Northwestern University

Lebesgue Integration and Measure Theory

MATH 321-3

size matters

Author:
Elliott Yoon

18 April 2023

1 The Lebesgue Measure

1.1 Desired Properties of the Lebesgue Measure

In our study of measure theory, we wish to find a function (or *measure*) that denotes size of sets, some $\mu(E) \in [0, \infty)$ for all sets $E \in \mathbb{R}$. Let's write down some intuitive axioms:

1. **Normalization of Length.** For an open interval $E = (a, b)$, we want $\mu(E) = b - a$.
2. **Translation Invariance**. First note that for some scalar c and a set A, the set $A + c = \{a + c \mid a \in A\}$. We want $\mu(E) = \mu(E + c)$ for all $c \in \mathbb{R}$.
3. **Countable Additivity** If $E_i \subset \mathbb{R}$, $i \in \mathbb{N}$, then $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu(E_i)$. Moreover, if the E_i's are pairwise disjoint (i.e. $E_i \cap E_j = \varnothing$ for all $i \neq j$), then $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)$.

Unfortunately, no such measure satisfying these properties exists. Rats :/

Fact: It's impossible to define μ satisfying (1)-(3) and defined for all (bounded) $E \subset \mathbb{R}$.

1.2 Null Sets

When working with Riemann integration, there's an often repeated motto that "finite sets don't matter". In the field of measure theory, we want to generalize this statement to be that sets of "generalized length 0", or **measure zero**, don't matter. In fact, we can explore these sets of measure zero without even needing to properly define the Lebesgue measure (though, of course, we will).

In our search for a measure of satisfactory compatibility with the previously proposed "measure axioms" of sorts, we will describe the notion of the **outer measure**, which is defined for all bounded sets of real numbers, satisfies Properties (1) and (2), and satifies the inequality of Property (3), called *subadditivity.* The outer measure fails to be additive (the equality portion of (3)) for certain disjoint sets, so we'll restrict its definition to a large collection of nice (measurable) sets to which additivity holds. What's a measurable set? Let's find out!

Before jumping into some definitions, let's first formalize a notion of length of intervals. We define the length of an open interval $I = (a, b)$ to be $\text{len}(I) = b - a$. Great! We're all set now.

Definition 1.1 (Lebesgue Outer Measure)**.** Suppose $A \subset \mathbb{R}$ is bounded and $\mathscr{U}(A)$ is the set of all *countable* coverings of A by open intervals. We define the **Lebesgue Outer Measure**, $\mu^*(A)$, by

$$\mu^*(A) = \inf_{\{U_n\} \in \mathscr{U}(A)} \left\{ \sum_{i=1}^{n} \text{len}(U_n) \right\},$$

where the infimum is taken over the set of all countable coverings of A by open intervals.

Remark 1.2. It seems silly, but just to be safe, let's note that $\inf\{\infty\} = \infty$.

Example 1.3

- Let $A = (a, b)$. Then $\mu^*(A) = b - a$. (Clearly, $A \subset (a, b)$, so $\mu^*(A) \leq b - a$. Why does $\mu^*(A) \geq b - a$ hold?).
- Let $A = \varnothing$. Then $\varnothing \subset (0, \epsilon)$ for all $\epsilon > 0$, so $\mu^*(A) \leq \inf_\epsilon \text{len}((0, \epsilon)) = \inf_\epsilon \epsilon = 0$.
- Let $A = \{c\}, where c \in \mathbb{R}$. Then $A \subset (c - \epsilon, c + \epsilon)$, so $\mu^*(A) = 0$.
- Let $A = \mathbb{Q}$. Then $\mu^*(A) = 0$. (Why?)

Proposition 1.4

The outer measure of a closed interval is the same as the outer measure of its correspondent open interval. In other words, if $A = [a, b]$, then $\mu^*(A) = b - a$.

Proof. We can encapsulate A inside an open interval: $A \subset (a - \epsilon, b + \epsilon)$, which has length $b - a + 2\epsilon$ for all ϵ. Thus $\mu^*(A) \leq b - a$. Now, note that if $\{U_n\}$ is a cover of A by open intervals, then compactness gives a finite subcover $A \subset \bigcup_{i=1}^n U_i$. Thus, it suffices to show that for any finite cover $\{U_i\}_{i=1}^n$, $\sum_{i=1}^n \text{len}(U_i) \geq b - a$. We'll do so by induction:

The $n = 1$ case is trivial. Now, suppose that for coverings of $n - 1$ intervals, the $(n - 1)$-sum of lengths of the covering open intervals is greater than or equal to $b - a$. Let $A \subset \bigcup_{i=1}^n U_i$. Since A is connected, then if $A \cap U_i$ for all $1 \leq i \leq n$, there are $i \neq j$ such that $U_i \cap U_j \neq \varnothing$. Reordering without loss of generality, assume $i = 1$ and $j = 2$, and let $V = U_1 \cup U_2$ (which is also an open interval). Then $A \subset V \cup \bigcup_{i=3}^n$, which is a union of $n - 1$ open sets, so we're done by the induction hypothesis. ☐

Definition 1.5 (Null Sets)**.** A set $A \subset \mathbb{R}$ is said to be a **null set** provided that $\mu^*(A) = 0$.

Remark 1.6. Null sets can also defined without the machinery of the Lebesgue outer measure as follows: If for all $\epsilon > 0$, there exists a collection of open intervals $\{U_i\}_{i=1}^\infty$ such that

$$\sum_{i=1}^\infty \text{len}(U_i) < \epsilon \quad \text{and} \quad A \subset \bigcup_{i=1}^\infty U_i.$$

then we say A is a null set.

Example 1.7

- $\varnothing$ is a null set.
- Finite sets are null sets.
- The countable collection of null sets $E = \bigcup_{i=1}^\infty E_i \subset \mathbb{R}$ is a null set.
- Countable sets are null sets.
- The Cantor 1/3-set is a null set.

