Notes on PMATH-450:
Lebesgue Integration and Fourier Analysis

Unversity of Waterloo

David Duan

Last Updated: May 7, 2021 (draft)

Contents

Chapter 1

Introduction

Section 1. The Axiom of Choice

1.1. Motivation: The **Axiom of Choice** is independent of other axioms of set theory and is important for doing analysis. Informally, it says that given any collection of bins, each containing at least one object, it is possible to make a selection of exactly one object from each bin, even if the collection is infinite.

1.2. Axiom (Axiom of Choice): *Let $\mathcal{F}$ be a non-empty collection of non-empty sets, say*

$$\mathcal{F} = \{A_\lambda : \lambda \in \Lambda\},$$

where the index set $\Lambda \neq \varnothing$ and each set $A_\lambda \neq \varnothing$. Then there is a function

$$f : \Lambda \to \bigcup_{\lambda \in \Lambda} A_\lambda$$

such that $f(\lambda) \in A_\lambda$ for each $\lambda \in \Lambda$.

1.3. This axiom becomes important when Λ is uncountable. Consider $\mathbb{R}$ with the equivalence $x \sim y$ if $x - y \in \mathbb{Q}$. This equivalence relation partitions $\mathbb{R}$ into uncountably many disjoint sets. How can you pick on representative from each class? This question has no straightforward answer.

1.4. Definition: Let S be a set. A relation $\leq$ on S is called a **partial order** if

(1). (**reflective**) $x \leq x$ for all $x \in S$;

(2). (**anti-symmetric**) $x \leq y$ and $y \leq x$ implies that $x = y$ for all $x, y \in S$;

(3). (**transitive**) $x \leq y$ and $y \leq z$ implies that $x \leq z$ for all $x, y, z \in S$.

A partial order on S is called a **total order** if for every $x, y \in S$, either $x \leq y$ or $y \leq x$.

1.5. Example: On $\mathbb{R}^2$, define the relation $(x, y) \leq (x', y')$ if $x \leq x'$ and $y \leq y'$. This is a partial order (easy) but not a total order. Indeed, consider $a = (1, 4)$ and $b = (2, 3)$. Since $1 \leq 2$ but $3 \leq 4$, neither $a \leq b$ nor $b \leq a$ holds.

1.6. Example: Let S be a set. Define $\leq$ on the power set of S by the rule $A \leq B$ if $A \subseteq B$. We will refer to this as the **partial order by inclusion**. It is easy to see that inclusion is a partial order. To see it is not a total order, consider $S = \{0, 1, 2\}$ with two subsets $S_1 = \{0, 1\}$ and $S_2 = \{1, 2\}$. Observe we cannot compare S_1 and S_2. This is therefore not a total order.

1.7. Definition: A partially ordered set $(S, \leq)$ is said to be **well-ordered** if every non-empty subset of S has a smallest element. In other words, if $T \subseteq S$ and $T \neq \varnothing$, then there exists $x \in T$ such that $x \leq y$ for every $y \in T$.

1.8. Remark: Any well-ordered set is totally ordered since within every size-2 set, we can tell which one is “smaller”.

1.9. Example: $\mathbb{N}$ is well-ordered as there is always a smallest element in a subset of natural numbers. $\mathbb{Z}$ is totally ordered but not well-ordered as itself is unbounded below. $\mathbb{Q}_{\geq 0}$ is not-well ordered since the subset $\mathbb{Q}_{>0}$ has no minimal element (even though $\mathbb{Q}_+$ it self has a minimal element). Note this example emphasizes that in the definition of well-ordered sets, we are considering *every subset* of S, not just S itself.

1.10. Example: Every countable set can be well-ordered. Indeed, recall that a set S is countable means that we could define a bijection f between $\mathbb{N}$ and S. Now for any two $x, y \in S$, define $x \leq y$ if $f(x) \leq f(y)$ in the usual order in $\mathbb{N}$. This is clearly a partial order. Then, for any subset of A, we can use f to find a minimal element in the set since $\mathbb{N}$ is well-ordered.

1.11. Note: One of the consequences of AC is known as the **Well-Ordering Principle**, which states that *every-set can be well-ordered.* In other words, on every set it is possible to define a partial order that makes the set well-ordered.

1.12. Definition: Let $(S, \leq)$ be a partially ordered set.

- An **upper bound** for $T \subseteq S$ is an element $s \in S$ such that $s \geq t$ for every $t \in T$.
- An element $s \in S$ is said to be **maximal** if whenever $x \in S$ and $x \geq s$, then $x = s$.
- A **chain** is a totally ordered subset of S.

1.13. Remark:

- Note the upper bound s need not be in the subset T in the definition above.
- Partially ordered sets need not have maximal elements, e.g., $\mathbb{N}$ is not bounded above with the usual ordering. Furthermore, maximal elements (if exist) need not be unique. For example, take $S = \{a, b, c\}$ with the ordering $a \leq b$, $a \leq c$, and $x \leq x$ for $x = a, b, c$. Then both b and c are maximal elements.

1.14. Example: Let $S \neq \varnothing$ be a set and define the partial order by inclusion on $\mathcal{P}(S)$. Then S is the unique maximal element for X. An example of a chain in $\mathcal{P}(S)$ is a set $\{B_i\}_{i=1}^{\infty}$ where $B_i \subseteq B_{i+1} \subseteq X$ for all i (note we can have $B_1 = \varnothing$).

1.15. Note: Another useful consequence of AC is **Zorn's Lemma**: *Suppose S is a non-empty, partially ordered set in which each chain has an upper bound. Then S has a maximal element.* In fact, AC, Zorn's Lemma, and the Well-Ordering Principle are equivalent.

1.16. We conclude this section with an important and standard application of Zorn's Lemma.

1.17. Theorem: *Every vector space has a basis.*

Proof. Let V be a vector space and S be the set of linear independent subsets of V. Partially order S by inclusion. Let C be any chain in S. We show that C has an upper bound and hence, by Zorn's Lemma, S has a maximal element. We then prove that this maximal element is a basis for V.

Let $Y = \bigcup_{W \in C} W$. Clearly, $Y \supseteq W$ for every $W \in C$, so assuming $Y \in S$, then it is an upper bound for the chain C. To check Y is an element of S, we need to show its linear independence. Let $x_1, \ldots, x_n \in Y$. Then for each $i \in \{1, \ldots, n\}$, there is some set $W_i \in C$ such that $x_i \in W_i$. Since C is totally ordered and there are only finitely many sets W_i, one of these sets must be maximal, i.e., one of the W_i, call it W^*, contains all the other sets W_i. This means $x_i \in W^*$ for all i. Now since $W^* \in C \subseteq S$, W^* is linearly independent. Thus

$$\exists \alpha_1, \ldots, \alpha_n \in \mathbb{F} : \sum_{i=1}^{n} \alpha_i x_i = 0 \implies \forall i \in \{1, \ldots n\} : \alpha_i = 0.$$

This shows that Zorn's Lemma can be applied, i.e., S contains some maximal elements.

Let M be a maximal element of S. Since $M \in S$, M is linearly independent. It remains to check it spans V. Suppose not. Then there exists some $x \in V$ with $x \notin \operatorname{span}(M)$. But then $M \cup \{x\}$ is linearly independent, so $M \cup \{x\} \in S$ and properly contains M, so M is no longer maximal in S. It follows that M spans V and is a basis for V. $\square$

1.18. Note that we followed the convention of commenting on the application of Zorn's Lemma.

Section 2. Review: Riemann Integral

2.1. Motivation: In this chapter, we review the basic definitions of the Riemann integral. See MATH-148 for more details.

2.2. Note: Let $f : [a, b] \to \mathbb{R}$ be bounded. Consider a **partition** $P : a = x_0 < x_1 < \cdots < x_n = b$ of $[a, b]$. Define the **upper** and **lower Riemann sums**:

$$U(f, P) = \sum_{i=1}^{n} \sup f\Big|_{[x_{i-1}, x_i]} (x_i - x_{i-1})$$
$$L(f, P) = \sum_{i=1}^{n} \inf f\Big|_{[x_{i-1}, x_i]} (x_i - x_{i-1}).$$

Obviously, we have $U(f, P) \geq L(f, P)$ for all such f, P.

We say the partition Q is a **refinement** of P if Q contains P (and possibly more points). If Q is a refinement of P, then $U(f, Q) \leq U(f, P)$ and $L(f, Q) \geq L(f, P)$. If P_1, P_2 are any two partitions of $[a, b]$ and Q is their common refinement (the union of two partitions), then $U(f, P_1) \geq U(f, Q) \geq L(f, Q) \geq L(f, P_2)$. Hence, every upper sum (for a fixed f) dominates every lower sum.

2.3. Definition: If $\inf\{U(f, P) : \text{all partitions } P\} = \sup\{L(f, P) : \text{all partitions } P\}$, we say that f is **Riemann integrable** over $[a, b]$ and write

$$R - \int_a^b f = \inf_P U(f, P) = \inf_P L(f, P).$$

2.4. Remark: We use $R - \int_a^b f$ to emphasize that this is the Riemann integral, as we will be defining a new integral which generalizes this integral.

2.5. Note: The Riemann integral has many good properties. For instance, as a consequence of the Fundamental Theorem of Calculus, we can (sometimes) determine the Riemann integral by finding an antideriative, rather than having to go back to the definition. Also, any continuous function, or even any function with only finitely many discontinuities, is Riemann integrable. However, not all functions are Riemann integrable. Later, we will show prove a characterization of the Riemann integrable functions.

