imply that $m([0, 1]) = 0$. Thus P must be non-measurable. □

Corollary 1.22. Any set with positive outer measure contains a non-measurable set.

Proof. Let P be as in Theorem 1.21. Then let $A_r = A \cap (P + r)$. At least one of these sets must be non-measurable. □

1.6 MEASURABLE FUNCTIONS

Our goal will be integration. We expect that if E is a measurable set, then $\int \chi_E = m(E)$. More generally, we want

$$\int \sum_{k=1}^N a_k \chi_{E_k} = \sum_{k=1}^N m(E_k).$$

Functions of this form are called *simple*. We also want to have some results where the integral of a limit is equal to the limit of the integral. Similar to continuity, we can try to define f to be measurable if $f^{-1}(E)$ is measurable for all measurable E . However, this is too hard to check, we find a minimal condition.

Definition 1.23. A function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is *measurable* if for all a , $f^{-1}(-\infty, a)$ is measurable.

³If we assume choice is false, then every subset of the real numbers is measurable.

Here are some properties of measurable functions:

1. f is measurable if and only if $f^{-1}(G)$ is measurable for all open G . Alternatively, $f^{-1}(F)$ is measurable for all closed F .
2. Continuous functions are measurable. In addition, if f is measurable and finite-valued and Φ is continuous, then $\Phi \circ f$ is measurable.
3. If $\{f_n\}$ is a sequence of measurable functions, then $\sup f_n, \inf f_n, \liminf f_n, \limsup f_n$, and $\lim f_n$ (if it exists) are all measurable.
4. If f, g are measurable, then $f + g, fg$ are both measurable.
5. Define $f = g$ *almost everywhere* if the set on which they differ has measure 0. If $f = g$ almost everywhere, then f is measurable if and only if g is measurable.

We will now prove some approximation theorems for measurable functions. First, we attempt to approximate them by simple functions.

Theorem 1.24. *If $f \geq 0$ is measurable on \mathbb{R}^d , then there exists an increasing sequence ϕ_k of simple functions converging pointwise to f .*

Proof. Define the sequence of functions

$$f_N(x) = \begin{cases} f(x) & x \in Q_N \text{ and } f(x) \leq N \\ N & x \in Q_N \text{ and } f(x) > N \\ 0 & x \notin Q_N \end{cases}.$$

Clearly $f_N(x) \nearrow f(x)$ for all x and $f_N(\mathbb{R}) \subseteq [0, N]$. Now define the set

$$E_{\ell, m}^N = \left\{ x \in Q_N \mid \frac{\ell}{m} < f_N(x) \leq \frac{\ell+1}{m} \right\}$$

and the simple function

$$f_{N, M}(x) = \sum_{\ell=0}^{MN} \frac{\ell}{M} \chi_{E_{\ell, M}^N}.$$

Clearly $f_{N, M}$ is increasing in M . We also see that $0 \leq f_N(x) - f_{N, M}(x) \leq \frac{1}{M}$. Now set $N = M = 2^k$. Then we see that

$$\lim_{k \rightarrow \infty} f_{2^k, 2^k} = f(x)$$

for all x . □

Corollary 1.25. *Suppose f is measurable. Then there exists a sequence ϕ_k simple such that $|\phi_k(x)| \nearrow |f(x)|$ and $\lim \phi_k(x) = f(x)$.*

Proof. Write $f(x) = f_+(x) - f_-(x)$, where $f_+(x) = \max\{0, f(x)\}$ and $f_-(x) = \max\{-f(x), 0\}$. The rest is left to the reader. □

Theorem 1.26. *If f is measurable, there exists a sequence ψ_k of step functions such that $\psi_k \rightarrow f$ almost everywhere.*

Proof. First, there exist φ_k simple converging to f everywhere. Write $\varphi_k = \sum_{j=1}^{N_k} a_{j,k} \chi_{E_{j,k}}$. Then for all $E_{j,k}$, there exists a finite collection of cubes Q_ℓ such that $m(E_\Delta \cup Q_\ell) < \varepsilon$. Therefore, there exist almost disjoint rectangles \tilde{R}_p such that $\bigcup Q_\ell = \bigcup \tilde{R}_p$. Shrink the \tilde{R}_p to obtain disjoint closed rectangles R_p such that $m(E_\Delta \cup R_p) < 2\varepsilon$. Then there exist step functions ψ_k and measurable sets F_k such that

1. $m(F_k) < \frac{1}{2^k}$;
2. $\{\varphi_k \neq \psi_k\} \subseteq F_k$.

Then set $F = \limsup F_k$, so F has measure 0. Then for all $x \notin F$, we have

$$\lim_{k \rightarrow \infty} f(x) - \psi_k(x) = \lim_{k \rightarrow \infty} f(x) - \varphi_k(x) = 0. \quad \square$$

Theorem 1.27 (Egorov). *Let f_k be a sequence of measurable functions on E with $m(E) < \infty$ and $f_k \rightarrow f$ pointwise. Then for all ε there exists a closed set $A_\varepsilon \subseteq E$ with $m(E \setminus A_\varepsilon) < \varepsilon$ and such that $f_k \rightarrow f$ uniformly on A_ε .*

Proof. For given n, k , set

$$E_{k,n} = \left\{ x \mid |f_j(x) - f(x)| < \frac{1}{n} \text{ for all } j > k \right\} = \left\{ x \mid \sup_{j>k} |f_j(x) - f(x)| < \frac{1}{n} \right\}.$$

If we fix n , then $E_k^n \subseteq E_{k+1}^n$ and $E_k^n \nearrow E$. Therefore there exists k_n such that $m(E \setminus E_{k_n}^n) < \frac{1}{2^n}$. Now choose N such that $2^{1-N} < \frac{\varepsilon}{2}$ and set $\tilde{A}_\varepsilon = \bigcap_{n \geq N} E_{k_n}^n$. Then

$$m(E \setminus \tilde{A}_\varepsilon) = m\left(E \setminus \bigcap_{n \geq N} E_{k_n}^n\right) = m\left(\bigcup_{n \geq N} E \setminus E_{k_n}^n\right) \leq \sum_{n \geq N} m(E \setminus E_{k_n}^n) < \frac{\varepsilon}{2}.$$

Choose $A_\varepsilon \subseteq \tilde{A}_\varepsilon$ such that $m(\tilde{A}_\varepsilon \setminus A_\varepsilon) < \frac{\varepsilon}{2}$. Then $m(E \setminus A_\varepsilon) < \varepsilon$. To prove uniform convergence on A_ε , choose M such that $M \geq N$ and $\frac{1}{M} \leq \delta$. Then if $x \in A_\varepsilon$, $x \in E_{k_n}^M$, so

$$\sup_{j < k_n} |f_j(x) - f(x)| < \frac{1}{M} < \delta. \quad \square$$

Theorem 1.28 (Lusin). *Let f be measurable and finite-valued on E and $m(E) < \infty$. Then for all ε , there exists $F_\varepsilon \subseteq E$ closed such that $m(E \setminus F_\varepsilon) < \varepsilon$ and $f|_{F_\varepsilon}$ is continuous.*

Proof. First, there exists a sequence of step functions f_n such that $f_n \rightarrow f$ pointwise. Then remove thin boundaries of rectangles to find E_n such that $m(E_n) < \frac{1}{2^n}$ and such that f_n is continuous off of E_n . By Egorov, there exists $A_{\varepsilon/3}$ closed such that $f_n \rightarrow f$ uniformly on $A_{\varepsilon/3}$ and $m(E \setminus A_{\varepsilon/3}) < \varepsilon/3$. Next, let $F' = A_{\varepsilon/3} \setminus \bigcup_{n \geq N} E_n$, where $2^{1-N} < \frac{\varepsilon}{3}$. Now if $m \geq N$, f_n is continuous on F' and $f_n \rightarrow f$ uniformly, so f is continuous on F' . Finally, approximate F' by a closed set $F_\varepsilon \subseteq F'$ such that $m(F' \setminus F_\varepsilon) < \frac{\varepsilon}{3}$. \square

Finally, we will construct a non-measurable function. First, we will define the *Cantor function* $f_c : [0, 1] \rightarrow [0, 1]$ which is given by

$$f_c(x) = \begin{cases} \left(0, \frac{d_1}{2}, \frac{d_2}{2}, \dots\right)_2 & x = (0.d_1 d_2 \dots)_3 \in C \\ \left(0, \frac{d_1}{2}, \dots, \frac{d_k}{2}, 0, \dots\right)_2 & x = (0d_1 \dots d_k 1 \dots)_3 \notin C \end{cases}.$$

Note that f is monotone and continuous on $[0, 1]$ and constant on C^c .

Then define $f(x) = f_c(x) + x$, so $f : [0, 1] \rightarrow [0, 2]$ is a continuous bijection. On any excluded interval I , f sends I to $I + f_c(I)$. Therefore, $m(f(\bigcup I)) = 1$, so $m(f(C)) = 1$. Then there exists a non-measurable set $S \subseteq f(C)$. Then χ_S is not measurable.

Integration

2.1 DEFINING THE INTEGRAL

Our goal is to integrate measurable functions. If $f = \sum a_k \chi_{E_k}$ is simple, then we define

$$\int f = \sum a_k m(E_k).$$

To check that this is well-defined, if we can write $f = \sum b_l \chi_{F_l}$, then we can consider $E_1 \cap E_2, E_1 \setminus E_2, E_2 \setminus E_1$. This gives a canonical representation of a simple function (where all of the E_j are disjoint). It is easy to see that we have linearity, monotonicity, and the triangle inequality. We want to extend this to a larger class of functions.

Next, we will consider bounded measurable functions on sets of finite measure. If f is supported on a set of finite measure and f is bounded, we define

$$\int f = \lim_{n \rightarrow \infty} \int \varphi_n$$

for some sequence φ_n of simple functions converging to f . We need to show this is independent of the sequence.

Lemma 2.1. *If $\varphi_n \rightarrow f$ pointwise on E with $m(E) < \infty$, and f is bounded, then*

1. *The limit $\lim \int \varphi_n$ exists;*
2. *If $f = 0$ almost everywhere, then $\lim \int \varphi_n = 0$.*

Proof. For the first part, we use Egorov. There exists a closed set $A_\varepsilon \subseteq E$ such that $\varphi_n \rightarrow f$ uniformly on A_ε and $m(E \setminus A_\varepsilon) < \varepsilon$. Let $I_n = \int_E \varphi_n$. Then

$$\begin{aligned} |I_n - I_m| &\leq \int |\varphi_n - \varphi_m| dx \\ &\leq \int_{A_\varepsilon} |\varphi_n - \varphi_m| dx + \int_{E \setminus A_\varepsilon} |\varphi_n - \varphi_m| dx. \end{aligned}$$

On A_ε , we have uniform convergence, so

$$\begin{aligned} \int_{A_\varepsilon} |\varphi_n - \varphi_m| &\leq m(A_\varepsilon) \sup_{A_\varepsilon} |\varphi_n - \varphi_m| \\ &\leq 2m(A_\varepsilon) \sup_{A_\varepsilon} |\varphi_n - f|, \end{aligned}$$

which clearly converges to 0. Then the second terms are $\int_{E \setminus A_\varepsilon} |\varphi_n - \varphi_m| dx \leq 2Mm(E \setminus A_\varepsilon)$, which also converges to 0, where $M = \sup |f|$. Thus I_n is Cauchy, so it converges.