The punchline of the tail end of the previous list of null-set examples is that all null sets are measurable, and for whatever reason, the existence of uncountable null sets implies that describing all measurable sets and functions is, well... complicated.

1.3 σ-algebras

Remark 1.8. Usually, the existence of σ in the nomenclature of an object is to denote that countable operations are allowed.

We're going to now delve into the wonderful mathematical structures called σ-algebras. It turns out that these will be imperative to the study of measurable sets. In fact, as motivation, we shall see that the following holds:

The collection of measurable sets has a structure of a σ-algebra.

First, let's recess quickly for a brief discussion of cardinality: Let X be a set, and write the power set of X as $\mathscr{P}(X) = \{A \subset X\}$. If X is finite and $\text{card}(X) = l$, then $\text{card}(\mathscr{P}(X)) = 2^l$. Instead, if X is countably infinite, then $\text{card}(\mathscr{P}(X))$ is uncountable. (To see why, use a diagonalization argument.)

Definition 1.9 (σ-algebra on X)**.** Suppose X is a set and A is a collection of subsets of X, i.e. $A \subset \mathscr{P}(X)$. A is a *sigma algebra of subsets of* X if

1. $\varnothing, X \in A$,
2. A is closed under complements, and
3. A is closed under countable unions, i.e. if $E_i \subset A$ for $i \in \mathbb{N}$, then $\bigcup_{i=1}^{\infty} E_i \in A$.

Remark 1.10. It's often written as fourth necessary condition that A be closed under countable intersections, but if $E_i \in A$ for $i \in \mathbb{N}$, then

$$\bigcap_{i=1}^{\infty} = \left(\bigcup_{i=1}^{\infty} E_i^C \right)^C \in A,$$

so closure under intersection follows immediately from (2) and (3). Moreover, if $U, V \in A$, then $U \setminus V = U \cap V^C \in A$.

Example 1.11 (Degenerate $\sigma-$algebras)

1. $\mathscr{P}(X)$,
2. $\{\varnothing, X\}$ (called the *trivial* σ*-algebra*)

Example 1.12 (The Null-Conull $\sigma-$algebra)

A more fun (and illuminating) example of a σ-algebra is defined as follows: the set $A \subset \mathscr{P}(\mathbb{R})$ such that $E \in A$ if either E is a null set or E is a null set.

Definition 1.13. Let $\mathscr{F} \subset \mathscr{P}(X)$. The σ-algebra generated by $\mathscr{F}$, written $\sigma(\mathscr{F})$, is the smallest $\sigma-$algebra containing $\mathscr{F}$.

Remark 1.14. Baked into the definition of generated σ-algebras is the guarantee that a $\sigma-$algebra containing $\mathscr{F}$ exists int he first place! (Proven in homework.)

Example 1.15 (The Borel σ-algebra)

Take $\mathscr{F} \subset \mathscr{P}(\mathbb{R})$ to be all open subsets of the real line. $\mathscr{B} \subset \sigma(\mathscr{F})$, the $\sigma-$algebra generated by open sets, is called the **Borel σ-algebra**.

Remark 1.16. Thinking about basic topology of the real line, closure under complements, unions, and intersections means that there are a lot of interesting structures contained in the Borel $\sigma-$algebra. A few of the more interesting ones are as follows:

- Countable unions of closed sets, and
- Countable intersections of open sets.

Indeed,

$$\mathscr{B} = \sigma(\text{open sets}) = \sigma(\text{closed sets}) = \sigma(\text{open intervals}) = \sigma(\text{open intervals of the form } (a, \infty)).$$

Theorem 1.17 (yo this bih kinda slaps)
The $\sigma-$algebra of **Lebesgue-measurable sets** is generated by (1) Borel sets and (2) Null sets.

Zoo wee mama! We don't have sufficient machinery to prove this right now, but it should serve as sufficient motivation for what's to come.

1.4 Properties of the Outer Measure μ^*

So far, we've defined the outer measure μ^* (which isn't a true measure) and checked that $\mu^*([a,b]) = b - a$. At the very beginning, we defined some desired properties of this theoretical notion of a measure, and we'll now explore which of these properties the outer measure has.

Proposition 1.18 (Monotonicity)
If $A \subset B \subset \mathbb{R}$, then $\mu^*(A) \leq \mu^*(B)$.

Proof. Since $A \subset B$, every countable cover of B by open intervals $\{U_n\} \in \mathscr{U}(B)$ also covers A. Thus

$$\inf_{\{U_n\} \in \mathscr{U}(A)} \sum_{i=1}^{\infty} \text{len}(U_n) \leq \inf_{\{U_n\} \in \mathscr{U}(B)} \sum_{i=1}^{\infty} \text{len}(U_n),$$

so $\mu^*(A) \leq \mu^*(B)$. □

We'd previously stated that $\mu^*((a,b)) = b - a$. Let's finish the proof from before:

Proof. Obviously, $\mu^*((a,b)) \leq b - a = \text{len}(a,b)$ since $(a,b) \subset (a,b)$. Moreover, note that $[a+\epsilon, b-\epsilon] \subset (a,b)$ for all sufficiently small $\epsilon > 0$. So $\mu^*((a,b)) \geq \mu^*([a+\epsilon, b-\epsilon] = b - a + 2\epsilon$. □

Corollary 1.19
$\mu^*(\mathbb{R}) = +\infty$ and $\mu^*((a,\infty)) = +\infty$.

Proof. $(a,m) \subset (a,\infty)$ for all $m > a$. Use monotonicity. □

Theorem 1.20 (Translation invariance)
For all subsets $E \subset \mathbb{R}$ and scalars $c \in \mathbb{R}$

$$\mu^*(E) = \mu^*(E + c).$$

Proof. Homework (use intervals). □

Theorem 1.21 (Countable subadditivity)
Given $E_i \subset \mathbb{R}$, $\mu^*\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu^*(E_i)$.