2.6. Definition: Given $A \subseteq X$, define the **characteristic function of** A by

$$\chi_A : X \to \{0, 1\}$$
$$x \mapsto \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

2.7. Example: Here is a function that is not Riemann integrable. Consider $\chi_{\mathbb{Q}}$ and $[a, b] = [0, 1]$. Then the sup of $\chi_{\mathbb{Q}}$ on any subinterval of $[0, 1]$ equals 1 and the inf is 0. Thus for all P, $U(f, P) = 1$ and $L(f, P) = 0$, so f is not Riemann integrable.

2.8. Besides some common functions are not necessarily Riemann integrable, another of its weaknesses is its poor behavior under limits.

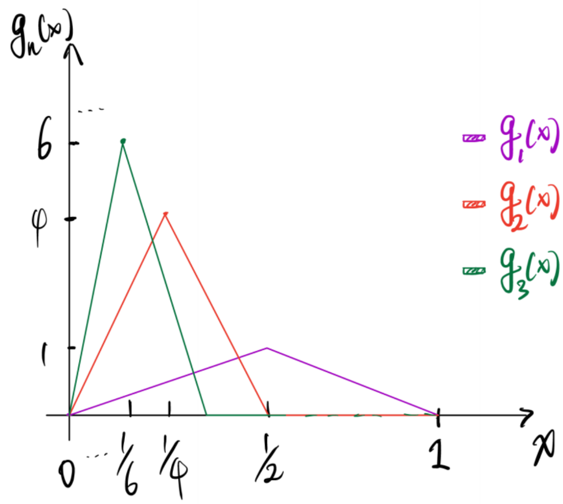

Figure 1.1: g_n. (Credit: Bill Zhuo)

2.9. Example: Consider g_n defined on $[0, 1]$ by

$$g_n(x) = \begin{cases} 0 & 0 \\ n & x = 1/2n \\ 0 & x \in [1/n, 1] \\ \text{piecewise linear} & \text{on } [0, 1/(2n)] \text{ and } [1/(2n), 1/n] \end{cases}$$

Each g_n is continuous and $g_n \to 0$ piecewise. Using the common area interpretation, we have

$$\forall n : R - \int_0^1 g_n = \frac{1}{2} \implies \lim_{n\to\infty} \left[R - \int_0^1 g_n \right] = \frac{1}{2}.$$

However,

$$R - \int_0^1 \left[\lim g_n \right] = 0.$$

In other words, we cannot switch the limit and the Riemann integral without uniform convergence. Indeed, Riemann integral itself is a limit process (approximating the "true" integral by Riemann sums), and whenever we are switching two limiting processes, there should be some concerns.

2.10. Example: Let $\{r_n\}_{n=1}^{\infty} = \mathbb{Q} \cap [0,1]$ be the set of rationals in $[0,1]$. Define

$$f_n(x) = \begin{cases} 1 & x = r_j, j \leq n \\ 0 & \text{otherwise} \end{cases}$$

As f_n has only finitely many discontinuities, it is Riemann integrable. All lower Riemann sums equal 0, hence $R - \int_0^1 f_n = 0$ for all n. But the sequence (f_n) converges pointwise to $f = \chi_{\mathbb{Q}\cap[0,1]}$ which is not Riemann integrable. Thus,

$$\lim_{n\to\infty} \left[R - \int_0^1 f_n \right] \neq R - \int_0^1 \left[\lim f_n \right]$$

as the RHS does not exist. In this case, we see that interchanging limit and Riemann integral might also make things undefined.

Section 3. Motivation: Lebesgue Integral

3.1. The Lebesgue integral will extend the Riemann integral to a larger class of functions. Instead of partitioning the domain, we now partition the range of this function. Say Range$(f) = [A, B]$ and we consider the partition $A = y_1 < \cdots < y_N = B$. Throughout this course, we use

$$f^{-1}(S) = \{x \in \text{dom}(f) : f(x) \in S\}$$

to denote the **preimage** of S under f. We put $E_i = f^{-1}((y_{i-1}, y_i])$.

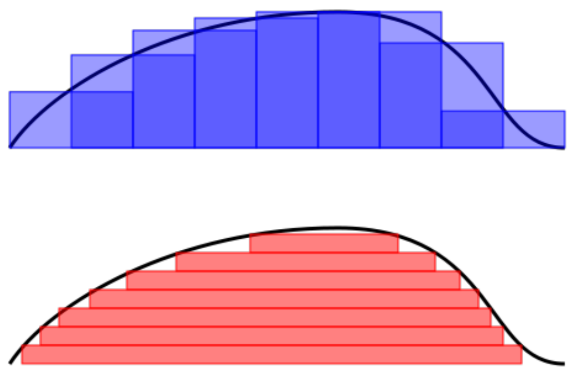

Figure 1.2: Riemann integral (blue) vs Lebesgue integral (red).

3.2. Example: Let $f = \chi_{\mathbb{Q}}$. Then

$$E_i = \begin{cases} \mathbb{Q} & \{1\} \subseteq S, \{0\} \not\subseteq S \\ \mathbb{Q}^c & \{0\} \subseteq S, \{1\} \not\subseteq S \\ \mathbb{R} & \{0,1\} \subseteq S \\ \varnothing & \{0,1\} \not\subseteq S \end{cases}$$

3.3. In genearl, $f(x)$ will lie between y_{i-1} and y_i for $x \in E_i$, so if the partition is very fine, then $f \sim y_i$ on E_i. Since the sets E_i are disjoint and their union is the domain of f, we have

$$f \sim \sum y_i \chi_{E_i}.$$

Roughly speaking, we define the Lebesgue integral of f over the domain to be

$$\lim \sum_i (y_i \cdot \text{length of } E_i)$$

where the limit should be taken over finer and finer partitions and the notion of length needs to be explained. This is where we will begin our study.

Section 4. The Ideal Case: Too Good to be True

4.1. As mentioned in the previous section, we need to extend the notion of intervals to much more general sets. Ideally, we would like to define a function m on all subsets of $\mathbb{R}$, taking on non-negative values or positive infinity,

$$m : \mathcal{P}(\mathbb{R}) \to [0, \infty] = \mathbb{R}_{\geq 0} \cup \{\infty\},$$

with the following properties:

(1). The empty set has zero length, i.e., $m(\varnothing) = 0$.

(2). The length of an interval equals its normal notion of length, i.e., $m(L) = |L|$.

(3). **σ-addivity**: If $\{E_n\}_{n=1}^{\infty}$ are disjoint sets, then

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

Note this also guarantees **monotonicity**, $A \subseteq B \implies m(A) \leq m(B)$.

(4). **Translation invariance**: Let $E \subseteq \mathbb{R}$ and $y \in \mathbb{R}$. Then $m(E) = m(E + y)$.

Unfortunately, it is impossible to define a function with all these properties. This is a consequence of the axiom of choice as we will now prove.

Proof. Define a relation on $\mathbb{R}$ by the rule $x \sim y$ if $x - y \in \mathbb{Q}$. This partitions $\mathbb{R}$ into disjoint equivalence classes, each consists of the numbers of the form $x_0 + \mathbb{Q}$ where $x_0 \in \mathbb{R}$. Thus each class is a translate of $\mathbb{Q}$ and hence is dense in $\mathbb{R}$. In particular, each class intersects $[0, 1/2]$. Use AC to pick a representative of each equivalence class that lies in $[0, 1/2]$. Let E denote this collection of real numbers. Note that if x_1 and x_2 are distinct rational numbers, then the two sets $E + x_1$ and $E + x_2$ are disjoint.

Indeed, if we had some $e_1 + x_1 = e_2 + x_2$ for $e_1, e_2 \in E$ and $x_1, x_2 \in \mathbb{Q}$, then $e_1 - e_2 = x_2 - x_1 \in \mathbb{Q}$. But then $e_1 \sim e_2$ which is impossible unless $e_1 = e_2$ as the elements of E are chosen from distinct equivalence classes. Thus, $e_1 + x_1 = e_2 + x_2 \implies e_1 = e_2$ and that forces $x_1 = x_2$, contradicting our initial assumption. Thus, we have $E + x_1$ and $E + x_2$ disjoint as claimed.

By σ-additivity and translation invariance,

$$m\left(\bigcup_{x \in \mathbb{Q}} E + x\right) = \sum_{x \in \mathbb{Q}} m(E + x) = \sum_{x \in \mathbb{Q}} m(E) = \begin{cases} 0 & m(E) = 0 \\ \infty & \text{otherwise} \end{cases}$$

For any $y \in \mathbb{R}$, $y \sim e$ for some $e \in E$. Then $y - e \in \mathbb{Q}$ so $y = e + x$ for some $x \in \mathbb{Q}$. This shows that $\bigcup_{x \in \mathbb{Q}} E + x = \mathbb{R}$. Hence,

$$m\left(\bigcup_{x \in \mathbb{Q}} E + x\right) = m(\mathbb{R}) = \infty$$

and therefore $m(E) > 0$.