If $f = 0$ almost everywhere, then

$$|I_n| \leq \int_{A_\varepsilon} |\varphi_n| + \int_{E \setminus A_\varepsilon} |\varphi_n| \leq m(A_\varepsilon) \cdot \varepsilon,$$

so $\lim I_n = 0$. □

By Lemma 2.1, integration is well-defined. Again, we have the desired properties. Before we extend the definition of the integral to all functions, we will prove a convergence theorem. This is the first exchange of limits result.

Theorem 2.2 (Bounded Convergence Theorem). *If f_n is a sequence of functions, uniformly bounded by M and uniformly supported on E of finite measure, and it $f_n \rightarrow f$ pointwise almost everywhere, then f is measurable and $\int f = \lim \int f_n$. In fact, $\int |f - f_n| \rightarrow 0$ as $n \rightarrow \infty$.*

Proof. The proof of this is the same as proof of the consistency of the definition of the integral. □

Corollary 2.3. *If $f \geq 0$ is bounded and supported on E with $m(E) < \infty$, and if $\int f = 0$, then $f = 0$ almost everywhere.*

Proof. Set $f_k(x) = f(x)\chi_{\{f(x) \geq 1/k\}}$. Then f_k is bounded, supported on E , and $f_k \leq f$. Then

$$0 = \int f(x) \geq \int f_k(x) \geq \frac{1}{k} m(\{f(x) \geq 1/k\}),$$

so $m(\{f(x) > 0\}) = 0$. □

Theorem 2.4. *If f is Riemann-integrable on $[a, b]$, then it is Lebesgue integrable, and the integrals coincide.*

Proof. Translate the condition about partitions into step functions $\varphi_i \leq f \leq \psi_i$, where $\varphi_i \nearrow f$ and $\psi_i \searrow f$. Clearly the ψ_i, φ_i are uniformly bounded. Therefore, by bounded convergence, $\tilde{\varphi}, \tilde{\psi}$ are both measurable and bounded, and $\int \tilde{\varphi} = \int \tilde{\psi}$, so $\tilde{\psi} = \tilde{\varphi}$ almost everywhere. Then clearly $\tilde{\varphi} = f$ almost everywhere, so f is measurable, and by monotonicity, $\int^R f = \int^L f$. □

Finally, we can define integration for all nonnegative functions. We define

$$\int f = \sup \left\{ \int g \mid 0 \leq g \leq f, m\{\text{supp}(g)\} < \infty, g \text{ bounded} \right\}.$$

For any measurable E , define $\int_E f = \int f \cdot \chi_E$. Note that it is easy to check that the definition makes sense. If $\int f < \infty$, we say that f is *integrable*.

Properties of the Integral.

1. Linearity;
2. Monotonicity;
3. Additivity on disjoint sets;
4. If g is integrable and $0 \leq f \leq g$, then f is integrable;
5. If f is integrable, then $f(x) < \infty$ almost everywhere;
6. If $\int f = 0$, then $f = 0$ almost everywhere.

Note that if $f_n \rightarrow f$ almost everywhere, then $\int f_n$ does not necessarily converge to $\int f$. For example, if we take $f_n = n\chi_{[0, 1/n]}$, then $\int f_n = 1$, but $f = 0$.

2.2 SOME CONVERGENCE RESULTS

We will state some convergence results where we can swap limits and integrals. We will also give the general definition of the integral here.

Lemma 2.5 (Fatou). *If f_n is a sequence of functions with $f_n \geq 0$ and if $f(x) = \liminf f_n(x)$, then*

$$\int f \leq \liminf \int f_n.$$

Proof. Suppose g bounded satisfies $m(\text{supp } g) < \infty$ and $0 \leq g < f$. Then set $g_n = \min\{g, f_n\}$. Then $\int g$ is bounded, so by bounded convergence, $\int g_n \rightarrow \int g$. Also, $\int g_n \leq \int f_n$, so $\int g \leq \liminf \int f_n$. Because $\int f = \sup \int g$, we get the desired result. \square

Corollary 2.6. *If $f_n \leq f$ with $f_n(x) \rightarrow f(x)$ almost everywhere, then $\lim \int f_n = \int f$.*

Proof. By monotonicity, $\limsup \int f_n \leq \int f \leq \liminf \int f_n$. \square

Corollary 2.7 (Monotone Convergence). *If $f_n(x) \nearrow f(x)$ almost everywhere, then $\lim \int f_n = \int f$.*

Corollary 2.8. *Given $a_k(x) \geq 0$ measurable, then $\int \sum a_k(x) \leq \sum \int a_k(x)$. Moreover, if $\sum \int a_k(x)$ converges, $\sum a_k(x)$ converges almost everywhere.*

Proof. Let f_n be the partial sums. Then by monotone convergence, if $\sum \int a_k$ converges, $\int \sum a_k$ also converges, so it is integrable. Thus $\sum a_k$ converges almost everywhere. \square

Example 2.9 (Borel-Cantelli). Given E_k measurable with $\sum m(E_k) < \infty$, then $m(\overline{\lim} E_k) = 0$. To see this, set $a_k = \chi_{E_k}$ and use the previous result.

Example 2.10. Let

$$f(x) = \begin{cases} \frac{1}{|x|^{d+1}} & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

Then $\int_{|x|=\varepsilon} f(x) dx \leq \frac{c}{\varepsilon}$ for all $\varepsilon > 0$.

To see this, we will bound $\int_{|x| \geq \varepsilon} f$. Consider the annulus $A_k = \{x \mid 2^k \varepsilon \leq |x| < 2^{k+1} \varepsilon\}$. Then set

$$a_k = \chi_{A_k} \sup_{x \in A_k} f(x) = \chi_{A_k} \frac{1}{(2^k \varepsilon)^{d+1}}.$$

Then set $g(x) = \sum a_k(x)$. Clearly $f \leq g$, so $\int f$ is finite if $\int g$ is. However, we see that

$$\begin{aligned} \int g &= \int \sum a_k \\ &= \sum \int a_k \\ &= \sum \frac{1}{(2^k \varepsilon)^{d+1}} m(A_k) \\ &= \sum \frac{1}{(2^k \varepsilon)^{d+1}} m(A) \\ &= \sum \frac{1}{2^{k(d+1)}} m(A) \\ &= \frac{m(A)}{\varepsilon^{d+1}}. \end{aligned}$$

We now give the general definition of the integral. Let f be a measurable function and write $f(x) = f_1(x) - f_2(x)$, where both f_1, f_2 are nonnegative. Then we want $\int f = \int f_1 - \int f_2$. Clearly this is defined if at most one of the two integrals on the RHS is infinite.

Definition 2.11. f is integrable if both $\int f_1, \int f_2$ are finite.

Theorem 2.12. The integrable is linear, additive, monotone, and satisfies the triangle inequality.

Lemma 2.13. Let f be integrable. Then

1. Given $\varepsilon > 0$, there exists a ball B of finite measure such that $\int_{B^c} |f| < \varepsilon$;
2. For all $\varepsilon > 0$, there exists δ such that whenever $m(E) < \delta$, then $\int_E |f| < \varepsilon$.

Proof. 1. We assume $f \geq 0$. Then for each N consider $B_N = \{|x| \leq N\}$ and $f_n = \chi_{B_n}$. Then we see that $f_n \nearrow f$, so by monotone convergence, $\int f_n \rightarrow \int f < \infty$. Therefore, there exists N such that

$$\int B_N^c f = \int f(1 - \chi_{B_N}) = \int f - \int f_N < \varepsilon.$$

2. Again assume $f \geq 0$. Then set $E_n = \{f(x) \leq n\}$ and $f_n = f \cdot \chi_{E_n}$. Again, $f_n \nearrow f$ and f is integrable, so there exists N such that $f - f_n < \frac{\varepsilon}{2}$ for all $n \geq N$. Then choose δ such that $N\delta < \frac{\varepsilon}{2}$. Now if $m(E) < \delta$, we see that

$$\int_E f = \int_E f - f_n + \int_E f_n = \int_E f - f_n + N \cdot m(E) < \varepsilon. \quad \square$$

Theorem 2.14 (Dominated Convergence). If $f_n \rightarrow f$ almost everywhere and $|f_n| \leq |g|$ almost everywhere for some integrable g , then f is integrable and $\int |f_n - f| \rightarrow 0$ as $n \rightarrow \infty$.

Proof. For $N \geq 0$, set $E_N = \{|x| |g(x)| \leq N\}$. Then define $g_n = g \cdot \chi_{E_n}$. As in the lemma above, there exists N such that $\int E_N^c g < \varepsilon$.

Note that $f_k \chi_{E_n} \leq n$ for all k and is supported on E_n , which has finite measure. Also, $f_k \chi_{E_n} \rightarrow f \chi_{E_n}$. By bounded convergence,

$$\int |f_k \chi_{E_n} - f \chi_{E_n}| = \int_{E_n} |f_k - f| < \varepsilon$$

for k large enough, so

$$\int |f_k - f| = \int_{E_n} |f_k - f| + \int_{E_n^c} |f_k - f| \leq \varepsilon + 3 \int_{E_n^c} g < 4\varepsilon. \quad \square$$

We now briefly consider functions $f : \mathbb{R}^d \rightarrow \mathbb{C}$. Then note that $f = u + iv$, so $|f| = \sqrt{u^2 + v^2}$. Also, because $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, $|f| \leq |u| + |v|$. We conclude that f is integrable if and only if u, v are both integrable.

2.3 VECTOR SPACE OF INTEGRABLE FUNCTIONS

Let L^1 denote the set of integrable functions. Define the L^1 norm $\|f\|_1 = \int |f|$. Note that this is only a seminorm, so we set $f \sim g$ if $f = g$ almost everywhere.

Theorem 2.15. L^1 is a normed vector space.

Theorem 2.16 (Riesz-Fischer). L^1 is a Banach space.¹

¹We will define this in 624, but a Banach space is just a complete normed vector space.

Proof. Suppose $\{f_n\}$ is Cauchy in L^1 . Then for all k choose n_{k+1} such that $\|f_{n_{k+1}} - f_{n_k}\| < \frac{1}{2^k}$. Then we define

$$\begin{aligned} f &= f_{n_1} + \sum_k (f_{n_{k+1}} - f_{n_k}) \\ g &= |f_{n_1}| + \sum_k |f_{n_{k+1}} - f_{n_k}|. \end{aligned}$$

Note that g is nonnegative, so by monotone convergence,

$$\int |g| = \int |f_{n_1}| + \sum \int |f_{n_{k+1}} - f_{n_k}| < \infty,$$

so by dominated convergence, f is integrable. In addition, $\|f - f_{n_{k+1}}\| \rightarrow 0$, so we have a convergent subsequence. But then because the sequence is Cauchy and by the triangle inequality, it converges to f . \square

Corollary 2.17. *If $f_n \rightarrow f$ in L^1 , there exists a subsequence f_{n_k} such that $f_{n_k} \rightarrow f$ almost everywhere.*

Definition 2.18. We say f_n converges to f in measure if $m\{x \mid |f_n(x) - f(x)| > \varepsilon\} \rightarrow 0$.