Proof. Fix $\epsilon > 0$. For each i, pick a cover $\{U_n^i\}$ of E_i by open intervals with

$$\sum_{n=1}^{\infty} \text{len}\left(U_n^i\right) - \frac{\epsilon}{2^i} \leq \mu^*(E_i) \leq \sum_{n=1}^{\infty} \text{len}(U_n^i).$$

Let $E = \bigcup_{i=1}^{\infty} E_i$. Now, the set $\{U_n^i \mid i, n \in \mathbb{N}\}$ is a cover of E by countably many open intervals, and

$$\mu^*(E) \leq \sum_{i=1}^{\infty}\left(\sum_{n=1}^{\infty} \text{len}(U_n^i)\right) \leq \sum_i \left(\mu^*(E_i) + \frac{\epsilon}{2^i}\right) = \left(\sum_{i=1}^{\infty} \mu^*(E_i)\right) + \epsilon.$$

□

Remark 1.22. Unlike our desired measure properties, we might not have equality even if all our subsets are pairwise disjoint! (this is really sad)

In fact, there exists $A, B \subset [0,1]$ such that

1. $A \cup B = [0,1]$,
2. $A \cap B = \varnothing$, *but*
3. $\mu^*(A) + \mu^*(B) > 1$.

This defect, of sorts, is why "outer measure" is not a measure.

1.5 A non-measurable set

To concretely illustrate the shortfall of the Lebesgue outer measure, we'll construct a non-measurable set.

Theorem 1.23

There is no $\lambda : \mathscr{P}(\mathbb{R}) \to [0, \infty)$ satisfying

1. λ is **translation invariant,**
2. **monotonicity** holds,
3. $\lambda([0,1]) = 1$ (this can be any non-zero, noninfinite value), and
4. **countable additivity** holds

Remark 1.24. Note that **countable additivity** in (4) can be split into **countable *sub*-additivity** (i.e. $\lambda\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \lambda(E_i)$), and the equality statement:

$$E_i \cap E_j = \varnothing \ \forall i \neq j \Rightarrow \lambda\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \lambda(E_i). \qquad (\star)$$

Moreover, Lebesgue Outer Measure μ^* satisfies (1)-(3) and countable subadditivity (but not the equality statement of (4)).

As hinted before, the obvious punchline of Theorem 1.23 is that we will need to restrict the real line $\mathbb{R}$ to a class of sets we "measure". To prove this theorem, we will "build" a non-measurable set.

First, let's define the following equivalence relation: Given $x, y \in \mathbb{R}$, say $x\ y$ if $x - y \in \mathbb{Q}$. (Feel free to check this yourself if the omission of the proof will keep you up at night.) Then, we'll define the following equivalence class:

$$E_x = \{y \in \mathbb{R} \mid y\ x\}.$$

Note that $x + \frac{k}{107}$ x for all $x \in \mathbb{Z}$, so $E_x \cap [0,1] \neq \varnothing$. For each equivalence class, we will pick a *unique* representative $Z_\alpha \in [0,1]$, where $\alpha \in \Delta$, an uncountable index set.

Definition 1.25 (The "Bad Set")**.** We will define the following set, and later show that it is unmeasurable:

$$B = \{Z_n \mid \alpha \in \Delta\}.$$

Remark 1.26. Note that

1. If $y \in \mathbb{R}$, there exists an index α and rational $q \in \mathbb{Q}$ such that $y = z\alpha + q$,
$$\bigcup_{q \in \mathbb{Q}} B + q = \mathbb{R}.$$
2. If $(B+q) \cap (B+p) \neq \varnothing$ for $p, q \in \mathbb{Q}$, then $p = q$. (This is not entirely obvious, so here's a quick proof: Take $y \in (B+q) \cap (B+p)$. Then there are α, β such that $y = Z_\alpha + q, y = Z_\beta + p$. Thus $Z_\alpha = Z_\beta + p - q$, so Z_α Z_β. Since the representatives in B are unique, $Z_\alpha = Z_\beta$, and thus $p = q$.

We can now prove Theorem 1.23:

Proof. Note that $B \subset [0,1]$. So, $\lambda(B) \leq \lambda([0,1]) \leq 1$. The proof of the theorem is immediate from the following two propositions:

1. If λ satisfies (1)-(3) and countable subadditivity, then $\lambda(B) > 0$. *Proof:* Enumerate $\mathbb{Q} = \{q_i\}$ and write $B_i = B + q_i$ for each $i \in \mathbb{N}$. Since $\mathbb{R} = \bigcup_{i=1}^\infty B_i$,
$$1 \leq \lambda(\mathbb{R}) \leq \sum_{i=1}^\infty \lambda(B_i) \leq \sum_{i=1}^\infty \lambda(B),$$
so $\lambda(B) > 0$.
2. If λ satisfies (1)-(4), then $\lambda([0,2]) = +\infty$. *Proof:* Enumerate $\mathbb{Q} \cap [0,1] = \{q_j\}$, and set $B_j = B + q_j$. Since $B \subset [0,1]$ and $0 \leq q_j \leq 1$, translation is limited and thus $\bigcup_{j=1}^\infty B_j \subset [0,2]$ so
$$\lambda([0,2]) \geq \lambda\left(\bigcup_{j=1}^\infty B_j\right) = \sum_{j=1}^\infty \lambda(B_j) = \sum_{j=1}^\infty \lambda(B) = +\infty.$$

□

Remark 1.27. Observe the following:

1. Our bad set B is non-measurable. If μ is our Lebesgue measure, then $\mu(B)$ is undefined.
2. μ^* satisfies (1)-(3) and countable subadditivity, so $0 \leq \mu^*(B) < 1$.
3. *Claim:* The set $N = [0,1] \setminus B$ is also non-measurable and $\mu^*(N) = 1$. (Think about we're building measurable sets up to have structure similar to σ-algebras.) So $[0,1] = B \cup N$, $B \cap N = \varnothing$, and $\mu^*(B) + \mu^*(N) > 1 = \mu^*(B \cup N)$.