Now assume $Q \cap [0,1/2] = \{r_i\}_{i=1}^{\infty}$. By σ-additivity and using the fact that $m(E) > 0$, we get

$$m\left(\bigcup_{i=1}^{\infty} E + r_i\right) = \sum_{i=1}^{\infty} m(E + r_i) = \sum_{i=1}^{\infty} m(E) = \infty.$$

But also, $E \subseteq [0,1/2]$, so $E + r_i \subseteq [0,1]$ and therefore by monotonicity,

$$m\left(\bigcup_{i=1}^{\infty} E + r_i\right) \leq m([0,1]) = 1,$$

which gives us $\infty \leq 1$. This contradiction shows that such a function m cannot exist. $\square$

> **4.2.** Which property are we willing to give up to obtain such a function m? The only two options are σ-additivity and the choice that m is defined for all subsets. As it turns out, even giving up comfortable but not finite additivity does not give the desired function m. Thus, we will only work with nice subsets of $\mathbb{R}$ where a function with all the desired properties exist.

Part I

Lebesgue Integration

Chapter 2

Lebesgue Measure

Section 5. Outer Lebesgue Measure

5.1. Motivation: As we saw previously, we cannot have a function m with all mentioned properties. The solution is to construct a function with all properties but σ-additivity. However, as this property is really important for limit theorems, we will relax the requirement that m be defined on all subsets of $\mathbb{R}$ by taking the function which we construct without σ-additivity and restricting it to a suitable subset of $\mathcal{P}(\mathbb{R})$, called the **measurable sets**, where it will have σ-additivity. This causes the technical complications as we always need to check whether the sets we want to act m upon are measurable. But, in practice, this is always the case. The function, denoted m^*, which has all but σ-additivity, is called **outer Lebesgue measure**.

5.2. Definition: Given $A \subseteq \mathbb{R}$, define [a]

$$\mathcal{C}(A) = \left\{ \{I_n\}_{n=1}^{\infty} \mid I_n \text{ are open intervals with } A \subseteq \bigcup_{n=1}^{\infty} I_n \right\}.$$

Let $\ell(I)$ denote the length of an interval I. Define [b]

$$m^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \ell(I_n) \mid \{I_n\}_{n=1}^{\infty} \in \mathcal{C}(A) \right\}.$$

[a] In words, $\mathcal{C}(A)$ contains the open covers of A where each cover consists of only intervals.
[b] In words, the outer measure of A is the total length of the "smallest" interval cover of A.

5.3. Note: Some easy properties of m^*:

(1). $m^* : \mathcal{P}(\mathbb{R}) \to [0, \infty]$, $m^*(\varnothing) = 0$, $m^*(\{x\}) = 0$.

(2). **monotonicity**: $m^*(A) \leq m^*(B)$ if $A \subseteq B$.

(3). m^* is **translation invariant** since the length of intervals is translation invariant.

5.4. Proposition: *$m^*(I) = \ell(I)$ if I is an interval.* [a]

[a] This result holds regardless of the type (open/closed/neither, bounded/unbounded).

Proof. We prove the result for $I = [a, b]$ and leave the rest as exercises. For every $\varepsilon > 0$, the interval $(a-\varepsilon, b+\varepsilon)$ is an open interval that covers $[a, b]$. By definition, $m^*(I) \leq \ell((a-\varepsilon, b+\varepsilon)) = b-a+2\varepsilon$. Since this holds for all $\varepsilon > 0$, we have $m^*(I) \leq b-a$. It remains to show that whenever I_n are open intervals with $I \subseteq \bigcup_{n=1}^{\infty} I_n$, then $\sum_{n=1}^{\infty} \ell(I_n) \geq b-a$.

As the sets I_n are an open cover of the compact set I, we can obtain a finite subcover. WLOG, assume it is $I_1, \ldots, I_N$. Since $\sum_{n=1}^{\infty} \ell \geq \sum_{n=1}^{N} \ell(I_n)$, it is enough to check that $\sum_{n=1}^{N} \ell(I_n) \geq b-a$ whenever $I_1 \cup \cdots \cup I_N \supseteq [a, b]$. The endpoint $a \in \bigcup_{n=1}^{N} I_n$, say $a \in I_{n_1} = [a_1, b_1]$. In other words, $a_1 < a < b_1$. If $b_1 > b$, then $(a_1, b_1) \supseteq [a, b]$ and consequently $\sum_{n=1}^{N} \ell(I_n) \geq \ell(I_{n_1}) = b_1 - a_1 \geq b-a$ and we are done.

Assume not. Then $b_1 \in [a,b]$ and thus $b_i \in I_{n_2}$ where $n_2 \neq n_1$ since $b_1 \notin I_{n_1}$. Say $I_{n_2} = (a_2, b_2)$. If $b_2 > b$, then as $b_1 > a_2$, we have

$$\sum_{n=1}^{N} \ell(I_n) \geq \ell(I_{n_1}) + \ell(I_{n_2}) = b_1 - a_1 + b_2 - a_2 \geq b_2 - a_1 \geq b - a$$

and we are done. Otherwise, we repeat to obtain sets $I_{n_j} = (a_j, b_j)$ with $n_j \neq n_1, \ldots, n_{j-1}$ and $a_j < b_{j-1} < b_j$. Since the collection of sets $\{I_n\}_{n=1}^N$ is finite, this process must terminate at some step k, with $b_k > b$. Then

$$\sum_{n=1}^{N} \ell(I_n) \geq \sum_{j=1}^{k} \ell(I_{n_j}) \geq \sum_{j=1}^{k} (b_j - a_j) \geq b - a$$

as desired. □

5.5. Proposition (σ-sub-additivity): *For all sets $A_k \subseteq \mathbb{R}$, we have*

$$m^*\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} m^*(A_k).$$

Note that we do not require the sets A_k to be disjoint.

Proof. Fix $\varepsilon > 0$. From the definition of m^*, it follows that for each A_k, we can choose $\{I_{n,k}\}_{k=1}^{\infty} \in \mathcal{C}(A_k)$ (the set of open interval covers of A) such that $\bigcup_{n=1}^{\infty} I_{n,k} \supseteq A_k$ and

$$\sum_{n=1}^{\infty} \ell(I_{n,k}) \leq m^*(A_k) + \varepsilon 2^{-k}.$$

The collection $\{I_{n,k}\}_{k,n=1}^{\infty}$ is a countable collection of open intervals whose union contains $A := \bigcup_{k=1}^{\infty} A_k$ and hence belongs to $\mathcal{C}(A)$. Thus,

$$m^*\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k,n=1}^{\infty} \ell(I_{n,k}) \leq \sum_{k=1}^{\infty} \left(m^*(A_k) + \varepsilon 2^{-k}\right) = \sum_{k=1}^{\infty} (m^*(A_k)) + \varepsilon$$

As $\varepsilon > 0$ was arbitrary, this completes the proof. □

5.6. Corollary: *If $A = \{x_n\}_{n=1}^{\infty}$ is any countable set, then $m^*(A) \leq \sum_{n=1}^{\infty} m^*(\{x_n\}) = 0$.*

5.7. Corollary: *For any $A \subseteq \mathbb{R}$ and for any $\varepsilon > 0$, there is an open set $O \supseteq A$ such that $m^*(O) \leq m^*(A) + \varepsilon$.*

Proof. Choose open intervals $\{I_n\}_{n=1}^{\infty} \in \mathcal{C}(A)$ such that $\sum_{n=1}^{\infty} \ell(I_n) \leq m^*(A) + \varepsilon$ and put $O = \bigcup_{n=1}^{\infty} I_n$. Then O is open and by σ-sub-additivity,

$$m^*(O) = m^*\left(\bigcup_{n=1}^{\infty} I_n\right) \leq \sum_{n=1}^{\infty} \ell(I_n) \leq m^*(A) + \varepsilon.$$

□

Section 6. Lebesgue Measure

6.1. Motivation: To obtain σ-additivity, we will restrict m^* to a suitable class of subsets of $\mathbb{R}$, known as *Lebesgue measurable sets*. Below is known as the *Caratheodory definition* of Lebesgue measurability.

6.2. Definition: A set $A \subseteq \mathbb{R}$ is said to be **Lebesgue measurable** if for every $E \subseteq \mathbb{R}$, we have

$$m^*(E) = m^*(E \cap A) + m^*(E \cap A^c).$$

6.3. Remark: Since $E = (E \cap A) \cup (E \cap A^c)$, by σ-sub-additivity of m^*, we automatically have $m^*(E) \leq m^*(E \cap A) + m^*(E \cap A^c)$, so we only ever have to check the $\geq$ direction.

6.4. Example: Here are some easy examples of Lebesgue measurable sets.

(1). $\mathbb{R}$ and $\varnothing$. Also any set A with $m^*(A) = 0$.

(2). A^c is measurable iff A is measurable (as the definition is symmetric in A and A^c).

6.5. Proposition: *The interval* (a, ∞) *is Lebesgue measurable.*

Proof. We wish to show that for all $E \subseteq \mathbb{R}$, $m^*(E) = m^*(E \cap (a, \infty)) + m^*(E \cap (-\infty, a])$. Fix E. Let $\varepsilon > 0$ and choose an open interval cover $\{I_n\}_{n=1}^{\infty} \in \mathcal{C}(E)$ of E such that

$$\sum_{n=1}^{\infty} \ell(I_n) \leq m^*(E) + \varepsilon.$$

Let $I_n^+ = I_n \cap (a, \infty)$ and $I_n^- = I_n \cap (-\infty, a]$. Notice that I_n^+ and I_n^- are either intervals or empty. Moreover, the additivity property of length and the fact that the outer Lebesgue measure of an interval is its length gives us $\ell(I_n) = \ell(I_n^+) + \ell(I_n^-) = m^*(I_n^+) + m^*(I_n^-)$.