Example 2.19. Write natural numbers n as $n = k + 2^v$, where $0 \leq k < 2^v$. Set $f_n = \chi_{\left[\frac{k}{2^v}, \frac{k+1}{2^v}\right]}$. Then this converges to 0 in measure but not pointwise anywhere.

Theorem 2.20. *If $f_n \rightarrow f$ in measure, there exists a subsequence $f_{n_k} \rightarrow f$ almost everywhere.*

Proof. Given v , there exists n_v such that $m\{x \mid |f_n(x) - f(x)| \geq 2^{-v}\} \leq 2^{-v}$ for $n \geq n_v$. Then choose

$$E_v = \{x \mid |f_{n_v}(x) - f(x)| \geq 2^{-v}\}.$$

Now if $x \notin E_v$, then $f_{n_v}(x) \rightarrow f(x)$ for $x \notin \bigcap_{k=1}^{\infty} \bigcup_{v=k}^{\infty} E_v$. However, $\sum m(E_v) < \infty$, so this is a set of measure 0. \square

Remark 2.21. In Fatou's lemma, monotone convergence, and dominated convergence, we can replace pointwise convergence by convergence in measure.

Recall that in a topological space X , a set A is dense if $A \cap \mathcal{O} \neq \emptyset$ for all open \mathcal{O} . In a normed space, this is equivalent to saying that for all $x \in X$ and $\varepsilon > 0$, there exists $a \in A$ such that $\|a - x\| < \varepsilon$.

Theorem 2.22. *The following classes of functions are dense in L^1 :*

1. *Simple functions;*
2. *Step functions;*
3. *Continuous functions of compact support.*

Proof. Take $f \geq 0$.

1. We know that there exists a sequence of increasing simple functions $\varphi_n \nearrow f$ pointwise, but f is integrable, so $\|f - \varphi_n\| = \int f - \varphi_n \rightarrow 0$.
2. We approximate simple functions by step functions. Let $\varphi = \sum a_i \chi_{E_i}$ be simple. Then there exist disjoint closed rectangles R_j such that $m(E_\Delta \cup R_j) < \varepsilon$. This bounds $\int \chi_E - \chi_{R_j}$.
3. It is easy to approximate χ_R by continuous functions. Just build side ramps.

\square

2.4 SYMMETRIES

First, note that the integral is translation invariant. To see this, we can translate step functions. In addition, it is easy to see that $n^d \int f(nx) dx = \int f(x) dx$.

Examples 2.23. First, note that

$$\int_{|x| \geq \varepsilon} \frac{dx}{|x|^a} = \varepsilon^{-a+d} \int_{x \geq 1} \frac{dx}{|x|^a}.$$

Similarly,

$$\int_{|x| \leq \varepsilon} \frac{dx}{|x|^a} = \varepsilon^{-a+d} \int_{|x| \leq 1} \frac{dx}{|x|^a}.$$

Example 2.24 (Convolution). Given integrable f, g , define their convolution $f * g$ by

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

It is easy to see that $f * g = g * f$.

Now write $d = d_1 + d_2$, so $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then for fixed y write $f^y = f(-, y)$. Similarly, for $E \subseteq \mathbb{R}^d$, write $E^y = \{x \mid (x, y) \in E\} \subseteq \mathbb{R}^{d_1}$. Define f_x and E_x similarly. The main issue is that if E is measurable, then E^y, E_x might not be.

Theorem 2.25 (Fubini). Suppose $f(x, y)$ is integrable on $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then

1. For almost every $y \in \mathbb{R}^{d_2}$, f^y is integrable on \mathbb{R}^{d_1} ;
2. The function $\int_{\mathbb{R}^{d_1}} f^y(x) dx$ is integrable on \mathbb{R}^{d_2} ;
3. $\int_{\mathbb{R}^{d_2}} \int_{\mathbb{R}^{d_1}} f^y(x) dx dy = \int_{\mathbb{R}^d} f(x, y) dx dy$.

Proof. Let \mathcal{F} be the set of integrable functions where all three conditions hold. We show that $\mathcal{F} = L^1$. First, we show that $\mathcal{F} \neq \emptyset$.

Given any rectangle $R \subseteq \mathbb{R}^d$, it is easy to see that $\chi_R \in \mathcal{F}$. If we write $R = R_1 \times R_2$, then $\chi_R^y(x) = \chi_{R_1}(x) \cdot \chi_{R_2}(y)$. Then the first condition holds everywhere. For the second condition, note that

$$\int \chi_R^y(x) dx = m(R_1) \cdot \chi_{R_2}(y).$$

For the third condition, note that

$$\int \int \chi_R^y(x) dx dy = m(R_1) \cdot m(R_2) = m(R) = \int \chi_R.$$

Therefore, all step functions are in \mathcal{F} .

Now we will show that

1. \mathcal{F} is closed under linear combinations;

To see this, let $f_k \in \mathcal{F}$ and $a_k \in \mathbb{R}$. Then use linearity of the integral.

2. \mathcal{F} is closed under limits.

We do this for monotone limits. Note that for all $f \in \mathcal{F}$, there exists A of measure 0 such that f^y is not integrable for $y \in A$. If $f_k \in \mathcal{F}$ and $f_k \nearrow f$ and f is integrable, we want to show that $f \in \mathcal{F}$. Let A_k correspond to f_k . Then their union has measure 0, so f^y is integrable almost everywhere.

Now set $g_k(y) = \int f_k^y \nearrow \int f^y = g$. Now each g_k is integrable and $g_k \nearrow g$, so $\int g_k = \int f_k \nearrow \int f$, so $\int g < \infty$ and thus $g < \infty$ almost everywhere. Thus f^y is integrable almost everywhere, so $f \in \mathcal{F}$.

Recall that if f is integrable, there exists a set φ_k of simple functions converging up to f . Each simple is a linear combination of χ_E for some measurable E of finite measure, so we prove $\chi_E \in \mathcal{F}$.

This reduces the problem to measurable sets:

3. G_δ sets;

If E is contained in the boundary of a rectangle, the result is obvious. Next, suppose E is a finite union of almost disjoint rectangles R_j . Then write χ_E as a linear combination of the χ_R and the χ_P , where $P \subseteq \partial R$. Thus $\chi_E \in \mathcal{F}$.

Now if E is open and of finite measure, write $E = \bigcup Q_j$, where the Q_j are almost disjoint cubes. Then write $\chi_E = \sum \chi_{Q_j}$, which is a monotone limit. Finally, if $E = \bigcap G_k$ for G_k open, then χ_E is also a monotone limit.

4. Null sets;

Recall that there exists a G_δ set B of measure 0 such that $E \subseteq B$. Then $\chi_B \in \mathcal{F}$, so $\int \chi_B^y = 0$, so $\int \chi_E^y = 0$ almost everywhere. Because $E^y \subseteq B^y$ for all y , E^y is measurable and null almost everywhere, so $\chi_E \in \mathcal{F}$.

5. Limiting procedure;

For any measurable E , E differs from a G_δ set by a null set.

Finally, for general f , write $f = f_+ - f_-$, which are both integrable and nonnegative. □

Theorem 2.26 (Tonelli). *Fubini extends to measurable functions for which $\int f$ is defined but infinite.*

Corollary 2.27 (Cavaleiri's Principle). *Let E be a measurable subset of $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Then for almost every y , the slice E^y is measurable, and $m(E) = \int m(E^y)$.*

Differentiation

3.1 INTEGRATION AND DIFFERENTIATION

Recall the fundamental theorem of calculus:

1. For f continuous, the function $F(x) = \int_a^x f(x) dx$ is differentiable and $F' = f$.
2. If F' is Riemann-integrable, then $F(b) - F(a) = \int_a^b F'(x) dx$.

This roughly tells us that integration is anti-differentiation. We can replace differential equations with formulations in terms of integrals. Then recall that

$$F'(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} f.$$

What happens when we generalize to higher dimensions and consider $\frac{1}{m(B)} \int_B f$? Define $f^*(x) = \sup \frac{1}{m(B)} \int |f|$. Note that f^* need not be integrable.

Theorem 3.1. 1. f^* is measurable;

2. $f^* < \infty$ almost everywhere;

3. There exists A depending only on d such that for all α ,

$$m(\{x \mid f^*(x) > \alpha\}) \leq \frac{A}{\alpha} \|f\|_1.$$

Proof. 1. Define $E_\alpha = \{x \in \mathbb{R}^d \mid f^*(x) > \alpha\}$. Then note that $x \in E_\alpha$ if and only if

$$\frac{1}{m(B)} \int_B |f| > \alpha$$

for some ball $B \ni x$. In fact, for all $y \in B$, $y \in E_\alpha$, so E_α is open.

2. This is a corollary of part 3.

3. The proof of this relies on the Vitali Covering Lemma, which is stated below.

Let $E_\alpha = \{x \mid f^*(x) > \alpha\}$. Then we know $x \in E_\alpha$ if and only if there exists $B \ni x$ such that $m(B) < \frac{1}{\alpha} \int_B |f|$. Given a compact $K \subseteq E_\alpha$, write $K \subseteq \bigcup_{x \in K} B_x$ and extract a finite subcover. Then we use Vitali, so

$$m(K) < 3^d \sum m(B_i) \leq \frac{3^d}{\alpha} \sum \int_{B_i} |f| \leq \frac{3^d}{\alpha} \int |f|.$$

Because K is arbitrary, we have the desired result. \square

Lemma 3.2 (Vitali's Covering Lemma). *Let $B = \{B_1, \dots, B_N\}$ be a finite set of open balls. Then there exists a set of disjoint open balls $\{B_{i_1}, \dots, B_{i_k}\} \subset B$ such that $\bigcup B_j \subseteq \bigcup 3 \cdot B_{i_k}$.*

Proof. Inductively choose B_{i_k} to be the ball of largest radius not disjoint to thre previous balls and not contained in $\bigcup 3 \cdot B_{i_j}$. We started with a finite collection, so the process terminates. \square

Theorem 3.3 (Lebesgue Differentiation Theorem). *If f is integrable, then for almost every x ,*

$$f(x) = \lim_{\substack{m(B) \rightarrow 0 \\ x \in B}} \frac{1}{m(B)} \int_B f(y) dy.$$

In \mathbb{R} , this implies $\frac{d}{dx} F(x) = f(x)$.

Proof. For $\alpha > 0$, define the set

$$E_\alpha = \left\{ x \mid \limsup_{\substack{m(B) \rightarrow 0 \\ x \in B}} \left| \frac{1}{m(B)} \int f(y) dy - f(x) \right| > 2\alpha \right\}.$$

We will show that $m(E_\alpha) = 0$.