Proposition 1.28 (Outer Regularity)

If $A \subset \mathbb{R}$ is a set with finite outer measure, then for any $\epsilon > 0$, there exists an open set v with

1. $A \subset V$, and
2. $\mu^*(A) \leq \mu^*(V) \leq \mu^*(A) + \epsilon$.

In particular, $\mu^*(A) = \inf\{\mu^*(V) \mid A \subset V, V \text{ open }\}$.

Proof. If $U = \{U_n\}$ is a cover by countably many open intervals with $\sum_{n=1}^{\infty} \text{len } U_n \le \mu^*(A) + \epsilon$. Take $V = \bigcup_{n=1}^{\infty} U_n$. Then $A \subset V$ and $\mu^*(V) \le \sum_{n=1}^{\infty} \mu^*(U_n) = \sum_{n=1}^{\infty} \text{len } U_n \le \mu^*(A) + \epsilon$. □

Zooming out a bit to gain some perspective, we can see that we've found sets A, B such that

$$\mu^*(A) + \mu^*(B) > \mu^*(A \cup B).$$

In particular, we found A, B, where $\mu^*(A \cap [0,1]) + \mu^*(A^C \cap [0,1]) > 1$. We will soon say that $A \subset \mathbb{R}$ is **measurable** if for any $E \subset \mathbb{R}$,

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

1.6 Measurable Sets

Definition 1.29. Let $M_0 \subset \mathscr{P}(\mathbb{R})$. Denote all sets $\mathscr{A}$ with the following property: for any $X \subset \mathbb{R}$,

$$\mu^*(A \cap X) + \mu^*(A^C \cap X) = \mu^*(X). \qquad (\star)$$

For a set $\mathscr{A} \in M_0$, define the **Lebesgue measure** of $\mathscr{A}$ to be $\mu(A) = \mu^*(A)$.

Proposition 1.30

Let $A \subset \mathbb{R}$.

1. $A \in M_0$ (is measurable) if, and only if, A^C is measurable.
2. $A \in M_0$ if, and only if, for all $X \subset \mathbb{R}$, $\mu^*(A \cap X) + \mu^*(A^C \cap X) \le \mu^*(X)$.

Proof. (obvious from definitions) □

1.7 M, the σ−algebra generated by Borel sets and Null sets

Surprise! M_0 is a σ−algebra, $M = M_0$, and μ defined on $M = M_0$ has the desired properties of a measure outlined in the beginning of the chapter.

Definition 1.31. M_0 (which we'll later show to be exactly M) is the σ−algebra of **Lebesgue measurable sets**.

Recall that $\mu^*(A \cap X) + \mu^*(A^C \cap X) \ge \mu^*(X)$. To show M_0 is Lebesgue measurable, it therefore suffices to check $(\star)$ for sets with bounded outer measure. By countable subadditivity, it further suffices to check for only bounded sets; in fact, it's enough to check $(\star)$ when X is an open set or interval. (Shown in week 2 problem set).

Proposition 1.32

IF $A \subset M_0$ is bounded (or even $\mu^*(A) < \infty$), then there exists a Borel set B and a null set $N = A^C \cap B$ such that $A = B \setminus N$.

Proof. For $\epsilon > 0$, there exists an open set V_ϵ with

- $A \subset V_\epsilon$, and
- $\mu^*(A) \le \mu^*(V_\epsilon) \le \mu^*(A) + \epsilon$.

Set $B = \bigcap_{k=1}^{\infty} V_{1/k}$. Then

- B is Borel,
- $A \subset B \subset V_{1/k}$ for all k, and
- $\mu^*(A) \le \mu^*(B) \le \mu^*(A) + \frac{1}{k}$ for all k.

So $\mu^*(A) = \mu^*(B)$. Let $N = B \setminus A$, so $\mu^*(A \cap B) + \mu^*(A^C \cap B) = \mu^*(B)$. Since $\mu^*(A) = \mu^*(B) = \mu^*(A \cap B)$, we have $\mu^*(N) = \mu^*(A^C \cap B) = 0$. □

Proposition 1.33
$A \subset \mathbb{R}$ is null if, and only if, $A \in M_0$ and $\mu(A) = 0$.

Proof. If $A \in M_0$ and $\mu(A) = 0$, then $\mu^*(A) = 0$, so A is null. On the other hand, suppose A is null, so $\mu^*(A) = 0$. Fix $X \subset \mathbb{R}$. Then monotonicity gives

$$\mu^*(A \cap X) + \mu^*(A^C \cap X) \le \mu^*(A) + \mu^*(X) = \mu^*(X).$$

□

Proposition 1.34
If $A, B \in M_0$, then $A \cup B$ and $A \cap B \in M_0$.

Remark 1.35. Since $A \cap B = \left(A^C \cup B^C\right)^C$, it suffices to show $A \cup B \in M_0$.

Proof. Fix $A, B \in M_0$ and pick any $X \subset \mathbb{R}$. Note that

1. $(A \cup B) \cap X = (B \cap X) \cup (A \cap B^C \cap X)$, ad
2. $(A \cup B)^C \cap X = A^C \cap B^C \cap X$.

So

$$\begin{aligned}\mu^*((A \cup B) \cap X) + \mu^*((A \cup B)^C \cap X) &\le \mu^*(B \cap X) + \mu^*(A \cap B^C \cap X) + \mu^*(A^C \cap B^C \cap X)\\ &= \mu^*(B \cap X) + \mu^*(B^C \cap X)\\ &= \mu^*(X).\end{aligned}$$

□

Remark 1.36. If A_i are in M_0, so are $\bigcup_{i=1}^{n} A_i$ and $\bigcap_{i=1}^{n} A_i$. (To prove, induct on n)

Proposition 1.37
All intervals are in M_0.