Since $E \cap (a, \infty) \subseteq \bigcup_{n=1}^{\infty} I_n \cap (a, \infty) = \bigcup_{n=1}^{\infty} I_n^+$, the monotonicity and σ-sub-additivity properties of the outer Lebesgue measure m^* gives

$$m^*(E \cap (a, \infty)) \leq m^* \left(\bigcup_{n=1}^{\infty} I_n^+ \right) \leq \sum_{n=1}^{\infty} m^* \left(I_n^+ \right).$$

Similarly, $E \cap (-\infty, a] \subseteq \bigcup_{n=1}^{\infty} I_n \cap (-\infty, a] = \bigcup_n I_n^-$, so we have

$$m^*(E \cap (-\infty, a]) \leq m^* \left(\bigcup_{n=1}^{\infty} I_n^- \right) \leq \sum_{n} m^* \left(I_n^- \right).$$

Combining both observations, we see that

$$m^*(E \cap (a, \infty)) + m^*(E \cap (-\infty, a]) \leq \sum_{n=1}^{\infty} \left(m^* \left(I_n^+ \right) + m^* \left(I_n^- \right) \right) = \sum_{n=1}^{\infty} \ell\left(I_n \right) \leq m^*(E) + \varepsilon$$

as desired. $\square$

6.6. As we will see, the class of all Lebesgue measurable sets has a nice structure, known as a σ-algebra.

6.7. Definition: A family Ω of subsets of $\mathbb{R}$ is called a **σ-algebra** if:

(1). $\varnothing \in \Omega$.

(2). **closed under complements**: $A \in \Omega \implies A^c \in \Omega$.

(3). **closed under countable union**: $A_1, A_2, \ldots \in \Omega \implies \bigcup_{n=1}^{\infty} A_n \in \Omega$.

6.8. Example: $\{\varnothing, \mathbb{R}\}$ is the smallest σ-algebra. $\mathcal{P}(\mathbb{R})$ is the largest σ-algebra.

6.9. Example: Any intersection of σ-algebras is again a σ-algebra. Let Ω_1, Ω_2 be two σ-algebras and consider $\Omega = \Omega_1 \cap \Omega_2$.

- $\varnothing \in \Omega_1 \wedge \varnothing \in \Omega_2 \implies \varnothing \in \Omega$.
- Let $\{A_n\} \subseteq \Omega$. Then $\forall n : A_n \in \Omega_1 \wedge A_n \in \Omega_2$ so $\bigcup_n A_n \in \Omega_1$, $\bigcup_n A_n \in \Omega_2$. Thus, $\bigcup_n A_n \in \Omega$.
- $A \in \Omega \implies A \in \Omega_1 \wedge A \in \Omega_2 \implies A^c \in \Omega_1 \wedge A^c \in \Omega_2 \implies A^c \in \Omega$.

6.10. Definition: The **Borel σ-algebra** is the intersection of all the σ-algebras containing all the open sets in $\mathbb{R}$. A **Borel set** is any set in the Borel σ-algebra.

6.11. Remark: As $\mathcal{P}(\mathbb{R})$ is a σ-algebra containing all the open sets, the intersection in the definition of the Borel σ-algebra is not vacuous. The Borel σ-algebra is the smallest (in the inclusion sense) σ-algebra containing all the open sets. Since very open set in $\mathbb{R}$ is a countable union of open intervals (see Proposition 0.1 in the appendix), we can equivalently define the Borel σ-algebra as the smallest σ-algebra containing all the open intervals. We will also say the Borel σ-algebra is the smallest σ-algebra "generated" by the open sets (or open intervals).

6.12. Theorem: *The set of Lebesgue measurable sets, denoted $\mathcal{M}$, is a σ-algebra which contains the Borel sets and includes all sets of outer Lebesgue measure zero. Moreover, if $A_n \in \mathcal{M}$ are disjoint, then*

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

Notice m^* restricted to the Lebesgue measurable sets is σ-additive.

Proof. See PMATH-451. □

6.13. Definition: By **Lebesgue measure** m on $\mathbb{R}$, we mean m^* restricted to $\mathcal{M}$.

6.14. Remark: Lebesgue measure m generalizes length of intervals, is σ-additive and translation invariant. However, it is only defined on the sets which are Lebesgue measurable. This is a very rich class of sets since it contains all Borel sets.

6.15. Example: The Cantor set C is compact and measurable. It has empty interior and every point is an accumulation point. It is an uncountable set of Lebesgue measure zero.

Section 7. Properties of Lebesgue Measure

7.1. Theorem: *Here are the important properties of Lebesgue measure.*

- $m(I) = \ell(I)$ *if* I *is an interval.*
- *If* $A \in \mathcal{M}$*, then* $A + t \in \mathcal{M}$ *for all* $t \in \mathbb{R}$ *and* $m(A) = m(A + t)$*.*
- m *is* σ*-additive, i.e., if* $A_1, A_2, \ldots \in \mathcal{M}$ *and are disjoint, then*

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

- m *is* σ*-sub-additive, i.e., if* $A_1, A_2, \ldots \in \mathcal{M}$ *(but are not necessarily disjoint), then*

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

- *If* $A, B \in \mathcal{M}$ *and* $A \subseteq B$*, then* $A^c \cap B \in \mathcal{M}$ *and since* $B = A \cup (A^c \cap B)$*, we have*

$$m(B) = m(A) + m(A^c \cap B) = m(A) + m(B \setminus A).$$

  If, in addition, $m(A) < \infty$*, then* $m(B \setminus A) = m(B) - m(A)$*.*
- m *is monotonic.*
- *A set* E *has Lebesgue measure zero iff for every* $\varepsilon > 0$*, there are open intervals* I_n *such that*

$$\bigcup_{n=1}^{\infty} I_n \supseteq E \quad \textit{and} \quad \sum_{n=1}^{\infty} \ell(I_n) < \infty.$$

- *For any compact set* A*,* $m(A) < \infty$ *as* A *is contained in* $[-N, N]$ *for large enough* N*.*

Proof. Omitted. □

7.2. Proposition: *If* E *is measurable and* $\varepsilon > 0$*, then there is an open set* $O \supseteq E$ *such that* $m(O \setminus E) < \varepsilon$*.*

Proof. First suppose $m(E) < \infty$. By Corollary 5.7, we know there exists an open set $O \supseteq E$ such that $m^*(O) \leq m^*(E) + \varepsilon$. Since every open set is in the Borel σ-algebra, we get $m(O) \leq m(E) + \varepsilon$. Finally, since $m(E) < \infty$, we have $m(O \setminus E) = m(O) - m(E) \leq \varepsilon$.

Now suppose $m(E) = \infty$. For each $n \geq 1$, define $E_n = E \cap [-n, n]$ where $[-n, n]$ is in the Borel σ-algebra and is thus measurable. Then, $m(E_n) \leq 2n < \infty$. By the first case, we have an open set $O_n \supseteq E_n$ for each $n \geq 1$ with $m(O_n \setminus E_n) \leq \varepsilon/2^n$. Let $O = \bigcup_{n=1}^{\infty} O_n$ be the countable union of open set (still open), then $E = \bigcup_{n=1}^{\infty} E_n \subseteq \bigcup_{n=1}^{\infty} O_n = O$ and

$$O \backslash E = \bigcup_{n=1}^{\infty} (O_n \backslash E) = \bigcup_{n=1}^{\infty} \left(O_n \backslash \bigcup_{i=1}^{\infty} E_i\right) \subseteq \bigcup_{n=1}^{\infty} (O_n \backslash E_n).$$

By σ-sub-additivity and monotonicity, we have $m(O \backslash E) \leq \sum_{n=1}^{\infty} m(O_n \backslash E_n) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$. □

7.3. Proposition: *Lebesgue measure works well in taking (inclusion) limits:*

- *Continuity of measure: If* $A_1 \subseteq A_2 \subseteq \cdots \subseteq \bigcup_{n=1}^{\infty} A_n = A$ *and* $A_n \in \mathcal{M}$ *for all* n*, then* $A \in \mathcal{M}$ *and*

$$m(A) = \lim_{n\to\infty} m(A_n).$$

- *Downward continuity of measure: If* $A_1 \supseteq A_2 \supseteq \cdots \supseteq \bigcap_{n=1}^{\infty} A_n = A$ *and* $A_n \in \mathcal{M}$ *for all* n*, then* $A \in \mathcal{M}$*. If, in addition,* $m(A_1) < \infty$*, then*

$$m(A) = \lim_{n\to\infty} m(A_n).$$

Proof. Let $B_1 = A_1$ and for all $n > 1$, let $B_n = A_n \setminus A_{n-1}$. Then each $B_n \in \mathcal{M}$ and these sets are disjoint. Furthermore, $\bigcup_{n=1}^{\infty} B_n = A$ and $\bigcup_{k=1}^{n} B_k = A_n$. By σ-additivity,

$$m(A_n) = \sum_{k=1}^{n} m(B_k) \to \sum_{k=1}^{\infty} m(B_k) = m(A).$$

This proves the first statement. The second proof is similar. Let $B_n = A_1 \setminus A_n$ and $B = A_1 \setminus A$. Then $B_n \in \mathcal{M}$, $B_n \subseteq B_{n+1}$, and $\bigcup_{n=1}^{\infty} B_n = B$. By the first statement, $m(B) = \lim_{n\to\infty} m(B_n)$. The result follows from the fact that $m(B) = m(A_1) - m(A)$ and $m(B_n) = m(A_1) - m(A_n)$. $\square$

7.4. Example: Why is $m(A_1) < \infty$ necessary? Consider $A_n = (n, \infty)$. Clearly that $A_{n+1} \subseteq A_n$ and $\bigcap_{n=1}^{\infty} A_n = \varnothing$. Here, $m(A_n) = \infty$ for all n, while $m(\bigcap_{n=1}^{\infty} A_n) = 0$.