Let $\varepsilon > 0$. Recall that continuous functions of compact support are dense in L^1 , so choose a continuous function with compact support g such that $\|f - g\| < \varepsilon$. Then g is uniformly continuous, so $\frac{1}{m(B)} \int g(y) dy \rightarrow g(x)$ as $m(B) \rightarrow 0$. Now use the triangle inequality:

$$\frac{1}{m(B)} \int f - f(x) = \frac{1}{m(B)} \int f(y) - g(y) dy + \frac{1}{m(B)} \int g(y) dy - g(x) + (g(x) - f(x)).$$

Therefore, we have

$$\limsup_{\substack{m(B) \rightarrow 0 \\ x \in B}} \left| \frac{1}{m(B)} \int_B f(y) dy - f(x) \right| \leq (f - g)^*(x) + |f(x) - g(x)|.$$

Let F_α be the set where $(f - g)^*(x) > \alpha$ and G_α where $|f - g| > \alpha$. By Theorem 3.1, $m(F_\alpha) \leq \frac{A}{\alpha} \|f - g\|$, and by Chebyshev, $m(G_\alpha) \leq \frac{1}{\alpha} \|f - g\|$. \square

Corollary 3.4. *For almost every x , $f^*(x) \geq |f(x)|$.*

Also note that we can replace integrability by local integrability in these results. If E is measurable, define x to be a *point of density* of E if

$$\lim_{\substack{m(B) \rightarrow 0 \\ x \in B}} \frac{1}{m(B)} \int_B \chi_E(y) dy = 1.$$

Corollary 3.5. *Let E be measurable. Then*

1. *Almost every $x \in E$ is a point of density;*
2. *Almost every $x \notin E$ is not a point of density.*

In a related notion, we can define the *Lebesgue set* $\mathcal{L}(f)$ of f as the set of all x such that $f(x) < \infty$ and

$$\lim_{m(B) \rightarrow 0} \frac{1}{m(B)} \int_B |f(y) - f(x)| dy = 0.$$

Clearly if $x \in \mathcal{L}(f)$, then the Lebesgue differentiation theorem holds. Also, if f is continuous at x , then $x \in \mathcal{L}(f)$.

Corollary 3.6. *For any $f \in L^1_{\text{loc}}$, almost every x belongs to $\mathcal{L}(f)$.*

Proof. For rationals $r \in \mathbb{Q}$, consider $h_r = |f(x) - r|$. Then there exists E_r with $m(E_r) = 0$ such that the Lebesgue differentiation theorem holds for h_r for all $x \notin E_r$. Thus if $x \notin E = \bigcup E_r$, then there exists r such that $|f(x) - r| < \varepsilon$ for all ε . For this ε, r , choose δ such that if $m(B) < \delta$, then

$$\left| \frac{1}{m(B)} \int_B |f(y) - r| - |f(x) - r| \right| < \varepsilon,$$

so $\frac{1}{m(B)} \int_B |f(y) - r| < |f(x) - r| + \varepsilon$. Therefore

$$\frac{1}{m(B)} \int_B |f(y) - f(x)| \leq \frac{1}{m(B)} \int_B |f(y) - r| + |f(x) - r| < 2|f(x) - r| + \varepsilon < 3\varepsilon. \quad \square$$

Note that any locally integrable function has different representatives up to sets of measure 0. We know that for almost every x , $f(x) = \lim_{m(B) \rightarrow 0} \frac{1}{m(B)} \int_B f$ and if $f = \bar{f}$ almost everywhere, $\int_B f = \int_B \bar{f}$, but $\mathcal{L}(f) \neq \mathcal{L}(\bar{f})$ up to a set of measure 0.

On the Lebesgue set, we can generalize average. Instead of balls, we can consider general sets U . Note that $\frac{1}{m(U)} \int_U f = \int \frac{1}{m(U)} \chi_U f$, which is morally equal to $\int K f$, where the kernel K satisfies $\int K = 1$. We will actually consider families of such kernels K_δ .

3.2 KERNELS

Consider sets U for which $\frac{1}{m(U)} \int_U f$ converges.

Definition 3.7. A family of sets U_α *shrinks regularly* to x if there exists $c > 0$ such that for all $B \ni x$, there exists α such that $x \in U_\alpha \subseteq B$, and $m(U_\alpha) \geq c \cdot m(B)$.

Corollary 3.8. *If f is locally integrable and U_α shrinks regularly to $x \in \mathcal{L}(f)$, then*

$$\lim_{m(U_\alpha) \rightarrow 0} \frac{1}{m(U_\alpha)} \int_{U_\alpha} f = f(x).$$

Proof. Note that

$$\frac{1}{m(U_\alpha)} \int_{U_\alpha} |f(y) - f(x)| dy \leq \frac{1}{m(U_\alpha)} \int_B |f(y) - f(x)| dy \leq \frac{1}{c} \frac{1}{m(B)} \int_B |f(y) - f(x)| dy \rightarrow 0. \quad \square$$

Note that E is measurable if and only if for all A , $m_*(A) = m_*(A \cap E) + m_*(A \cap E)^c$. Also, note that $\frac{1}{m(B)} \int_B f = \frac{1}{m(B)} \int_{\mathbb{R}^d} \chi_B \cdot f = \int f \cdot \frac{1}{m(B)} \chi_B$. Instead of multiplication, we will convolve. We will show later that $*$ commutes with differentiation.

Note that $f * \chi$ is a restriction of f . Far from x , $f * \chi = 0$ and $f * \chi = f$ near x . However, χ_E is not usually differentiable. Instead, we will convolve with different kernels.

Definition 3.9. A good kernel K_δ satisfies

1. $\int K_\delta = 1$ for all $\delta > 0$.
2. $\int |K_\delta| \leq A$ uniformly. Also, for all $\xi > 0$,

$$\int_{|x| \geq \xi} |K_\delta(x)| dx \rightarrow 0$$

as $\delta \rightarrow 0$.

Example 3.10. We will construct a smooth bump function. Let

$$f_1(x) = \begin{cases} e^{-1/x} & 0 < x < 1 \\ 0 & x \leq 0 \\ \frac{1}{e} & x \geq 1 \end{cases}.$$

Now set $f_2(x) = f_1(x) \cdot f_1(1-x)$. Then set $\varphi(x) = \frac{f_2(x+1/2)}{\int f_2}$. This is a smooth even function with integral 1 supported on $[-1/2, 1/2]$. Now we will introduce scaling, to define a family $\phi_\varepsilon : \mathbb{R}^d \rightarrow \mathbb{R}$ defined by

$$\phi_\varepsilon(x) = K \cdot \varphi\left(\frac{|x|}{\varepsilon}\right),$$

where K is chosen such that $\int \varphi_\varepsilon(x) = 1$. If we define $f_\varepsilon = f * \phi_\varepsilon$, then f_ε is a smooth approximation to f that is C^∞ . In addition, the derivative of f_ε is $D\phi_\varepsilon * f$.

The bottom line is that for all $\varepsilon > 0$, $\phi_\varepsilon * f$ is C^∞ . Also, $(\phi_\varepsilon * f)(x) \rightarrow f(x)$ as $\varepsilon \rightarrow 0$ for all Lebesgue points of f .

We call a kernel K_δ an *approximation to the identity* if it satisfies

1. $\int K_\delta = 1$;
2. $|K_\delta(x)| \leq A\delta^{-d}$ for all $\delta > 0$;
3. $|K_\delta(x)| \leq \frac{A\delta}{|x|^{d+1}}$ for all $\delta > 0$, $x \in \mathbb{R}^d$, $x \neq 0$.

Therefore, $K_\delta \rightarrow \infty$ as $\delta \rightarrow 0$ for $x = 0$. If $x \neq 0$, $K_\delta(x) \rightarrow 0$ as $\delta \rightarrow 0$. Note that the *Dirac mass*¹ is a “generalized function” which satisfies $\int \delta = 1$, $\delta(x) = 0$ for $x \neq 0$, and $\int \delta \cdot f(x) = f(0)$.

Example 3.11 (Poisson Kernel). Consider the upper half plane (or equivalently the disk). Then define

$$P_y = \frac{1}{\pi} \frac{y}{x^2 + y^2}$$

¹In 624, we will see that the Dirac mass is a *distribution*.

for $x \in \mathbb{R}, y > 0$. Then for $x \neq 0, P_y \rightarrow 0$ as $y \rightarrow 0$. On the disk, define

$$P_r(x) = \begin{cases} \frac{1}{2\pi} \cdot \frac{1-r^2}{1-2r\cos x+r^2} & |x| \leq \pi \\ 0 & |x| > \pi \end{cases}.$$

Note that $\Delta P = 0$ away from the origin. If we convolve $u = P * g$, u solves $\Delta u = 0$ on $y > 0$ and $u(x, 0) = g(x)$.

Example 3.12 (Heat Kernel). Define the heat kernel

$$H_t = \frac{1}{(4\pi t)^{d/2}} e^{-|x|^2/4t}.$$

Again, if $x \neq 0, H_t(x) \rightarrow 0$ as $t \rightarrow 0$. If $x = 0$, then $H_t(0) \rightarrow \infty$. Also, for $t > 0, \int_{\mathbb{R}^d} H_t(x) dx = 1$. Note that for $t > 0$, the heat equation $u_t = \Delta u$ with initial conditions $u(0, x) = g(x)$ is solved by $u(t, x) = H_t * g$.

Because the heat equation is linear, since $\delta = \lim H_t$ is a convolution identity, $*$ commutes with ∂_x , so given f , set $u = H_t * f$. Because differentiation commutes with $*$, we see that

$$\partial_t(H_t * f) - \Delta(H_t * f) = (\partial_t H_t - \Delta H_t) * f,$$

but at $t = 0$, then $\delta * f = f$.

Example 3.13 (Fejer Kernel). This kernel is defined by

$$\frac{1}{2\pi} F_N(x) = \begin{cases} \frac{1}{2\pi N} \frac{\sin^2 \frac{Nx}{2}}{\sin^2 \frac{x}{2}} & |x| \leq \pi \\ 0 & |x| > \pi \end{cases}.$$

It is related to the Fourier transform. Recall that the Fourier transform \tilde{f} is defined by

$$\tilde{f}(\xi) = \int f(x) e^{-2\pi i x \xi} dx$$

and its inverse

$$f(x) = \int \tilde{f}(\xi) e^{2\pi i x \xi} d\xi.$$

Note that $f(x) e^{2\pi i x \xi}$ as a function of ξ is a wave.

Theorem 3.14. If f is measurable and $x \in \mathcal{L}(f)$, then $(f * K_\delta)(x) \rightarrow f(x)$ as $\delta \rightarrow 0$.