Remark 1.38. Using complements and finite intersections/unions, we can build any interval from intervals of the form $(-\infty, a], [b, \infty)$.

Example 1.39

$(1, 7] = (-\infty, 7] \cap ((-\infty, 1])^C$.

This finally leads to the following claim: If U is an interval, set $U^- = (-\infty, b) \cap U$ and $U^+ = [b, \infty) \cap U$. Then $\mu^*(U) = \text{len } U$, $\mu^*(U^-) = \text{len } U^-$, and $\mu^*(U^+) = \text{len } U^+$. So by additivity of length,

$$\mu^*(U) = \text{len } U = \mu^*(U^-) + \mu^*(U^+).$$

1.8 TA recitation 1

Definition of outer measure can take open intervals disjoint (uses lemma: any open $U \subset \mathbb{R}$ is countable union of disjoint intervals)

1.9 An equivalence of σ-algebras

An equivalence that we'll end up using naively, going forward is that the $\sigma-$algebra generated by Borel sets and Null sets is exactly the same set as the $\sigma-$algebra of measurable sets. In other words, $M = M_0$.

Proposition 1.40

Intervals are in M_0.

Proof. By the magic of complements, countable unions, and the like, it suffices to show that $[b, \infty) \in M_0$ (and $(-\infty, a] \in M_0$). Since the argument to prove either is the same, we'll proceed by showing $[b, \infty) \in M_0$. Let $A = [a, \infty), X \subset \mathbb{R}$. Fix $\epsilon > 0$ and countably many open intervals N_n such that

$$X \subset \bigcup_{n=1}^{\infty} U_n \quad \text{and} \quad \mu^*(X) \leq \sum_{n=1}^{\infty} \text{len } U_n \leq \mu^*(X) + \epsilon.$$

Set $X^+ = A \cap X$, $X^- = A^C \cap X$, $U_n^+ = U_n \cap X$, $U_n^- = U_n^C \cap X$. Note that U_n^- is an open interval and U_n^+ is a half-open interval, so $\text{len } U_n^+ + \text{len } U_n^- = \text{len } U_n$. So

$$\begin{aligned}
\mu^*(X^+) &\leq \sum_{n=1}^{\infty} \mu^*(U_n^+) = \sum_{n=1}^{\infty} \text{len } U_n^+ \\
\mu^*(X^-) &\leq \sum_{n=1}^{\infty} \mu^*(U_n^-) = \sum_{n=1}^{\infty} \text{len } U_n^-,
\end{aligned}$$

thus

$$\begin{aligned}
\mu^+(A \cap X) + \mu^*(A^C \cap X) &= \mu^*(X^+) + \mu^*(X^-) \\
&\leq \sum_{n=1}^{\infty} \text{len } U_n^+ + \text{len } U_n^- \\
&= \mu^*(X) + \epsilon.
\end{aligned}$$

□

Proposition 1.41 (Finite additivity, of sorts)
Let $A \in M_0$. Fix $B, X \subset \mathbb{R}$. Suppose that $A \cap B = \varnothing$ Then $\mu^*(A \cap X) + \mu^*(B \cap X) = \mu^*((A \cup B) \cap X)$. In particular, if $A, B \in M_0$, then $\mu(A) + \mu(B) = \mu(A \cup B)$.

Proof. Notice that

$$A \cap ((A \cup B) \cap X) = A \cap X,$$

and because $A \cap B = \varnothing$,

$$A^C \cap ((A \cup B) \cap X) = B \cap X.$$

Thus,

$$\begin{aligned}\mu^*((A \cup B) \cap X) &= \mu^*(A \cap ((A \cup B) \cap X)) + \mu^*(A^C \cap ((A \cup B) \cap X) \\ &= \mu^*(A \cap X) + \mu^*(B \cap X).\end{aligned}$$

□

Corollary 1.42
Given countably many $E_i \in M_0$ such that $E_i \cap E_j = \varnothing$ for all $i \neq j$ and $X \subset \mathbb{R}$,

$$\mu^*\left(\bigcup_{i=1}^{n} E_i \cap X\right) = \sum_{i=1}^{n} \mu^*(E_i \cap X).$$

Great, so we now have everything we need to proceed with the main proof of the section!

Theorem 1.43
M_0 is a σ-algebra.

Proof. By definition, if $A \subset M_0$, then $A^C \subset M_0$. Moreover, $\mathbb{R}, \varnothing \in M_0$ since $\mu^*(\mathbb{R} \cap X) + \mu^*(\varnothing \cap X) = \mu^*(X)$. With the first two conditions out of the way, we just need to show that M_0 closed under countable unions. Take countably many $A_i \in M_0$, and let $E_n = \bigcup_{i=1}^{n} A_i$. Set $B_1 = A_1$, and $B_{n+1} = A_{n+1} \setminus E_n$. (This constructs pairwise disjoint sets from $\{E_i\}$.) Then $E_{n+1} = E_n \cup B_{n+1}$ and $E_n \cap B_{n+1} = \varnothing$. Both $E_n, B_n \in M_0$. If $i < j$ and $B_i \subset E_i \subset E_{j-1}$, then $B_i \cap B_j = \varnothing$.