7.5. Remark: Read about the inner Lebesgue measure.

Section 8. Lebesgue's Characterization of Riemann Integrability

8.1. Theorem: *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable over $[a, b]$ iff the set of discontinuities of f has measure zero.*

Proof. Later. □

8.2. Remark: It is immediate from this theorem that every function that is continuous or has countably many discontinuities is Riemann integrable. $\chi_{\mathbb{Q}}$ is an example of a function having uncountably many discontinuities and hence is not Riemann integrable on any interval $[a, b]$.

Appendix: Extra Proofs

0.1. Proposition: *Let $U \subseteq \mathbb{R}$ be open. Then U is a countable union of disjoint intervals.*

Proof. For each $x \in U$, let I_x denote the largest open interval containing x and is contained in U. More precisely, since U is open, x is contained in some small (non-trivial) interval, and therefore if

$$a_x := \inf\{a < x : (a, x) \subseteq U\} \quad \text{and} \quad b_x := \sup\{b > x : (x, b) \subseteq U\},$$

we must have $a_x < x < b_x$. If we now let $I_x = (a_x, b_x)$, then by construction, we have $x \in I_x$ as well as $I_x \subseteq U$. Hence, these intervals cover our open set U:

$$U = \bigcup_{x \in U} I_x.$$

Now suppose that two intervals I_x and I_y intersect. Then their union (which is also an open interval) is contained in U and contains x. Since I_x is maximal, we must have $(I_x \cup I_y) \subseteq I_x$ and similarly $(I_x \cup I_y) \subseteq I_y$. This can happen only if $I_x = I_y$. Therefore, any two distinct intervals in the collection $\mathcal{I} = \{I_x\}_{x \in U}$ must be disjoint.

It remains to show there are only countably many distinct intervals in the collection $\mathcal{I}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, every open interval I_x contains at least one rational number. Since different intervals are disjoint, they must contain distinct rationals. Finally, since $\mathbb{Q}$ is countable, $\mathcal{I}$ must also be countable. This concludes the proof. $\square$

Chapter 3

Measurable Functions

Section 1. Lebesgue Measurable Functions

1.1. Definition: Suppose $X \subseteq \mathbb{R}$ is a Lebesgue measurable set. A function $f : X \to [-\infty, \infty]$ is called **Lebesgue measurable** if the preimages of all open sets are measurable, i.e., for every $\alpha \in \mathbb{R}$, the set [a]

$$\{x \in X : f(x) < \alpha\} = f^{-1}([-\infty, \alpha))$$

is Lebesgue measurable. A complex-valued function $f : X \to \mathbb{C}$ is called **Lebesgue measurable** if both its real and imaginary parts are Lebesgue measurable.

[a] If f is real-valued, then $f^{-1}([-\infty, \alpha)) = f^{-1}((-\infty, \alpha))$.

1.2. Remark: Any constant function f is measurable since $f^{-1}([-\infty, \alpha))$ is either empty or all of $\mathbb{R}$. Any continuous, real-valued function f is measurable since $f^{-1}([-\infty, \alpha)) = f^{-1}((-\infty, \alpha))$ is open and hence is a Borel set.

1.3. Proposition: *$\{x : f(x) < \alpha\}$ is measurable iff $\{x : f(x) > \alpha\}$ is measurable for all $\alpha \in \mathbb{R}$.*

Proof. We show that $\{x : f(x) < \alpha\}$ is measurable iff $\{x : f(x) \leq \alpha\}$ is measurable and then observe that $\{x : f(x) \leq \alpha\}^c = \{x : f(x) > \alpha\}$.

($\Rightarrow$) Suppose $\{x : f(x) < \alpha\}$ is measurable for all $\alpha \in \mathbb{R}$. Then for each $n \in \mathbb{N}$, $\{x : f(x) < \alpha + 1/n\}$ is Lebesgue measurable. We claim that

$$\{x : f(x) \leq \alpha\} = \bigcap_{n=1}^{\infty} \left\{x : f(x) < \alpha + \frac{1}{n}\right\}.$$

This holds since if $y \in \{x : f(x) \leq \alpha\}$, $f(y)$ is certainly less than any $\alpha + 1/n$. Conversely, if

$$y \in \bigcap_{n=1}^{\infty} \left\{x : f(x) < \alpha + \frac{1}{n}\right\} \implies \forall n \in \mathbb{N} : f(y) < \alpha + \frac{1}{n} \implies f(y) \leq \inf_{n \in \mathbb{N}} \left\{\alpha + \frac{1}{n}\right\},$$

so we have $y \in \{x : f(x) \leq \alpha\}$. Finally, since

$$\bigcap_{n=1}^{\infty} \{x : f(x) < \alpha + 1/n\}$$

is the intersection of countably many measurable sets, it is still measurable.

($\Leftarrow$): Now suppose $\{x : h(x) \leq \alpha\}$ is measurable for every $\alpha \in \mathbb{R}$. Then $\{x : f(x) \leq \alpha - 1/n\}$ is Lebesgue measurable. A similar argument as above gives us

$$\{x : f(x) \leq \alpha\} = \bigcup_{n=1}^{\infty} \left\{x : f(x) \leq \alpha - \frac{1}{n}\right\}.$$

Since the RHS being a countable union of measurable sets is measurable, so is the LHS. By our previous comments, the proof is complete. $\square$

1.4. Proposition: *If f is measurable, then $f^{-1}(\{\infty\}) = \{x \in X : f(x) = \infty\}$ is measurable.*

Proof. Observe that

$$\begin{aligned} f^{-1}(\{\infty\}) &= \{x \mid \forall a \in \mathbb{R} : f(x) > a\} \\ &= \{x \mid \forall a \in \mathbb{Q} : f(x) > a\} \\ &= \bigcup_{a \in \mathbb{Q}} \{x : f(x) > a\} \end{aligned}$$

is a countable union of measurable sets and is thus measurable. □

1.5. Proposition: *$f = \chi_E$ is measurable iff $E \subseteq \mathbb{R}$ is measurable.*

Proof. Note that

$$\{x : f(x) < \alpha\} = \begin{cases} \mathbb{R} & \alpha > 1 \\ E^c & 0 < \alpha < 1 \\ \varnothing & \alpha \leq 0 \end{cases}$$

Therefore, $\{x : f(x) < \alpha\}$ is measurable iff E^c is measurable iff E is measurable. □

Section 2. Simple Functions

2.1. Motivation: In the Riemann integral, step functions play a pivotal role. With the Lebesgue integral, what are called *simple functions* are the basic building blocks.

2.2. Definition: A **simple function** is a measurable function that takes finitely many values.

2.3. Example: Let $A_1, \ldots, A_n \subseteq X$ be disjoint sets and $a_1, \ldots, a_n \in \mathbb{R}$ be distinct. Then

$$f = \sum_{i=1}^{n} a_i \chi_{A_i}$$

defines a simple function since $f^{-1}((-\infty, a))$ is a union of at most finitely many measurable sets:

$$f^{-1}((-\infty, a)) = \{x : f(x) < a\} = \bigcup_{i=1}^{N} \{A_i \mid a_i < a\}$$

2.4. Definition: Let f be a simple function whose distinct values are $a_1, \ldots, a_n$. Define $A_i = \{x : f(x) = a_i\}$. Then the **canonical form** of f is given by

$$f = \sum_{i=1}^{n} a_i \chi_{A_i}.$$

Section 3. Properties of Measurable Functions

3.1. Motivation: In this section, we show that measurability is preserved under arithmetic operations and behaves well under limits.

3.2. Proposition: *If f, g are real-valued, measurable functions (with the same domain), then so are $f \pm g$, fg, and f/g if $g \neq 0$.*

Proof for $\pm$. First, note that $\{x : (f+g)(x) < \alpha\} = \{x : f(x) < \alpha - g(x)\}$. Further, $f(x) < \alpha - g(x)$ iff there exists some $r \in \mathbb{Q}$ with $f(x) < r < \alpha - g(x)$. It follows that

$$\{x : (f+g)(x) < \alpha\} = \bigcup_{r \in \mathbb{Q}} \left(\{x : f(x) < r\} \cap \{x : g(x) < \alpha - r\}\right).$$

As f and g are measurable, the set $\{x : f(x) < r\}$ and $\{x : g(x) < \alpha - r\}$ are measurable for each choice of α and r. Since countable unions and intersections of measurable sets are measurable, we conclude that $\{x : (f+g)(x) < \alpha\}$ is measurable for each α. □

Proof for fg and f/g. We first show that f is measurable implies f^2 is measurable. Let $\alpha \in \mathbb{R}$. If $\alpha < 0$, then $\{x : f^2(x) < \alpha\} = \varnothing$ which is trivially measurable. Now assume that $\alpha \geq 0$. Note that

$$\{x : f^2(x) < \alpha\} = \{x : f(x) < \sqrt{\alpha}\} \cup \{x : f(x) > -\sqrt{\alpha}\},$$

which is the union of two measurable sets and thus measurable. In particular, the second set is the complement of $\{x : f(x) \leq -\sqrt{\alpha}\}$ which we know is measurable by measurability of f. Now use

$$fg = \frac{1}{2}((f+g)^2 - f^2 - g^2)$$

to conclude the proof. We now show that $1/g$ is measurable when $g \neq 0$ is measurable. Let $a \in \mathbb{R}$. Then $\{x : 1/g(x) < a\} = \{x : g(x) > 1/a\}$ which is measurable if $a \neq 0$. If $a = 0$, then the set becomes $\{x : g(x) < 0\}$ which is still measurable. □