Proof. We calculate

$$\begin{aligned} (f * K_\delta)(x) - f(x) &= \int f(x-y) K_\delta(y) dy - f(x) \\ &= \int (f(x-y) - f(x)) K_\delta(y) dy. \end{aligned}$$

We will show that $\int |f(x-y) - f(x)| K_\delta(y) dy \rightarrow 0$. We will use properties to K_δ and split the integral into parts $|y| \leq \delta$ and $|y| > \delta$. Define for $x \in \mathcal{L}(f)$,

$$A(r) = \frac{1}{W_d r^d} \int_{|y| \leq r} |f(x-y) - f(x)| dy,$$

where $W_d = m(B_0(1))$. We then need Lemma 3.15, which is stated after this proof. This yields boundedness on $(0, R]$. Then for large e , we use the triangle inequality:

$$A(r) \leq \frac{1}{W_d r^d} \|f\| + \frac{1}{W_d r^d} \int_{|y| \leq r} |f(y)| \leq \frac{1}{W_d r^d} \|f\| + |f(x)|.$$

Turning back to our original problem, note that

$$\int |f(x-y) - f(x)| |K_\delta(y)| dy \leq \int_{|y| \leq \delta} |f(x-y) - f(x)| |K_\delta(y)| dy + \sum_{k=0}^{\infty} \int_{2^k \delta < |y| \leq 2^{k+1} \delta} |f(x-y) - f(x)| |K_\delta(y)| dy.$$

The first term is given by

$$\begin{aligned} \int_{|y| \leq \delta} |f(x-y) - f(x)| |K_\delta(y)| dy &\leq \frac{c}{\delta^d} \int_{|y| \leq \delta} |f(x) - f(y)| dy \\ &\leq cA(\delta) \rightarrow 0. \end{aligned}$$

On the annulus, we will estimate $|K_\delta| \leq \frac{C\delta}{|x|^{d+1}}$. Then we can write

$$\begin{aligned} \int_{2^k \delta < |y| \leq 2^{k+1} \delta} |f(x-y) - f(x)| |K_\delta(y)| dy &\leq \frac{c\delta}{(2^k \delta)^{d+1}} \int_{|y| \leq 2^{k+1} \delta} |f(x-y) - f(x)| dy \\ &\leq (2^{k+1} \delta)^d A(2^{k+1} \delta) W_d \frac{c\delta}{(2^k \delta)^{d+1}} \\ &\leq A(2^{k+1} \delta) W_d c 2^{d-k} \\ &\leq A(2^{k+1} \delta) (W_d c 2^d) 2^{-k}, \end{aligned}$$

so

$$\int |f(x-y) - f(x)| |K_\delta(y)| dy \leq cW_d \left(A(\delta) + \sum_{k=0}^{\infty} 2^d 2^{-k} A(2^{k+1} \delta) \right).$$

For large k , $A(2^{k+1} \delta) \leq M$, so for any N , we see that

$$\int |f(x-y) - f(x)| |K_\delta(y)| dy \leq cW_d \left(A(\delta) + \sum_{k=1}^N 2^d 2^{-k} A(2^{k+1} \delta) + 2^d M 2^{-N} \right).$$

Therefore, given ε , choose N such that $cW_d 2^d M 2^{-N} < \frac{\varepsilon}{2}$. Then for all sufficiently small δ ,

$$cW_d \left(A(\delta) + \sum_{k=1}^N 2^d 2^{-k} A(2^{k+1} \delta) \right) < \frac{\varepsilon}{2}. \quad \square$$

Lemma 3.15. *If $f \in L^1$ and $x \in \mathcal{L}(f)$, then $A(r)$ is continuous on $(0, \infty)$ and $A(r) \rightarrow 0$ as $r \rightarrow 0$.*

Proof. Because the integral is absolutely continuous, $A(r)$ is continuous: absolute continuity of the integral means that for all $\varepsilon > 0$, there exists $\delta > 0$ such that if $m(U) < \delta$, then $\int_U |f| < \varepsilon$.

Now consider small annuli. For $r > 0$ fixed, the annuli have measure $W_d((r + \delta)^d - (r - \delta)^d)$, which is small. Then $A(r) \rightarrow 0$ is just the statement that $x \in \mathcal{L}(f)$. \square

3.3 DIFFERENTIATION

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be monotone increasing. Given only two points x_L, x_R , for any increasing sequence $x_L \leq \dots < x_{-1} < x_0 < x_1 < \dots < x_R$, we have

$$\sum_{k=-N}^M f(x_{k+1}) - f(x_k) = f(x_M) - f(x_N) \leq f(x_R) - f(x_L).$$

In particular, this series converges absolutely, so there are only countably many discontinuities, and moreover $\sum_j f(x_j, +) - f(x_j, -) < \infty$, where $f(x_j, +) = \lim_{x \rightarrow x_j^+} f(x)$ and $f(x_j, -)$ is defined similarly.

We know that $f(x, -) \leq f(x) \leq f(x, +)$ for all x and that the three values are equal if and only if f is continuous at x . Then $\sum f(x_k, +) - f(x_k, -) < \infty$ means that there are countably many jumps and that their sum is finite. Now assume that f is continuous and increasing on $[a, b]$.

Note that the Cantor function is continuous, monotone, and takes $f : [0, 1] \rightarrow [0, 1]$. Recall that f_c is constant on each excluded interval, so $f'_c = 0$ almost everywhere. Therefore, $\int_0^1 f'_c(x) dx = 0$, but $f_c(1) - f_c(0) = 1$.

Theorem 3.16. *If f is continuous and monotone on $[a, b]$, then the derivative exists almost everywhere and*

$$f(b) - f(a) \geq \int_a^b f'(x) dx.$$

Before we prove the theorem, we need to define the Dini derivatives. Define

$$\begin{aligned} D^+ f(x) &= \limsup_{h \rightarrow 0^+} \Delta_h f(x) & D^- f(x) &= \limsup_{h \rightarrow 0^-} \Delta_h f(x) \\ D_+ f(x) &= \liminf_{h \rightarrow 0^+} \Delta_h f(x) & D_- f(x) &= \liminf_{h \rightarrow 0^-} \Delta_h f(x). \end{aligned}$$

Then the Dini numbers are defined everywhere and that $D_- \leq D^-$, $D_+ \leq D^+$. Then f is differentiable at x if and only if $D^+ \leq D_-$ and $D^- \leq D_+$ and all these values are finite.

Lemma 3.17 (Rising Sun Lemma). *Let G be continuous on $[a, b]$. Then the set $\{x \mid \exists h > 0 (G(x+h) > G(x))\}$ is a relatively open union of disjoint intervals $\bigcup_i (a_i, b_i)$, and for all $x \in (a_i, b_i)$, $G(a_i) = G(b_i) > G(x)$.*

Proof. It is clear that the set is open. Now if $G(b_i) > G(a_i)$, then we can extend the interval to the left, and similarly to the right if $G(a_i) > G(b_i)$. \square

Lemma 3.18 (Vitali's Covering Lemma). *A collection \mathcal{G} of open/closed intervals is a Vitali cover of E if for all $\varepsilon > 0$ and $x \in E$, there exists $G \in \mathcal{G}$ such that $x \in G$ and $|G| < \varepsilon$. Let E be of finite measure and \mathcal{G} a Vitali cover. Then for all $\varepsilon > 0$, there exists a finite set of disjoint $G_n \in \mathcal{G}$ such that $m(E) \leq \sum |G_j| + \varepsilon$.*

Proof. Assume without loss of generality that each $G \in \mathcal{G}$ is closed. Fine \mathcal{O} of finite measure with $E \subseteq \mathcal{O}$. Now choose a series of intervals as follows. Suppose that G_1, \dots, G_n have already been chosen. Then let

$$k_n = \sup_{G \subseteq \mathcal{O}, G \in \mathcal{G}, G \cap (\bigcup_{j=1}^n G_j) = \emptyset} |G|.$$

Either $E \subseteq \bigcup G_i$ or there exists $G_{n+1} \in \mathcal{G}$ contained in \mathcal{O} and disjoint from the other sets such that $|G_{n+1}| > \frac{1}{2} k_n > 0$. Continue the process. If this does not stop, then choose N such that $\sum_{n=N+1}^{\infty} |G_j| < \frac{\varepsilon}{5}$. Let $R = E \setminus \bigcup_{j=1}^N G_j$.

To show that $m(R) < \varepsilon$, choose $x \in R$. Because the G_j are closed, there exists $G \in \mathcal{G}$ such that $x \in G$ and $|G| < \frac{1}{2} d(x, \bigcup_{j=1}^N G_j)$. Thus G does not intersect any of the G_j . From our construction, there exists $n > N$

such that $|G| \leq k_n \leq 2|G_{n+1}|$. however, $|G_{n+1}|$ tends to 0, so there exists a smallest $n > N$ such that $G \cap G_n$ is nonempty. For this n , if m_n is the midpoint of G_n , we see that $|x - m_n| \leq |G| + \frac{1}{2}|G_n|$. However, $|G| \leq 2|G_n|$, so $|x - m_n| \leq \frac{5}{2}|G_n|$. Therefore $x \in J_n$ for some J_n with $|J_n| = 5|G_n|$, so

$$m(R) \leq \sum |J_j| \leq \sum 5|G_j| < \varepsilon. \quad \square$$

Proof of Theorem 3.16. We want to show that $D^+F \leq D_-F$ almost everywhere and $D^+F < \infty$ almost everywhere. Fix γ and set $E_\gamma = \{x \mid D^+F(x) > \gamma\}$. We will show that E_γ is measurable. Note that $\Delta_h F$ is measurable and f is continuous, so replace h with $1/n$. Set $G(x) = F(x) - \gamma x$.

The idea is that on E_γ , $D^+F(x) > \gamma$. Then there exists h_0 such that $\frac{F(x+h)-F(x)}{h} > \gamma$ for $0 < h < h_0$. Thus $F(x+h) - F(x) > \gamma h$, so $G(x+h) > G(x)$. G is not monotone but is continuous. We will now eliminate $\{D^+F = \infty\}$ and $\{D^+F > D_-F\}$. Note that $D^+F(x) > \gamma$ if and only if $G_\gamma(x+h) > G_\gamma(x)$ for some h . By rising sun, we know that $E_\gamma = \{D^+F > \gamma\}$ is open and of finite measure, so $m(\{D^+F = \infty\}) = 0$.

Now choose $R > r$ rational and set $E = \{x \mid D^+F > R > r > D_-F\}$. We will show that $m(E) = 0$. If not, then there exists an open $\mathcal{O} \supseteq E$ with $m(\mathcal{O}) < m(E)\frac{R}{r}$. On some connected component I_n of \mathcal{O} , apply rising sun to $G(x) = -F(-x) + rx$ on $-I_n$. Then we obtain intervals (c_k, d_k) , so we get $(a_k, b_k) = (-d_k, -c_k)$. Then we see that $-F(b_k) - rb_k = -F(a_k) - ra_k > -F(x) - rx$ for $x \in (a_k, b_k)$.

To show that $m(E) = 0$, find an open $\mathcal{O} \supseteq E$ such that $m(\mathcal{O}) < m(E) + \varepsilon$. Then for all $x \in E$, there exists an interval $[x-h, x] \subseteq \mathcal{O}$ such that $F(x) - F(x-h) < rh$. This is a Vitali covering.

Now choose I_1, \dots, I_N such that $m(E) \leq \sum |I_n| + \varepsilon$. Summing over this finite subcover, we see that

$$\sum_{j=1}^N [F(x_j) - F(x_j - h_j)] < r \sum h_j < rm(\mathcal{O}) < r(m(E) + \varepsilon).$$

Then if we set A to be the interior of $\bigcup I_j$, then for all $y \in A$ there exists δ such that $(y, y + \delta) \subseteq I_j$ and $F(y + \delta) - F(y) > R\delta$. By Vitali, there exists J_1, \dots, J_m of the form $J_i = [y_i, y_i + \delta_i]$ such that $m(\bigcup J_i) \geq m(A) - \varepsilon \geq m(E) - 2\varepsilon$. For each J_i , we have $F(y_i + \delta_i) - F(y_i) > R\delta_i$, so adding, we get that

$$\sum_{i=1}^m F(y_i + \delta_i) - F(y_i) > R \sum \delta_i > R(m(E) - 2\varepsilon).$$

Each J_i is contained in an I_ℓ , so we can sum over i to see that

$$R(m(E) - 2\varepsilon) < \sum_{i=1}^m F(y_i + \delta_i) - F(y_i) \leq \sum_{j=1}^N F(x_j) - F(x_j - h_j) \leq r(m(E) + \varepsilon).$$

Because ε is arbitrary and $Rm(E) \leq rm(E)$, we see that $m(E) = 0$. \square

We now claim that F' is measurable. In addition, F' is integrable and

$$\int_a^b F'(x) dx \leq F(b) - F(a).$$

To see this, note that F' is an almost everywhere limit of measurable functions. By Fatou, we see that

$$\begin{aligned} \int_a^b F'(x) dx &\leq \liminf_{n \rightarrow \infty} \int_a^b n[F(x + 1/n) - F(x)] dx \\ &= n \int_{a+1/n}^{b+1/n} F(y) dy - n \int_a^b F(x) dx \\ &= n \int_b^{b+1/n} F(y) dy - n \int_a^{a+1/n} F(y) dy. \end{aligned}$$

Because F is continuous,

$$n \int_b^{b+1/n} F(y) - F(b) dy \leq n \cdot \frac{1}{n} \cdot \varepsilon$$

for n large enough, so $n \int_b^{b+1/n} F(y) dy \rightarrow F(b)$. The same argument holds for the integral near a .