Now, let $E = \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} E_i$. Note that $E^C = \bigcap_{i=1}^{\infty} B_i^C \subset \bigcap_{i=1}^{n} B_i^C = \left(\bigcup_{i=1}^{n} B_i\right)^C = E_n^C$. Fix any $X \subset \mathbb{R}$. Then for *any* n, **Proposition 1.41** and $E_n \in M_0$ gives

$$\begin{aligned}\left(\sum_{i=1}^{n} \mu^*(B_i \cap X)\right) + \mu^*(E^C \cap X) &= \mu^*(E_n \cap X) + \mu^*(E^C \cap X) \\ &\leq \mu^*(E_n \cap X) + \mu^*(E_n^C \cap X) \\ &= \mu^*(X).\end{aligned}$$

So

$$\left(\sum_{i=1}^{\infty} \mu^*(B_i \cap X)\right) + \mu^*(E^C \cap X) \leq \mu^*(X)$$

and thus

$$\begin{aligned}\mu^*(E \cap X) + \mu^*(E^C \cap X) &= \mu^*\left(\bigcup_{i=1}^{\infty} B_i \cap X\right) + \mu^*(E^C \cap X) \\ &\leq \left(\sum_{i=1}^{\infty} \mu^*(B_i \cap X)\right) + \mu^*(E_C \cap X) \\ &\leq \mu^*(X).\end{aligned}$$

So $E = \bigcup_i A_i = \bigcup_i B_i \in M_0$. □

Corollary 1.44

$M_0 = M$.

Proof. First note that both are σ−algebras.

1. Since null sets are in M_0 and open intervals—and *all open sets*—are in M_0, so $M \subset M_0$.
2. Fix $A \in M_0$. Given $n \in \mathbb{Z}$, set $A_[n] = [n, n+1] \cap A$. Since $[n, n+1], A_n \in M_0$, $A_n \in M$. (We've showed before that bounded measurable sets are in M.) In particular, there is a Borel set B_n and Null set N_n such that $A_n = B_n \setminus N_n$. So $A = \bigcup_{n \in \mathbb{Z}} A_n \in M$. So $M_0 \subset M$.

□

It might be helpful to take a step back and list some notes that might be helpful going forward:

- M $(= M_0)$ is the σ−algebra of Lebesgue measurable sets.
- $A \subset \mathbb{R}$ is **measurable** if $A \in M$.
- M is generated by Borel sets and null sets.
- In particular, the following are measurable
  - Intervals,
  - Open sets,
  - Closed sets (like the **Cantor middle 1/3 set**!),
  - Null sets (and thus countable sets).
- Define the Lebesgue measure $\mu : M \to [0, \infty)$, $\mu(A) = \mu^*(A)$. If $A, B \in M$, then

$$\mu(A \cap B) + \mu(A^C \cap B) = \mu(B).$$

Example 1.45

Suppose $A_i, B_i \in M$, $A_1 \subset A_2 \subset A_3 \subset \cdots$, and $B_1 \supset B_2 \supset B_3 \supset \cdots$. Set $A = \bigcup_{i=1}^{\infty} A_i$ and $B = \bigcap_{i=1}^{\infty} B_i$.

1. Show $\mu(A_i) \to \mu(A)$.
   - Set $E_j = A_j \setminus A_{j-1}$. Then E_j are pairwise disjoint and $E_j \in M$, $A = \bigcup_{j=1}^{\infty} E_j$. So
   $$\mu(A_n) \leq \mu(A) = \mu\left(\bigcup_{i=1}^{\infty} E_i\right) < \sum_{i=1}^{\infty} \mu(E_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(E_i) = \lim_{n\to\infty} \mu(A_n).$$
   Note that the last equality holds because the E_i's are measurable.
2. If $\mu(B_1) < \infty$, show $\mu(B_i) \to \mu(B)$.
   - Set $F_j = B_1 \setminus B_j$, so $F_1 \subset F_2 \subset \cdots$. Then
   $$\bigcup_{j=1}^{\infty} F_j = B_1 \setminus B = B^C \text{ (inside } B_1).$$
   By bullet (1) above,
   $$\mu(F_i) \to \mu(B_1 \setminus B) = \mu(B_1) - \mu(B).$$
   Also,
   $$\mu(F_i) = \mu(B_1 \setminus B_i) = \mu(B_1) - \mu(B_i),$$
   so $(\mu(B_1) - \mu(B_i)) \to \mu(B_1) - \mu(B)$.
3. If $\mu(A) < \infty$, show $\mu(A \setminus A_n) = 0$.
   - Set $G_n = A \setminus A_n$. Since $G_1 \supset G_2 \supset \cdots$, notice $\bigcap_{n=1}^{\infty} G_n = \varnothing$. By bullet (2),
   $$\mu(A \setminus A_n) = \mu(G_n) \to \mu(\varnothing) = 0.$$

Remark 1.46. In (2) of the previous example, the assumption that $\mu(B_1 < \infty$ is imperative. To see why, set $B_n = [n, \infty)$. Notice that

1. $\bigcap_{n=1}^{\infty} = \varnothing$, and
2. $\mu(B_n) = \infty$, but $\mu(B) = 0$.

Clearly, this is an issue.

Remark 1.47. Similarly, in (3) of the previous example, the assumption that $\mu(A) < \infty$ is also necessary. Set $A_n = [-n, n]$. Then the union over all n is $\bigcup_{n=1}^{\infty} A_n = \mathbb{R} = A$, but

$$\mu(A \setminus A_n) = \mu((-\infty, -n) \cup (n, \infty)) = \infty \nrightarrow 0.$$

With the results from Example 1.45 under our belt, we can now prove that the Lebesgue Measure actually exists, and is not instead some wacky figment of our horribly rotten mathematical brains:

Theorem 1.48 (Existence of the Lebesgue Measure)
Specifically, there exists a measure (function) $\mu : M \to [0, \infty)$ satisfying:

1. $\mu((a, b)) = b - a$,
2. **translation invariance**, and
3. **countable additivity**.

Moveover, from (1)-(3), we get the following for free:

- **Monotonicity**,
- the null sets are measurable, and
- **outer regularity**: If $A \in M$, then $\mu(A) = \inf\{\mu(U) \mid U \text{ open}, U \supset A\}$.

Proof. We've pretty much proved everything here except countable additivity, so we'll say that it suffices to show (3) holds.