3.3. Proposition: *If $\{f_n\}_{n=1}^{\infty}$ are measurable functions, then so are $\sup_{n \in \mathbb{N}} f_n$ and $\inf_{n \in \mathbb{N}} f_n$. Moreover, if $f_n \to f$ pointwise as $n \to \infty$, then f is measurable.*

Proof. The set

$$\{x : \sup_n f_n > a\} = \bigcup_{n=1}^{\infty} \{x : f_n > \alpha\}$$

and hence is a countable union of measurable sets. Thus $\sup_n f_n$ is a measurable function. The proof for $\inf_n f_n$ is similar. Now if $f_n \to f$, then

$$f = \liminf f_n = \sup_n (\inf_{k \geq n} f_k)$$

and thus f is measurable. □

3.4. Theorem: *The positive, real-valued function $f : \mathbb{R} \to [0, \infty)$ is measurable iff there are simple functions ϕ_n, $n \in \mathbb{N}$, such that $\phi_n \leq \phi_{n+1}$ and $\phi_n \to f$ pointwise.*

Proof. One direction is easy: simple functions are measurable and their pointwise limit is measurable. Now suppose f is measurable and let we construct the functions ϕ_n. For $k = 0, 1, \ldots, n2^n - 1$, define measurable sets $E_{n,k}$ as follows:

$$E_{n,k} = f^{-1}([k2^{-n}, (k+1)2^{-n})) = \{x : k2^{-n} \leq f(x) < (k+1)2^{-n}\}.$$

Let E_n be the measure set given by

$$E_n = \bigcup_{k=0}^{n2^n-1} E_{n,k} = f^{-1}([0, n))$$

and put

$$\begin{aligned}\phi_n(x) &= \begin{cases} k2^{-n} & x \in E_{n,k} \\ n & x \notin E_n \end{cases} \\ &= \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \chi_{E_{n,k}} + n\chi_{E_n^c}.\end{aligned}$$

The functions ϕ_n are simple and the choice of partitioning the range into subintervals of width 2^{-n} ensures that $\phi_n \leq \phi_{n+1}$.

It remains to check that $\phi_n \to f$ pointwise. For this, fix x and assume $f(x) < N$. Then $x \in E_n$ for all $n \geq N$, say $x \in E_{n,k}$ for the suitable choice of k. The definitions ensure that then

$$f(x) \in [k2^{-n}, (k+1)2^{-n})$$

and

$$\phi_n(x) = k2^{-n}.$$

Hence $|f(x) - \phi_n(x)| < 2^{-n}$ for all $n \geq N$ and certainly implies $\phi_N(x) \to f(x)$. □

3.5. Remark: The fact that we can arrange for the simple functions ϕ_n to be increasing to f will be helpful later.

Section 4. Almost Everywhere

4.1. Motivation: Two functions are said to equal almost everywhere if they agree except on a set that is "invisible" as far as Lebesgue measure is concerned.

4.2. Definition: We say $f = g$ **almost everywhere**, and write $f = g$ **a.e.**, if

$$m(\{x : f(x) \neq g(x)\}) = 0.$$

4.3. Example: $\chi_{\mathbb{Q}} = 0$ a.e.

4.4. Proposition:

- *If $f = 0$ a.e., then f is measurable.*
- *If $f = g$ a.e. and f is measurable, then g is measurable.*

Proof. For $a \in \mathbb{R}$, if $a \leq 0$, then $\{x : f(x) < a\} \subseteq \{x : f(x) \neq 0\}$ is measurable with measure 0. If $a > 0$, then $\{x : f(x) > a\} \subseteq \{x : f(x) \neq 0\}$ is measurable with measure 0. Thus, f is measurable. For the second claim, let $h = f - g = 0$ a.e. Apply Statement 1 on h, we see h is measurable. It follows that $g = f - h$ is measurable. $\square$

Chapter 4

Lebesgue Integral

Section 1. Lebesgue Integral

1.1. Motivation: Recall that for a step function

$$\phi = \sum_{k=1}^{N} a_k \chi_{I_k}$$

where I_k are intervals and $a_j \in \mathbb{R}$, the Riemann integral is given by

$$R - \int_a^b \phi = \sum_k a_k \cdot \ell(I_k \cap [a,b]).$$

This motivates our definition of the Lebesgue integral of simple functions.

1.2. Definition: Let $\phi \geq 0$ be a simple function with standard representation

$$\phi = \sum_{k=1}^{N} a_k \chi_{E_k}.$$

We define the **Lebesgue integral** of ϕ over the measurable set E as

$$\int_E \phi = \sum_{k=1}^{N} a_k \cdot m(E_k \cap E).$$

1.3. Remark: Sometimes we write $\int_E \phi \, dm$ to emphasize that this is the integral wrt the Lebesgue measure. This will be more relevant in PMATH-451 as we will be integrating wrt arbitrary measures. Sometimes we write $\phi_E \phi(x)\, dx$ to emphasize that x is the independent variable.

1.4. Example:

- If ϕ is simple and $E = [a,b]$, then $\int_E \phi = R - \int_a^b \phi$.
- If $\phi = \chi_{\mathbb{Q}}$ and $E = [0,1]$, then as ϕ has standard representation $\phi = 1 \cdot \chi_{\mathbb{Q}} + 0 \cdot \chi_{\mathbb{Q}^c}$, we have

$$\int_E \phi = 1 \cdot m(\mathbb{Q} \cap [0,1]) + 0 \cdot m(\mathbb{Q}^c \cap [0,1]) = 0.$$

1.5. Definition: Suppose $f : E \subseteq \mathbb{R} \to [0, \infty]$ is measurable and E is a measurable set. The **Lebesgue integral** of f over E is given by

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

1.6. Remark: It's easy to see that Lebesgue integral is **monotonic**. If f and g are measurable functions with $0 \leq f \leq g$, then $\int_E f \leq \int_E g$ as any $\phi \leq f$ also satisfies $\phi \leq g$.

1.7. Lemma: *If $\phi = 0$ a.e., then $\int_E \phi = 0$ for any measurable set E.*

Proof. Since $\phi = 0$ a.e., $\{x : \phi(x) \neq 0\}$ has measure 0. Let $E \subseteq \mathbb{R}$ be any measurable set. By definition, we have

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

For each simple function φ in this set, $\{x : \varphi(x) \neq 0\} \subseteq \{x : \phi(x) \neq 0\}$ so each of them are also equal to 0 a.e. Then

$$\int_E \varphi = \sum_{k=1}^{N} a_k \cdot m(E_k \cap E) = 0.$$

Taking the supremum over all φ, we get $\int_E \phi = 0$ as desired. $\square$

1.8. Definition: For measurable E and $f : E \to \mathbb{R}$, the **Lebesgue integral** of f over E is

$$\int_E f = \int_E f^+ - \int_E f^-$$

provided this is not a $\infty - \infty$ form, where

$$f^+(x) = \max(f, 0) = \begin{cases} f(x) & f(x) \geq 0 \\ 0 & \text{else} \end{cases}$$

$$f^-(x) = \max(-f, 0) = \begin{cases} 0 & f(x) \geq 0 \\ -f(x) & \text{else} \end{cases}$$

1.9. Intuition: Geometrically, we are adding up the areas above and below the x-axis.

1.10. Remark: Note that both $f^+, f^- : E \to [0, \infty)$ and $f = f^+ - f^-$. Moreover, both f^+ and f^- are measurable. This implies that $|f| = f^+ + f^-$ is also measurable.

1.11. Definition: We say the measurable function f is **integrable** on E if $\int_E |f| < \infty$.

1.12. Remark: Since $f^+, f^- \leq |f|$, by the monotonicity property of the integral of non-negative functions, $\int_E f^+$ and $\int_E f^-$ are finite for any integrable function and hence $\int_E f$ is well-defined.

1.13. Definition: If $f : E \subseteq \mathbb{R} \to \mathbb{C}$, we say f is **integrable** if both real and imaginary parts are integrable and we define the **Lebesgue integral** of f over E as

$$\int_E f = \int_E \mathrm{Re}(f) + i \int_E \mathrm{Im}(f).$$

Section 2. Properties of Lebesgue Integral

2.1. All functions and sets in this section are assumed to measurable.

2.2. Proposition: *$\int_E f = \int_{\mathbb{R}} f\chi_E$. As a consequence, there is no loss in assuming $E = \mathbb{R}$ and in this case, we will simply write $\int f$.*

Proof. First note that the measurability of f and the set E also ensures the measurability of the function $f\chi_E$ (product of measurable functions are measurable). The proof now follows a standard line of reasoning:

- holds for positive, simple functions;
- holds for positive, measurable functions;
- holds for real-valued, integrable functions;
- holds for complex-valued, integrable functions.