3.4 BOUNDED VARIATION

Given $[a, b]$, create a finite partition $a = x_0 < x_1 < \dots < x_n = b$. Then set

$$p = \sum [f(x_i) - f(x_{i-1})]_+ \quad n = \sum [f(x_i) - f(x_{i-1})]_- \quad t = n + p = \sum |f(x_i) - f(x_{i-1})|.$$

Then define the increasing, decreasing, and total variations of f by

$$IV(f) = \sup p \quad DV(f) = \sup n \quad TV(f) = \sup t.$$

We say f is of *Bounded Variation* if $TV(f)$ is finite.

Lemma 3.19. *If f is BV, then $TV(f) = IV(f) + DV(f)$ and $f(b) - f(a) = IV(f) - DV(f)$.*

Proof. Choose a finite partition. It is easy to see that $f(b) - f(a) = p - n$, so for all partitions, $n + f(b) - f(a) = p$. Then $n + f(b) - f(a) \leq IV(f)$, but this means $DV(f) + f(b) - f(a) \leq IV(f)$. Similarly, if we take the supremum over n first, we get the inequality reversed. The first part of the lemma should be easy. \square

Theorem 3.20. *A function f is BV if and only if it is the difference of two bounded monotone increasing functions.*

Proof. If f is BV, set $g(x) = IV_{[a,x]}(f)$ and $h(x) = DV_{[a,x]}(f)$. Because f is BV, g, h are monotone increasing and bounded, and by Lemma 3.19, $f(x) - f(a) = IV_{[a,x]}(f) - DV_{[a,x]}(f) = g(x) - h(x)$. On the other hand, if $f = g - h$ with g, h monotone and bounded, then $TV(f) \leq TV(g) + TV(h)$ is finite. \square

Corollary 3.21. *If f is BV, then f' exists almost everywhere.*

Now we will consider jumps. Assume f is bounded and monotone increasing. At each point, the limits $f(x+) = \lim_{x \rightarrow x^+} f(x)$ and $f(x-)$ (defined analogously) exist. By monotonicity, $f(x+) \geq f(x) \geq f(x-)$. Define the jump set to be the set $J = \{x \mid f(x-) < f(x+)\}$. We know that J is countable, so enumerate it by $J = \{x_1, x_2, \dots\}$. For each n , define $[f]_n = f(x_n+) - f(x_n-)$. At each n , we have $f(x_n) = f(x_n-) + \theta[f]_n$ for some $0 \leq \theta \leq 1$. Now define

$$j_n(x) = \begin{cases} 1 & x > x_n \\ \theta_n & x = x_n \\ 0 & x < x_n \end{cases}$$

and define the jump function of f by $J_f(x) = \sum_n [f]_n j_n(x)$.

Lemma 3.22. 1. J_f is monotone increasing;

2. $f - J_f$ is monotone increasing and continuous;

3. J_f is differentiable almost everywhere and $J'_f(x) = 0$ almost everywhere.

Proof. 1. This is obvious.

2. The series defining J_f converges uniformly and $J_f(x) \leq F(b) - F(a)$, so if $x \notin J$, each $j_n(-)$ is continuous at x , so J_f is continuous at x . If x is a jump, then $J_f(x) = \sum_{x_k < x} [f]_k j_k(x) + \sum_{x_k > x} [f]_k j_k(x) + [f]_n j_n(x)$. In particular, $J_f(x-) = \sum_{x_k < x} [f]_k$, $J_f(x+) = \sum_{x_k > x} [f]_k$, and $J_f(x) = J_f(x-) + \theta_n [f]_n$.

Now $f(x+) - J_f(x+) - (f(x-) - J_f(x-)) = 0$ at $x = x_n$. Also, $f(x) - J_f(x) = 0$ provided that $f - J_f$ is increasing, because it is continuous. If $x < y$, note that

$$\begin{aligned} J_f(y) - J_f(x) &= J_f(y) - J_f(y-) + \sum_{x < x_n < y} [f]_n + J_f(x-) - J_f(x) \\ &\leq f(y) - f(y-) + \sum_{x < x_n < y} [f]_n + f(x+) - f(x-) \\ &\leq f(y) - f(x), \end{aligned}$$

so we see that $f - J_f$ is increasing.

3. Note that J_f is discontinuous on a countable set. Given any $\varepsilon > 0$, there exists a closed F such that $J \subseteq F^\circ$ and $m(F) < \varepsilon$. Then F^c is open and J_f is constant on each component of F^c . Therefore, the set where J'_f does not exist or is nonzero has measure less than ε for all ε . \square

Morally, we say that $J'_f(x) = \sum [f]_n H'_n$, where H is the Heaviside step function, and $H' = \delta$.

We now consider the question of which continuous monotone functions satisfy the Fundamental Theorem of Calculus. We consider properties of $\int_a^x f(t) dt$, where f is integrable on $[a, b]$. We will refer to the integral as a function $F(x)$.

Lemma 3.23. *F is continuous and of bounded variation.*

Proof. Continuity is easy. Use the fact that for all ε , there exists δ such that if $m(E) < \delta$, then $\int_E |f| < \varepsilon$. To show that F is BV, choose $a = x_0 < x_1 < \dots < x_n = b$. Then we see that

$$\sum_i |F(x_i) - F(x_{i-1})| = \sum_i \left| \int_{x_{i-1}}^{x_i} f(t) dt \right| \leq \sum_i \int_{x_{i-1}}^{x_i} |f(t)| dt = \int_a^b |f(t)| dt,$$

which is finite. \square

Lemma 3.24. *If f is integrable and $\int_a^x f(t) dt = 0$ for all x , then $f(x) = 0$ almost everywhere.*

Proof. Set $E = E_+ \cup E_-$, where $E_+ = \{f > 0\}$ and $E_- = \{f < 0\}$. If $m(E_+) > 0$, then there exists closed $F \subseteq E_+$ with $m(F) > 0$. Then $G = F^c$ is relatively open. Either we have $\int_a^b f(t) dt \neq 0$ or $0 = \int_a^b f(t) dt = \int_G f(t) dt + \int_F f(t) dt$. Then $\int_G f \neq 0$, so write $G = \bigcup (a_i, b_i)$. Therefore, at least one of the $\int_{a_i}^{b_i} f(t) dt$ is nonzero, so either $\int_{a_i}^{a_i} f(t) dt$ or $\int_{a_i}^{b_i} f(t) dt$ is nonzero. \square

Theorem 3.25. *If f is integrable and if $F(x) = F(a) + \int_a^x f(t) dt$, then $F'(x) = f(x)$ almost everywhere.*

Proof. If f is bounded by K , then we know F is BV, so F' exists almost everywhere. Let $f_n(x) = \Delta_{1/n} F(x)$, so

$$f_n(x) = n \int_x^{x+h} f(t) dt \leq K.$$

Also, $f_n \rightarrow F'$ almost everywhere. By bounded convergence, we then have

$$\begin{aligned}
 \int_a^c F'(x) dx &= \int_a^c \lim_{n \rightarrow \infty} f_n(x) dx \\
 &= \lim_{n \rightarrow \infty} \int_a^c f_n(x) dx \\
 &= \lim_{h \rightarrow 0} \frac{1}{h} \int_a^c F(x+h) - F(x) dx \\
 &= \lim_{h \rightarrow 0} \left[\int_c^{c+h} F(x) dx - \int_a^{a+h} F(x) dx \right] \\
 &= F(c) - F(a) \\
 &= \int_a^c f(x) dx,
 \end{aligned}$$

so $F'(x) = f(x)$ almost everywhere.

For the general case, assume that $f > 0$. Then set $f_n(x) = \min\{f(x), n\}$. Then $f - f_n \geq 0$, so $G_n(x) = \int_a^x f(x) - f_n(x) dx$ is monotone increasing. Thus it has nonnegative derivative almost everywhere. By the above, we know that

$$\frac{d}{dx} \int_a^x f_n(t) dt = f_n(x)$$

almost everywhere. Therefore

$$F'(x) = \frac{d}{dx} G_n(x) + \frac{d}{dx} \int_a^x f_n(t) dt \geq f_n(x)$$

almost everywhere. Thus $\int_a^b F'(x) dx \geq \int_a^b f_n(x) dx$, so $\int_a^b f(x) dx \leq \liminf \int_a^b f_n(x) dx \leq \int_a^b F'(x) dx$, so then

$$\int_a^b F'(x) dx \geq F(b) - F(a). \quad \square$$

3.5 ABSOLUTE CONTINUITY

Definition 3.26. A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is *absolutely continuous* if for all $\varepsilon > 0$, there exists $\delta > 0$ such that if x_i, x'_i for a finite set of non overlapping intervals and if $\sum_i |x'_i - x_i| < \delta$, then

$$\sum_i |f(x'_i) - f(x_i)| < \varepsilon.$$

Note that absolute continuity implies uniform continuity. From the previous lemma, then integrability of f means that $\int_a^x f$ is absolutely continuous.

Lemma 3.27. If f is absolutely continuous, then it is also BV.

Proof. Let f be absolutely continuous and choose $\varepsilon = 1$. Then given any partition $x_0 < x_1 < \dots < x_n$, choose $K > 1 + \frac{b-a}{\delta}$. Add in many partition elements so that $[a, b]$ is covered by K sets of subintervals each of length less than δ . Now

$$\sum |f(y_{i+1}) - f(y_i)| \leq K \sum_{\sum |y_{i+1} - y_i| < \delta} |f(y_{i+1}) - f(y_i)| \leq K.$$

Therefore $TV(f) \leq K$. \square

As a consequence, if f is absolutely continuous, then its derivative exists almost everywhere.

Lemma 3.28. *If f is absolutely continuous on $[a, b]$ and if $f' = 0$ almost everywhere, then f is constant.*

Proof. Fix $t \in (a, b)$ and set $E = (a, t) \cap \{f' = 0\}$, so $m(E) = t - a$. Let two parameters be given: ξ an upper bound for f' on E and ε a bound for $m((a, t) \setminus E)$. Use absolute continuity on $m((a, t) \setminus E)$: Choose δ according to the definition of f depending on ε .