If $A = \bigcup_{i=1}^{\infty} A_i$, then $\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$ by subadditivity of outer measure. Moreover, if A_j pairwise disjoint, then $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \geq \mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$. So $\sum_{i=1}^{\infty} \mu(A_i) \leq \mu\left(\bigcup_{i=1}^{\infty} A_i\right)$. And thus $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$ if A_j are pairwise disjoint. □

Remark 1.49. Note that $\mathbb{R}$ conventionally has measure $\mu(\mathbb{R}) = +\infty$. But sometimes we want to restrict to a total space of finite measure. Arbitrarily, fix $I = [0, 1]$ (any other bounded closed interval will also do!). Then we write

$$M(I) = \{A \cap I \mid A \in M\} \subset M.$$

If we do this, then we're restricting μ to $M(I)$, so $0 \leq \mu(A) \leq 1$ for all $A \in M(I)$. It's important to be aware of context, as one might be working in either $\mathbb{R}$ or I, and it's often up to the reader to figure out the total space when taking complements of sets, and the like.

Example 1.50
Say $A \in M(I)$. Then $A^C = I \setminus A$, and we can do things like

$$\mu(A^C) = \mu(I) - \mu(A) = 1 - \mu(A).$$

Proposition 1.51 (Inner Regularity)
Suppose that $A \in M(I)$. Then for any $\epsilon > 0$, there exists a closed set $C \subset A$ where

$$\mu(A) - \epsilon \leq \mu(C) \leq \mu(A).$$

In particular,

$$\mu(A) = \sup\{\mu(C) \mid C \subset A \text{ is closed}\}.$$

Proof. Pick an open U (in either $\mathbb{R}$ or I) with $A^c \subset U$. Then $A^C = I \setminus A$ and $\mu(U) \leq \mu(A^c) + \epsilon$. Set $C = U^c$ ($= I \setminus U$). Then C is closed, $C = U^c \subset (A^c)^c = A$, and

$$\mu(C) = \mu(U^c) = 1 - \mu(U) \geq 1 - (\mu(A^c) + \epsilon) \geq 1 - \mu(A^c) - \epsilon = \mu(A) - \epsilon.$$

□

1.10 A slightly upsetting freak of nature

Let's restrict our perspective to that of $I = [0, 1]$ and say we have a set $A \subset I$, with $\mu(A) = 2/3$. Now, if we pick a point in A and examine a small neighborhood of this point, what should we expect to be the density (in the non-mathematical sense) of points of A in this neighborhood? If you have more than two brain cells, you'd probably guess 2/3, 66%, whatever. Unfortunately, Lebesgue definitely only had one.

Theorem 1.52

Define

$$f_{A,X}(\delta) = \frac{\mu((x-\delta, x+\delta) \cap A)}{\mu(x-\delta, x+\delta)} = \frac{((x-\delta, x+\delta) \cap A)}{2\delta}.$$

For any *typical* (whatever that means...) $x \in A$, $f_{X,A}(\delta) \to 1$ as $\delta \to 0$.

Proof. Our professor didn't prove this for us, so I won't for you. Because fuck you, that's why. (Did I mention how mad this theorem makes me?) □

Theorem 1.53

Let A be measurable with $0 < \mu(A) < \infty$. Given $0 < p < 1$, there is an open interval U such that $\mu(U \cap A) \geq p \cdot \mu(U)$.

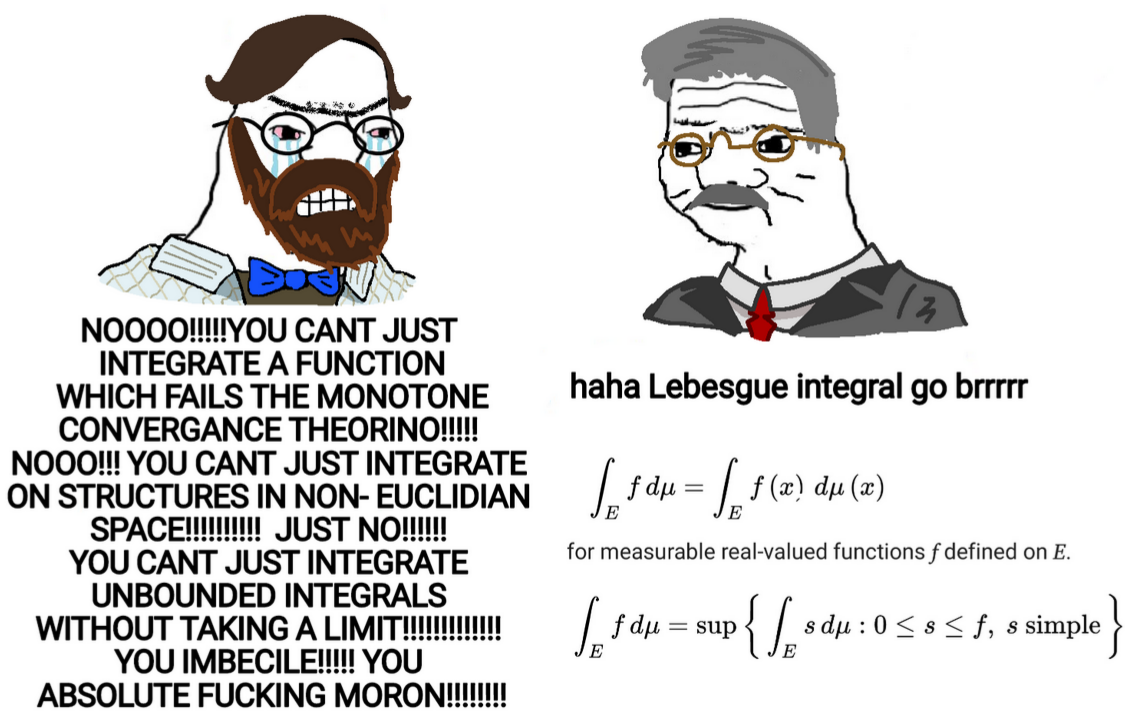
Proof. Yea, this is a proof by contradiction. Enjoy reading through this joke of a proof that's less illuminating than a snuffed out candle. That being said, fix $0 < p < 1$, set $0 < \epsilon = (1-p)\mu(A) < \infty$, and pick disjoint open intervals $\{U_n\}$ such that

$$A \subset \bigcup_{i=1}^{\infty} U_i \quad \text{and} \quad \mu\left(\bigcup_{i=1}^{\infty} Ui\right) \leq \mu(A) + \epsilon.$$