First, suppose $f = \sum_{k=1}^N a_k \chi_{E_k}$. Then

$$\int_E f = \sum_{k=1}^N a_k m\left(E_k \cap E\right) = \int_{\mathbb{R}} \sum_{k=1}^N a_k \chi_{E_k \cap E} = \int_{\mathbb{R}} \left(\sum_{k=1}^N a_k \chi_{E_k}\right) \chi_E = \int_{\mathbb{R}} f\chi_E.$$

Next, assume $f : \mathbb{R} \to [0, \infty]$. Note that if ϕ is a simple function with $0 \leq \phi \leq f$, then $\phi\chi_E$ is a simple function with $0 \leq \phi\chi_E \leq f\chi_E$. Thus

$$\int_E \phi = \int_{\mathbb{R}} \phi\chi_E \leq \sup\left\{\int_E \psi : 0 \leq \psi \leq f\chi_E\right\} = \int_{\mathbb{R}} f\chi_E.$$

As this is true for all such ϕ, we have

$$\int_E f = \sup\left\{\int_E \phi : 0 \leq \phi \leq f\right\} \leq \int_{\mathbb{R}} f\chi_E.$$

On the other hand, if ψ is a simple function with $0 \leq \psi \leq f\chi_E$, then also $0 \leq \psi \leq f$ and $\psi = \psi\chi_E$. Thus

$$\int_{\mathbb{R}} \psi = \int_{\mathbb{R}} \psi\chi_E = \int_E \psi \leq \sup\left\{\int_E \phi : 0 \leq \phi \leq f\right\} = \int_E f.$$

Therefore

$$\int_{\mathbb{R}} f\chi_E = \sup\left\{\int_{\mathbb{R}} \psi : 0 \leq \psi \leq f\chi_E\right\} \leq \int_E f$$

and hence $\int_{\mathbb{R}} f\chi_E = \int_E f$.

If f is real-valued and integrable over E, then since $(f\chi_E)^{\pm} = f^{\pm}\chi_E$, the previous step implies $\int (f\chi_E)^{\pm} = \int f^{\pm}\chi_E = \int_E f^{\pm}$. It follows that $f\chi_E$ is integrable over $\mathbb{R}$ and

$$\int_{\mathbb{R}} f\chi_E = \int_{\mathbb{R}} f^+\chi_E - \int_{\mathbb{R}} f^-\chi_E = \int_{\mathbb{E}} f^+ - \int_{\mathbb{E}} f^- = \int_E f$$

The argument when f is complex-valued and integrable over E is similar. $\square$

2.3. Proposition (Monotonicity):

- *If* $0 \le f \le g$, *then* $\int_E f \le \int_E g$.
- *If* $|f| < M$, *then* $\int_E |f| \le \int_E M = M \cdot m(E)$.

Proof. We have seen this already. □

2.4. Proposition: *If* $m(E) = 0$, *then* $\int_E f = 0$ *(even if* $f = \infty$ *on* E*).*

Proof. If $f = \sum a_k \chi_{E_k}$ is simple, then

$$\int_E f = \int_E \sum a_k \chi_{E_k} = \sum a_k m\left(E_k \cap E\right) = 0.$$

If $f \ge 0$, then

$$\int_E f = \sup\left\{\int_E \phi : 0 \le \phi \le f\right\} = 0$$

by the first step. The arguments for f real (or complex)-valued and integrable follow easily. □

2.5. Proposition (Homogeneity): *For all scalars* α, $\int \alpha f = \alpha \int f$.

Proof. Omitted. □

2.6. Proposition (Additivity): *For all* ϕ, ψ *simple,* $\int (\phi + \psi) = \int \phi + \int \psi$.

Proof. Omitted. □

2.7. Remark: We will later use this to prove the additivity property of the integral for all integrable functions f, g. Combined with homogeneity, we see the Lebesgue integral is linear.

2.8. Proposition (Triangle inequality): $|\int f| \le \int |f|$.

Proof. Choose α with $|\alpha| = 1$ satisfying $\left|\int f\right| = \alpha \int f$. It follows from homogeneity and the fact that $\left|\int f\right|$ is real-valued that

$$\left|\int f\right| = \int \alpha f = \int \operatorname{Re} \alpha f + i \int \operatorname{Im} \alpha f = \int \operatorname{Re} \alpha f.$$

But Re $\alpha f \le |\alpha f| = |f|$ and thus monotonicity implies

$$\left|\int f\right| = \int \operatorname{Re} \alpha f \le \int |\alpha f| = \int |f|.$$

□

2.9. Proposition (Translation Invariance): $\int_{\mathbb{R}} f(x+y)\,dx = \int_{\mathbb{R}} f(x)\,dx$.

Proof. First, f being measurable implies the function $x \mapsto f(x+y)$ is measurable for each fixed y. Now the translation invariance of Lebesgue measure gives

$$\int \chi_E(x+y)dx = \int \chi_{E-y}(x)dx = m(E-y) = m(E) = \int \chi_E(x)dx.$$

To complete the proof, we argue in the standard fashion, first checking for f simple, which follows very easily from the previous statement, then f positive and finally f integrable. $\square$

Chapter 5

Monotone and Dominated Convergence Theorems

Section 1. Monotone Convergence Theorem

1.1. Motivation: Recall that one of the weaknesses of the Riemann integral was that it was poorly behaved under limits. It is not necessarily true that

$$\lim_{n\to\infty}\left[R-\int f_n\right]=R-\int \lim f_n,$$

even when f_n and $\lim f_n$ are Riemann integrable. Unfortunately, we still don't have this full result for the Lebesgue integral. Here are two examples.

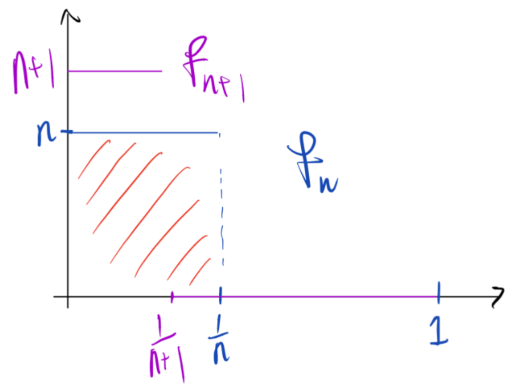

Figure 5.1: f_n in the example below (credit: Bill Zhuo).

1.2. Example: Define f_n on $[0,1]$ by $f_n(x)=n$ for $x\in(0,1/n)$ and $f_n(x)=0$ else. Then f_n is a simple function (a step function, even) and $f_n\to 0$ pointwise. However, $\int_{[0,1]} f_n = 1$ for all n while $\int_{[0,1]}\lim_n f_n=\int 0=0$.

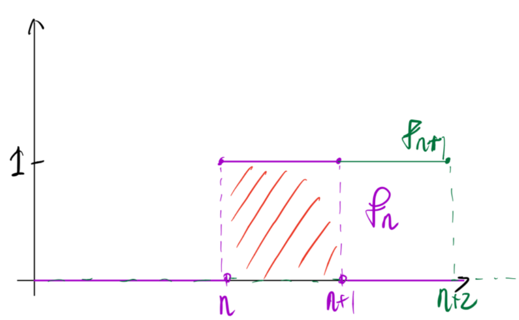

Figure 5.2: f_n in the example below (credit: Bill Zhuo).

1.3. Example: Define f_n on $\mathbb{R}$ by $f_n(x)=\chi_{[n,n+1]}$. Again $f_n\to 0$ pointwise, $\int f_n=1$ and $\int \lim f_n = 0$.

1.4. However, by adding some additional hypotheses, we can prove theorems for the Lebesgue integral that do not necessarily hold for the Riemann integral.

1.5. Theorem (Monotone Convergence Theorem): *Suppose $f_n \geq 0$ and measurable. If $f_n(x) \leq f_{n+1}(x)$ for all x and n, and $f_n \to f$ pointwise, then*

$$\lim_{n\to\infty} \int_E f_n = \int_E \lim_n f_n = \int_E f$$

for any measurable E.

1.6. Remark:

- We allow f_n and f to take on the value ∞ and allow for the possibility of $\lim_n \int_E f_n = \infty$.
- The extra hypothesis here is that the sequence f_n converges ***up*** to f.
- This theorem does not hold for the Riemann integral. Consider the following example.

1.7. Example: Define $f_n(x) = 1$ if $x \in \{r_1, \dots, r_N\} \subseteq \{r_n\}_{n=1}^{\infty} = \mathbb{Q}$ and 0 otherwise. Then $f_n \uparrow \chi_{\mathbb{Q}}$. The functions f_n are all Riemann integrable and $R - \int_0^1 f_n = 0$. But the limit function $\chi_{\mathbb{Q}}$ is not Riemann integrable, so $R - \int \lim f_n$ does not exist. However, using Lebesgue integral, we have seen that $\int_{[0,1]} f_n = 0 = \int_{[0,1]} \chi_{\mathbb{Q}}$.

1.8. Remark: We begin the proof of the theorem with a lemma which shows that the principle idea behind the theorem is **continuity of measure**. Note that this lemma is the special case of MCT where $f_n = \phi\chi_{A_n}$ and $f = \phi\chi_A$.