If $x \in E$, then $f'(x) = 0$. Thus there exists h such that $[x, x + h] \subseteq [a, t]$ and $|f(x + h) - f(x)| < \xi h$, so $E \subseteq \bigcup_{x \in E} [x, x + h]$. Now use Vitali's covering lemma, so there exists a finite non-overlapping set of intervals and B such that $m(B) < \delta$ and

$$E \subseteq \bigcup_{k=1}^n [x_k, y_k] \cup B.$$

Order the x_k, y_k and set $y_0 = a, x_{n+1} = t$. Now we have

$$\begin{aligned} |f(t) - f(a)| &\leq \sum_{i=1}^n |f(y_i) - f(x_i)| + |f(x_1) - f(a)| + \sum_{i=1}^n |f(x_{i+1}) - f(y_i)| + |f(t) - f(y_n)| \\ &\leq \sum_{i=1}^n |f(x_i + h_i) - f(x_i)| + \varepsilon \\ &\leq \sum_i \xi h_i + \varepsilon \\ &\leq (b - a)\xi + \varepsilon. \end{aligned}$$

□

Theorem 3.29 (Fundamental Theorem of Calculus, Part 2). *A function F has integrable f such that*

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

if and only if F is absolutely continuous.

Proof. Suppose that F is absolutely continuous. Then F is BV, so $F(x) = F_1(x) - F_2(x)$ for monotone increasing F_1, F_2 . Therefore $F' = F'_1 + F'_2$ exists almost everywhere and $F'_1 \cdot F'_2 = 0$. Therefore

$$\int |F'(x)| = \int F'_1(x) + F'_2(x) = F_1(b) + F_2(b) - F_1(a) - F_2(a).$$

Set $G(x) = \int_a^x F'(t) dt$ and $h(x) = F(x) - G(x)$. Then h is also absolutely continuous and $G'(x) = F'(x)$ almost everywhere from the definition of G , so $h' = 0$ almost everywhere. Thus h is constant and must be equal to $F(a)$. □

We will now characterize functions of bounded variation.

1. If F is BV, then $F = IV(F) - DV(F)$.
2. F has a jump part $J_F = \sum_k [F]_k j_k(x)$, where $[F]_k = F(x_k+) - F(x_k-)$.
3. $F_c = F - J_F$ is continuous and the difference of two monotone continuous functions. Thus F'_c exists almost everywhere. If we set

$$F_{ac}(x) = \int_a^x F'_c(t) dt,$$

then F_{ac} is absolutely continuous. Then the function $F_{sc} = F_c - F_{ac}$ is monotone increasing, but $F'_{sc} = 0$ almost everywhere, so it does not satisfy FTC (behaves like the Cantor function).

In conclusion, we can write $F = J_F + F_{ac} + F_{sc}$.

General Measures

Let X be a set and Σ be a σ -algebra of X . Then M is a *measure* on (X, Σ) if M is a nonnegative σ -additive function on Σ .

- Examples 4.1.**
1. The Lebesgue measure as defined previously;
 2. The counting measure;
 3. The discrete measure on $\Sigma = \mathcal{P}(X)$ where each element is assigned a weight;
 4. Let X be an uncountable set, set Σ to be the countable or cocountable sets, and then let $\mu(E) = 0$ when E is countable and $\mu(E) = 1$ otherwise.

Lemma 4.2. *The following properties hold for all measures:*

1. *Monotonicity;*
2. *Countable subadditivity;*
3. *If $E_i \in \Sigma$ and $\mu(E_i) < \infty$, then $\mu(\bigcap E_i) = \lim \mu(E_n)$ if the E_i are nested.*

Proof. 1. Observe that $B = A \cup (B \setminus A)$ and then use additivity and nonnegativity.

2. Let $F_i = E_n \setminus (\bigcup_{i < n} E_i)$. Then $\mu(\bigcup E_k) = \sum \mu(F_k) \leq \sum \mu(E_k)$.

3. Note that $E_1 = E \cup \bigcup_{k=1}^{\infty} E_k \setminus E_{k+1}$, so

$$\begin{aligned} \mu(E_1) &= \mu(E) + \sum_{k=1}^{\infty} \mu(E_k \setminus E_{k+1}) \\ &= \mu(E) + \sum_{k=1}^{\infty} \mu(E_k) - \mu(E_{k+1}) \\ &= \mu(E) + \mu(E_1) - \lim_{n \rightarrow \infty} \mu(E_{n+1}). \end{aligned}$$

□

We say that μ is finite if $\mu(X) < \infty$ and that μ is σ -finite if X is a countable union of sets of finite measure.

Example 4.3. 1. The Lebesgue measure is σ -finite (just take $[a, a + 1]$ for $a \in \mathbb{Z}$).

2. The counting measure is finite if and only if S is finite.

3. Then discrete measure is finite if and only if countably many points have positive weights and the sum of the weights is finite. It is σ -finite if and only if only countably many points have positive weight and each weight is finite.
4. This is a finite measure. In fact, any measure such that $\mu(X) = 1$ is a *probability measure*.

4.1 OUTER MEASURES

An *outer measure* is a countable subadditive function defined on $\mathcal{P}(X)$ such that $m_*(\emptyset) = 0$. Our main question is to extract a measure from an outer measure when we have no recourse to open sets, unlike in the Lebesgue case.

Definition 4.4. E is (Caratheodory)-measurable if for all $A \subseteq X$, $\mu_*(A) = \mu_*(E \cap A) + \mu_*(E^c \cap A)$.

Note that

1. This condition is equivalent to $\mu_*(A) \geq \mu_*(A \cap E) + \mu_*(A \cap E^c)$;
2. The condition holds for the empty set;
3. E is measurable if and only if E^c is.

This tells us that $\Sigma_c = \{E \mid E \text{ is measurable}\}$ is a σ -algebra and that μ_* is a measure on Σ_c .

To prove the last claim, we need to show countable additivity. First, we will do it for finite unions:

$$\begin{aligned} m_*(A) &\geq m_*(A \cap E_1) + m_*(A \cap E_1^c) \\ &\geq m_*(A \cap E_1 \cap E_2) + m_*(A \cap E_1 \cap E_2^c) + m_*(A \cap E_1^c \cap E_2) + m_*(A \cap E_1^c \cap E_2^c) \\ &\geq m_*(A \cap E_1 \cap E_2 \cup A \cap E_1 \cap E_2^c \cup A \cap E_1^c \cap E_2) + m_*(A \cap E_1^c \cap E_2^c). \end{aligned}$$

Also, additivity on disjoint sets is easy. Now we will consider countable disjoint unions. Set $F_n = \bigcup_{j=1}^n E_j$ and $F = \bigcup_{j=1}^\infty E_j$. Then

$$m_*(F_n \cap A) = m_*(E_n \cap (F_n \cap A)) + m_*(E_n^c \cap (F_n \cap A)) = m_*(E_n \cap A) + m_*(E_{n-1} \cap A).$$

Therefore, $m_*(F_n \cap A) = \sum_{j=1}^n m_*(E_j \cap A)$, so $m_*(A) \geq \sum_{j=1}^n m_*(E_j \cap A) + m_*(F_n^c \cap A)$, so if we allow $n \rightarrow \infty$, then

$$\begin{aligned} m_*(A) &= \sum_{i=1}^\infty m_*(E_i \cap A) + m_*(F^c \cap A) \\ &\geq m_*(\bigcup E_j \cap A) + m_*(F^c \cap A) \\ &= m_*(F \cap A) + m_*(F^c \cap A). \end{aligned}$$

Now let X be a metric space. As before, the Borel sets are the smallest σ -algebra containing all open balls. An outer measure is a *metric outer measure* if it has the separation property.

Theorem 4.5. If m_* is a metric outer measure, then the Borel sets are measurable and $m_*|_{\mathcal{B}_X}$ is a measure.

Proof. We show that closed sets are measurable. Assume F is closed and $m_*(A) < \infty$. Then set

$$A_n = \left\{ x \in F^c \cap A \mid d(x, F) \geq \frac{1}{n} \right\}.$$

Then clearly the A_n are increasing. We will show that $\lim m_*(A_n) = m_*(F^c \cap A)$.

Set $B_n = A_{n+1} \cap A_n^c$. Then $d(B_{n+1}, A_n) \geq \frac{1}{n} - \frac{1}{n+1}$. Then we see that

$$m_*(A_n \cup B_{n+1}) = m_*(A_n) + m_*(B_{n+1}) \leq m_*(A_{n+2}).$$

Then we see that $m_*(A) \geq \sum_{\text{odd}} m_*(B_n)$ and $m_*(A) \geq \sum_{\text{even}} m_*(B_n)$, so in particular, $\sum_n m_*(B_n) \leq 2m_*(A)$. Then we obtain

$$m_*(A_n) \leq m_*(F^c \cap A) \leq m_*\left(A_n \cup \bigcup_{k \geq n} B_k\right) \leq m_*(A_n) + \sum_{k \geq n} m_*(B_k).$$

Because the sum is finite, we have the desired result. \square

Theorem 4.6. *Let μ be a Borel measure that is finite on compact sets. Then for every $E \in \mathcal{B}$ and $\varepsilon > 0$, there exists an open $G \supseteq E$ and closed $F \subseteq E$ such that $\mu(G \setminus E) < \varepsilon$ and $\mu(E \setminus F) < \varepsilon$.*

Proof. Let b be a collection of sets satisfying both properties. We show this is a σ -algebra containing \mathcal{B} . Clearly, it is closed under complements. Thus suppose $E_k \in b$ for each k . First, there exist $G_k \supseteq E_k$ open such that $\mu(G_k \setminus E_k) < \frac{\varepsilon}{2^k}$, so their union is open and $\mu(G \setminus E) < \varepsilon$.

Similarly, there exists $F_k \subseteq E_k$ such that $\mu(E_k \setminus F_k) < \frac{\varepsilon}{2^k}$, but the union $F^* = \bigcup F_k$ is not closed. We prove that there exists a closed $F \subseteq F^*$ such that $\mu(F^* \setminus F) < \varepsilon$. Assume that F_k are increasing, so choose x_0 and set B_n . Then let $A_n = \overline{B_n} \setminus B_{n-1}$, so write $F^* = \bigcup_n (F^* \cap A_n)$, where $F^* \cap A_n = \lim_k F_k \cap A_n = \bigcup_k (F_k \cap A_n)$.

For each n , there exists $k(n)$ such that $\mu(F^* \setminus F_{k(n)} \cap A_n) < \frac{\varepsilon}{2^n}$ with $\mu(\overline{B_n}) < \infty$. Then we see that $F = \bigcup_{n \geq 1} (F_{k(n)} \cap A_n)$, so $\mu(F^* \setminus F) < \varepsilon$ and each $F \cap \overline{B_k}$ is closed, so F is closed.

Finally, we show that open $G \in b$. Set $F_k = \left\{x \in \overline{B_k} \mid d(x, G^c) \geq \frac{1}{k}\right\}$. Then we use the argument above. \square

4.2 RESULTS AND APPLICATIONS

Definition 4.7. A *premeasure* is a measure-like function on an algebra A (closed under complement and finite union).

We see that a premeasure defines an outer measure. Define f to be *measurable* if the preimage of any open set is a Borel set and define f to be *simple* if

$$f = \sum a_n \chi_{E_n}$$

for E_n measurable. Then given a measure μ , define

$$\int f d\mu = \sup_{\psi \leq f \text{ simple}} \int \psi d\mu,$$

where $\int \psi d\mu = \sum a_n \mu(E_n)$.

Theorem 4.8. $\mu_*|_{\Sigma_A}$ is a measure where Σ_A is the smallest σ -algebra containing A .