1.9. Lemma: *Let $\phi \geq 0$ be a simple function and assume $A_1, A_2, \dots$ are measurable with $A_n \subseteq A_{n+1}$ for each n. Define $A = \bigcup_{n=1}^{\infty} A_n$. Then*

$$\lim_{n\to\infty} \int_{A_n} \phi = \int_A \phi.$$

Proof. Assume $\phi = \sum_{i=1}^{N} a_i \chi_{E_i}$ where the sets E_i are measurable. Then

$$\int_{A_n} \phi = \sum_{i=1}^{N} a_i m(E_i \cap A_n).$$

For each i, $E_i \cap A_n \subseteq E_i \cap A_{n+1}$ and also $\bigcup_n (E_i \cap A_n) = E_i \cap A$. By continuity of measure, $m(E_i \cap A_n) \to m(E_i \cap A)$ as $n \to \infty$. Hence,

$$\int_{A_n} \phi \to \sum_{i=1}^{N} a_i m\left(E_i \cap A\right) = \int_A \phi.$$

□

1.10. We are ready to prove MCT: Suppose $f_n \geq 0$ and measurable. If $f_n(x) \leq f_{n+1}(x)$ for all x and n, and $f_n \to f$ pointwise, then

$$\lim_{n\to\infty} \int_E f_n = \int_E \lim_n f_n = \int_E f$$

for any measurable E.

Proof of MCT. Since each f_n is measurable and f is the pointwise limit, f is also measurable. As $f_n \uparrow f$, we have $f_n(x) \leq f(x)$ for all n and x, and thus by monotonicity of the integral, $\int_E f_n \leq \int_E f$. Since the sequence of integrals $(\int_E f_n)_n$ is also increasing, it must have a limit (allowing for the possibility that this limit might be ∞) and hence $\lim_n \int_E f_n \leq \int_E f$.

It remains to prove $\int_E f \leq \lim_n \int_E f_n$. We do this by checking that for every simple function ϕ with $0 \leq \phi \leq f$, we have $\int_E \phi \leq \lim_n \int_E f_n$. We then appealing to the definition of the integral of a non-negative function. Fix such a simple function $\phi \leq f$ and take any $\alpha \in (0,1)$. Define

$$A_n = \{x \in E : f_n(x) \geq \alpha\phi(x)\} = (f_n - \alpha\phi)^{-1}([0,\infty]) \cap E$$

These are measurable sets and since $f_n \leq f_{n+1}$, we have $A_n \subseteq A_{n+1}$ for each n.

Suppose $x \in E$. If $\phi(x) = 0$, then we clearly have $f_n(x) \geq \alpha\phi(x)$ for all n and hence $x \in A_n$ for all n. If $\phi(x) \neq 0$, then $\alpha\phi(x) < \phi(x) \leq f(x)$ as $\alpha < 1$. Since $f_n \to f$, we must have $f_n(x) \geq \alpha\phi(x)$ eventually, i.e., for all $n \geq N$ (the choice of which may depend on x). In any case, $x \in \bigcup A_n$. Since all $A_n \subseteq E$, we have $E = \bigcup A_n$.

The monotonicity property of the integral implies that

$$\alpha \int_{A_n} \phi = \int_{A_n} \alpha\phi \leq \int_{A_n} f_n = \int f_n \chi_{A_n} \leq \int f_n \chi_E = \int_E f_n \to \lim_n \int_E f_n.$$

Appealing to the Lemma, we have

$$\alpha \int_E \phi = \alpha \lim_n \int_{A_n} \phi \leq \lim_n \int_E f_n.$$

But $\int_E f = \sup_{\text{simple } \phi \leq f} \int_E \phi$, thus

$$\alpha \int_E f \leq \lim_n \int_E f_n \text{ for all } \alpha < 1.$$

Letting $\alpha \to 1$ completes the proof. □

1.11. Proposition (Additivity): *If $f, g \geq 0$ are measurable, then $\int_E (f+g) = \int_E f + \int_E g$.*

Proof. We have already seen this result for simple functions. We have also seen that there are positive, simple functions $\phi_n \uparrow f$ and $\psi_n \uparrow g$. Hence, also, $\phi_n + \psi_n \uparrow f + g$. By the MCT we have

$$\int_E (f+g) = \lim_n \int_E (\phi_n + \psi_n) = \lim_n \left(\int_E \phi_n + \int_E \psi_n \right) = \int_E f + \int_E g.$$

□

1.12. Lemma (Fatou): *For $f_n \geq 0$ and measurable, we have*

$$\int \liminf_n f_n \leq \liminf_n \int f_n.$$

Proof. Let $g_n = \inf_{k \geq n} f_k$. Then g_n is measurable and $g_n \leq f_k$ for all $n \leq k$. Thus $\int g_n \leq \int f_k$ for all $n \leq k$ and thus

$$\int g_n \leq \inf_{k \geq n} \int f_k.$$

Moreover, the sequence $(g_n)_n$ is increasing in n, consequently $\lim_n \int g_n$ exists (possibly equal to ∞). Likewise, the sequence $(\inf_{k \geq n} \int f_k)_n$ is increasing in n and hence has a limit. Thus,

$$\lim_n \int g_n \leq \lim_n \left(\inf_{k \geq n} \int f_k \right) = \liminf_n \int f_n$$

Put $F(x) = \liminf_n f_n(x)$. Since $g_n \uparrow F(x)$, MCT gives

$$\lim_n \int g_n = \int F = \int \liminf_n f_n$$

and together these statements give the claimed result. $\square$

1.13. Example: Suppose $f_n \to f$ pointwise and $\int |f_n| \leq 1$ for all n. By Fatou's Lemma,

$$\int |f| = \int \liminf |f_n| \leq \liminf \int |f_n| \leq 1.$$

Thus f is integrable.

Section 2. Dominated Convergence Theorem

2.1. Theorem (Dominated Convergence Theorem): *Suppose f_n are measurable functions with $f_n \to f$ pointwise. Assume there is an integrable function g such that $|f_n(x)| \leq g(x)$ for all x and n. Then*

$$\int f = \lim_n \int f_n.$$

In fact,

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

2.2. Remark: The extra and critical hypothesis is the *existence of the (single) function g which must be integrable and must dominate all the functions $|f_n|$.*

Proof of DCT. We apply Fatou's Lemma to the sequence of functions $2g - |f - f_n|$. Since $|f_n| \leq g$, we also have $|f| \leq g$ and thus $2g - |f - f_n| \geq 0$. It is also measurable being a linear combination of measurable functions. Thus,

$$\begin{aligned}\int 2g &= \int \liminf_n (2g - |f - f_n|) \\ &\leq \liminf_n \int (2g - |f - f_n|) = \int 2g - \limsup_n \int |f - f_n|.\end{aligned}$$

Since g is integrable, $|\int 2g| < \infty$, so we can subtract off $\int 2g$ from both sides to obtain

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

But

$$0 \leq \liminf_n \int |f - f_n| \leq \limsup_n \int |f - f_n| \leq 0$$

so we have equality throughout and, in particular,

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

Since $\int |f_n(x)| \leq \int g(x) < \infty$, each f_n is integrable (so $\int f_n$ is well defined). Furthermore, another application of Fatou's lemma shows that

$$\int |f(x)| \leq \liminf_n \int |f_n(x)| \leq \int g(x) < \infty$$

so also f is integrable and $\int f$ is well defined. Finally, we note that

$$\left| \int f_n - \int f \right| = \left| \int f_n - f \right| \leq \int |f_n - f| \to 0,$$

hence $\int f_n \to \int f$ as desired. $\square$

Section 3. Riemann Integral and Lebesgue Integral

3.1. Motivation: The following theorem proves that the Lebesgue integral generalizes the Riemann integral.

3.2. Theorem: *If f is Riemann integrable over $[a,b]$, then f is Lebesgue integrable over $[a,b]$ and the integrals coincide.*

Proof. Recall that a Riemann integrable function, f, is bounded, say $|f(x)| \le C$ for all x, and the measure of the set of discontinuities is zero. If we call the set of discontinuities E, then $f = f\chi_E + f\chi_{E^c}$. The function $f\chi_E$ is measurable being equal to 0 a.e. The function $f\chi_{E^c}$ is measurable being continuous. Hence f is measurable.

Since $\int_{[a,b]} |f| \le \int_{[a,b]} C \le C(b-a)$, f is an (Lebesgue) integrable function. Take any partition of $[a,b]$, $P : a = x_0 < x_1 < \cdots < x_n = b$, and let

$$M_i = \sup\left\{f(x) : x \in [x_{i-1}, x_i]\right\} \quad \text{and} \quad m_i = \inf\left\{f(x) : x \in [x_{i-1}, x_i]\right\}.$$

Notice that the upper/lower Riemann sums are the Lebesgue integrals of related simple functions:

$$U(f,P) = \sum_{i=1}^{n} M_i\,(x_i - x_{i-1}) = \int_{[a,b]} \sum_{i=1}^{n} M_i \chi_{[x_{i-1},x_i)}$$
$$L(f,P) = \sum_{i=1}^{n} m_i\,(x_i - x_{i-1}) = \int_{[a,b]} \sum_{i=1}^{n} m_i \chi_{[x_{i-1},x_i)}.$$

We have

$$\sum_{i=1}^{n} M_i \chi_{[x_{i-1},x_i)} \ge f \ge \sum_{i=1}^{n} m_i \chi_{[x_{i-1},x_i)}$$

so by monotonicity of the Lebesgue integral

$$\int_{[a,b]} \sum_{i=1}^{n} M_i \chi_{[x_{i-1},x_i)} \ge \int_{[a,b]} f \ge \int_{[a,b]} \sum_{i=1}^{n} m_i \chi_{[x_{i-1},x_i)}$$

Consequently, for all partitions P we have

$$U(f,P) \ge \int_{[a,b]} f \ge L(f,P)$$

By definition of the Riemann integral,

$$\inf_P U(f,P) = \sup_P L(f,P) = R - \int_a^b f.$$

Hence,

$$R - \int_a^b f = \int_{[a,b]} f\ .$$

$\square$