We have the following results, in analogy with the Lebesgue case.

Theorem 4.9 (Egorov). *Let f_k be defined on E with $\mu(E) < \infty$. Then for all $\varepsilon > 0$, there exists $A_\varepsilon \subseteq E$ with $\mu(E \setminus A_\varepsilon) < \varepsilon$ and such that $f_k \rightarrow f$ uniformly on A_ε .*

Lemma 4.10 (Fatou). *If f is nonnegative, then $\int \liminf f_n d\mu \leq \liminf \int f_n d\mu$.*

As before, if $f = f_+ - f_-$ with both $f_+, f_- \geq 0$, then $\int f d\mu$ is defined provided that not both $\int f_+ d\mu$ and $\int f_- d\mu$ are infinite. Then f is

1. *Integrable* if $\int_X |f| d\mu < \infty$. We say $f \in L^1(X, \mu)$;
2. *Square-integrable* if $\int_X f^2 d\mu < \infty$. We say that $f \in L^2(X, \mu)$.

Theorem 4.11 (Monotone Convergence). *If $f_n \geq 0$ and $f_n \nearrow f$ almost everywhere, then $\int f d\mu = \lim \int f_n d\mu$.*

Theorem 4.12 (Dominated Convergence). *If $f_n \rightarrow f$ almost everywhere and $|f_n| \leq g$ for some integrable g , then $\int_X |f_n - f| d\mu \rightarrow 0$.*

Example 4.13. Consider the Stieltjes integral. Recall that the Riemann integral is $\int f dx$, and the Stieltjes integral is $\int f dg$ for some monotone g . Finally, the Lebesgue integral is the correct definition of integration.

To construct a measure associated to a monotone function F , first normalize by changing F at the jumps to ensure right continuity.

Theorem 4.14. *Given a monotone, right-continuous F , there exists a unique Borel measure μ_F such that $\mu_F([a, b]) = F(b) - F(a)$ for all $a < b$. Moreover, if μ is a Borel measure that is finite on compact sets, there exists a unique F (up to constants) such that $\mu = \mu_F$.*

Proof. The first part is straightforward, but painful (Robin's words), so it is not given here. On the other hand, if μ is a given Borel measure and is finite on compact sets, how do we define F ? We want $\mu((a, b]) = F(b) - F(a)$. If $x > 0$, then we must have $\mu((0, x]) = F(x)$ and similarly for $x < 0$. It is easy to see that this function is monotone, so we show right continuity:

$$\begin{aligned}
 F(0+) &= \lim_{\varepsilon \rightarrow 0} F(\varepsilon) \\
 &= \lim_{\varepsilon \rightarrow 0} \mu(0, \varepsilon] \\
 &= \mu\left(\lim_{\varepsilon \rightarrow 0} (0, \varepsilon]\right) \\
 &= \mu(\emptyset) \\
 &= 0.
 \end{aligned}$$

□

4.3 SIGNED MEASURES

Note that we can add measures and multiply them by nonnegative constants. How can we obtain full linearity?

Definition 4.15. A function ν on a σ -algebra Σ is a *signed measure* if

1. $\nu(\emptyset) = 0$;
2. ν does not take on both $\pm\infty$;
3. ν is additive on disjoint sets. In particular, if $\nu(E)$ is finite and $E = \bigcup E_j$, then $\sum \nu(E_j)$ converges absolutely.

We say that $A \subseteq X$ is *positive with respect to ν* if $A \in \Sigma$ and for each $E \subseteq A$ with $E \in \Sigma$, we have $\nu(E) \geq 0$. Define similar for negativity. Then a set N is *null* if it is both positive and negative. However, having measure 0 does **not** imply that a set is null.

Lemma 4.16. *Subsets and countable unions of positive sets are positive.*

Proof. Subsets of positive sets are positive by definition. For the second part, if $E \subseteq \bigcup A_n$, set $E_n = E \cap A_n \cap A_{n-1}^c \cap \dots \cap A_1^c$. Then $v(E_n) \geq 0$ and $v(E) = \sum v(E_n) \geq 0$. \square

Lemma 4.17. *If E satisfies $0 < v(E) < \infty$, then there exists a positive $A \subseteq E$ with $v(A) > 0$.*

Proof. If E is positive, we are done. Otherwise, let n_1 be the smallest integer such that there exists $E_1 \subseteq E$ with $v(E_1) < -\frac{1}{n_1}$. Inductively, choose n_k to be the smallest integer such that there exists $E_k \subseteq E \setminus \left(\bigcup_{j < k} E_j\right)$ with $v(E_k) < -\frac{1}{n_k}$. This process may or may not terminate, so set $A = E \setminus \bigcup E_k$. Then clearly A has positive measure because $v(A) = v(E) - \sum v(E_k) > 0$ and $\sum v(E_k)$ converges absolutely. Therefore $\sum \frac{1}{n_k} < -\sum v(E_k)$ is finite, so either the process stops or $n_k \rightarrow \infty$, so $v(A) < \infty$.

To show that A is positive, it is clear if the process stops. Otherwise, for all $\varepsilon > 0$, there exists k such that $\frac{1}{n_k - 1} < \varepsilon$. Then $A \subseteq E \setminus \bigcup_{j=1}^{n_k} E_j$, so A contains no subset with measure less than $-\varepsilon$. \square

Theorem 4.18 (Hahn Decomposition Theorem). *Given a signed measure ν , there exists A positive and B negative such that $X = A \cup B$ and $A \cap B = \emptyset$.*

Proof. Assume that ν does not take the value ∞ . Define $\lambda = \sup_{A \text{ positive}} \nu(A)$. Note that $\lambda \geq 0$. Then choose A_i positive such that $\lambda = \lim \nu(A_i)$ and set $A = \bigcup A_i$ is positive and that $\lambda \geq \nu(A)$.

Also, $A \setminus A_i \subseteq A$, so each $\nu(A \setminus A_i) \geq 0$. Thus $\nu(A) \geq \nu(A_i)$ for all i , so $\nu(A) = \lambda$. Set $B = A^c$. We will show that B is negative. To see this, if $E \subseteq B$ with $\nu(E) > 0$, then $A \cup E$ is positive with measure strictly greater than λ . \square

Note that this A, B are not unique. Then define the positive and negative parts of ν by $\nu_+(E) = \nu(E \cap A)$ and $\nu_-(E) = -\nu(E \cap B)$. These are both measures. Then we see that $\nu = \nu_+ - \nu_-$. This decomposition is unique. Also, define the *total variation* of ν as $|\nu| = \nu_+ + \nu_-$.

Note that $|\nu|$ is defined because ν_+, ν_- cannot both be infinite. If $|\nu| < \infty$, we say that ν is *finite*. If two measures μ_1, μ_2 have sets A_1, A_2 such that $A_1 \cap A_2 = \emptyset, A_1 \cup A_2 = X$ and $\mu_1(A_2) = \mu_2(A_1) = 0$, then we call them *mutually singular* and write $\mu_1 \perp \mu_2$.

Example 4.19. The discrete measure is mutually singular with the Lebesgue measure. The Cantor measure is also mutually singular with the Lebesgue measure.

The opposite notion is absolute continuity of measures.

Definition 4.20. We say that ν is *absolutely continuous* with respect to μ , written $\nu \ll \mu$, if for all E with $\mu(E) = 0$, $\nu(E) = 0$.

Example 4.21. If E is measurable and f is integrable, then $\nu_f(E) = \int_E f d\mu$ is absolutely continuous with respect to μ .

Theorem 4.22 (Radon-Nikodym). *Let (X, \mathcal{B}, μ) be a measure space with Borel sets and suppose μ is σ -finite and $\nu \ll \mu$. Then there exists a nonnegative μ -measurable f such that for all $E \in \mathcal{B}$, $\nu(E) = \int_E f d\mu$. Here, f is called the Radon-Nikodym derivative of ν , sometimes written $f = \left[\frac{d\nu}{d\mu} \right]$.*

Before we prove this result (following Royden), we need two lemmas:

Lemma 4.23. *If B_α is a collection of Borel sets ordered by some countable $\alpha \in D$ satisfying $B_\alpha \subseteq B_\beta$ for $\alpha < \beta$, then there exists a measure f such that $f \leq \alpha$ on B_α and $f \geq \alpha$ on B_α^c .*

Proof. Given $x \in X$, set $f(x) = \inf\{\alpha \mid x \in B_\alpha\}$. Then if $x \in B_\alpha$, $f(x) \leq \alpha$. Then if $x \notin B_\alpha$, $f(x) \geq \alpha$ because $x \notin B_\beta$ for $\beta < \alpha$. To show that f is measurable, note that

$$\{x \mid f(x) < \alpha\} = \bigcup_{\beta < \alpha} B_\beta,$$

which is measurable. \square

Lemma 4.24. *Let $\alpha \in D$ countable and $B_\alpha \in \mathcal{B}$ such that if $\alpha < \beta$, then $\mu(B_\alpha \setminus B_\beta) = 0$. Then the same conclusion as in Lemma 4.23 holds.*

Proof. Let $C = \bigcup_{\alpha < \beta} B_\alpha \setminus B_\beta$. Set $B'_\alpha = B_\alpha \cup C$. Note that C has measure 0. Then apply Lemma 4.23 to the B'_α . \square

Proof of Radon-Nikodym. Assume μ is finite. For all rational α , consider the signed measure $\nu - \alpha\mu$. By the Hahn decomposition theorem, there exist $A_\alpha, B_\alpha = A_\alpha^c$ with $\nu - \alpha\mu \geq 0$ on A_α and $\nu - \alpha\mu \leq 0$ on B_α . Then set $A_0 = X, B_0 = \emptyset$ and note that $B_\alpha \setminus B_\beta = B_\alpha \cap A_\beta$, so $\nu - \alpha\mu \leq 0$ on B_α and $\nu - \beta\mu \geq 0$ on A_β . Therefore, if $\beta > \alpha$, then $\mu(B_\alpha) \setminus B_\beta = 0$, so by Lemma 4.24, there exists f such that $f \geq \alpha$ on A_α and $f \leq \alpha$ on B_α . Because $B_0 = \emptyset$, $f \geq 0$.

We now show that $\int_E f d\mu = \nu(E)$. Fix n and set

$$E_k = E \cap \left(B_{\frac{k+1}{n}} \setminus B_{\frac{k}{n}} \right), E_\infty = E \setminus \bigcup_k B_{\frac{k}{n}}.$$

Then E is the disjoint union $E = \bigsqcup_{k=0}^{\infty} E_k \cup E_\infty$, so

$$\nu(E) = \sum \nu(E_k) + \nu(E_\infty).$$

Now $E_k \subseteq B_{\frac{k+1}{n}} \cap A_{\frac{k}{n}}$, so $\frac{k}{n} \leq f \leq \frac{k+1}{n}$ on E_k , so

$$\frac{k}{n} \mu(E_k) \leq \int_{E_k} f d\mu \leq \frac{k+1}{n} \mu(E_k).$$

By definition of B, A , we know that $\frac{k}{n} \mu(E_j) \leq \nu(E_k) \leq \frac{k+1}{n} \mu(E_k)$, so

$$\nu(E_k) - \frac{1}{n} \mu(E_k) \leq \int_{E_k} f d\mu \leq \frac{1}{n} \mu(E_k).$$

Then add and take $n \rightarrow \infty$. \square