Suppose $\mu(U_n \cap A) < p \cdot \mu(U_n)$ for all n. Then

$$\mu(A) = \mu\left(\bigcup_{n=1}^{\infty} A \cap U_n\right) = \sum_{n=1}^{\infty} \mu(A \cap U_n) < \sum_{n=1}^{\infty} p \cdot \mu(U_n) \leq p(\mu(A) + \epsilon).$$

Then $(1-p)\mu(A) < p \cdot \epsilon$. ϵ is stupidly defined, so $1 < p$. Oh wait—that's bad, isn't it? Whatever. □



2 Measurable Functions

Now that we've properly formulated the Lebesgue Measure, it's time to do stuff with it—integrate, specifically. Whereas Riemann integration uses step functions as its auxiliary approximations, Lebesgue integration will utilize a much broader class of functions called simple functions.

2.1 Simple Functions

Before defining simple functions, we need a few elementary pieces of machinery:

Definition 2.1. Given $A \subset \mathbb{R}$, the **characteristic function** of A is defined by

$$X_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

Definition 2.2. A **finite measurable partition** of $I = [0,1]$ is a collection $\{A_i\}_{i=1}^{l}]$ such that

1. $\bigcup_{i=1}^{l} A_i = I$,
2. $\{A_i\}$ are pairwise disjoint, and
3. $A_i \in M$ (is measurable).

For convenience, we'll allow some $A_i = \varnothing$.

Now that we have the Characteristic function formulated and the notion of a finite measurable partition, we're ready to define the class of simple functions:

Definition 2.3. A function $\delta : [0,1] \to \mathbb{R}$ is **simple** if the following are satisfied:

1. There exists a measurable partition $\{A_i\}|_{i=1}^{l}$ of $[0,1]$, and
2. $r_i \in \mathbb{R}$ for $1 \le i \le l$ such that $f = \sum_{i=1}^{l} r_i X_{A_i}$.

While your mathematical spidey senses tingle, you might be tempted to define an integral over simple functions as your intuition suggests—and for once, you would be correct. If $f = \sum_{i=1}^{l} r_i X_{A_i}$ is simple, we can define

$$\int_0^1 f = \int_I f = \int f \, d\mu := \sum_{i=1}^{l} r_i \mu(A_i),$$

where the $d\mu$ is notational sugar to remind us that we're working with the Lebesgue measure.

Remark 2.4. If $f = \sum_{i=1}^{l} r_i X_{A_i}$ is simple, $f^{-1}((a,\infty)) = \bigcup_{r_i > a} A_i$. So $f^{-1}((a,\infty))$ is measurable for all a.

Remark 2.5. Simple functions form a vector space; i.e. if $f = \sum_{i=1}^{l} r_i X_{A_i}$, $g = \sum_{j=1}^{m} s_j X_{B_j}$ are simple:

- 0 is simple,
- cf is simple for all $c \in \mathbb{R}$, and
- $f + g$ is simple.

Proof. Everything is trivial except for closure under addition; write $C_{ij} = A_i \cap B_j$. Then $\{C_{ij}\}_{i \in \{1,\ldots,l\}}^{j \in \{1,\ldots,m\}}$ is a finite measurable partition of $I = [0,1]$. So

$$(f+g) = \sum_{i=1}^{l} \sum_{j=1}^{m} (r_i + s_j) X_{C_{ij}}.$$

□

Lemma 2.6 (Properties of integration on simple functions)
Let $f, g : I \to \mathbb{R}$ be simple functions. Then

1. Linearity: $\int c_1 f + c_2 g \, d\mu = c_1 \int f \, d\mu + c_2 \int g \, d\mu$,
2. If $f \le g$, then $\int f \, d\mu \le \int g \, d\mu$,
3. $|f|$ is simple and $|\int f \, d\mu| \le \int |f| \, d\mu$.

Proof. Write $f = \sum_{i=1}^{l} r_i X_{A_i}$, $g = \sum_{j=1}^{m} s_j X_{B_j}$, $c_1 f + c_2 g = \sum_{i,j} (c_1 r_i + c_2 g_2) X_{C_{ij}}$, where $C_{ij} = A_i \cap B_j$. Then

1. Notice that $\sum_j \mu(C_{ij} = \mu(A_i)$ and vice versa, so

$$\begin{aligned}
\int c_1 f + c_2 g \, d\mu &= \sum_{i,j} (c_1 r_i + c_2 s_j) \mu(C_{ij}) \\
&= \sum_i \left(\sum_j c_1 r_i \, \mu(C_{ij}) \right) + \sum_j \left(\sum_i c_2 s_j \mu(C_{ij}) \right) \\
&= \sum_i c_1 r_i \left(\sum_j \mu(C_{ij}) \right) + \sum_j c_2 s_j \left(\sum_i \mu(C_{ij}) \right) \\
&= c_1 \sum_i r_i \mu(A_i) + c_2 \sum_j s_j \mu(B_j).
\end{aligned}$$

2. If $f \le g$, then $g - f \ge 0$, so $\int (g-f)\,d\mu \ge 0$. By (1), $\int g - \int f \ge 0$.

3. Notice that $|f| = \sum_{i=1}^{l} |r_i X_{A_i}$. By the triangle inequality,

$$|f\,d\mu| = \left|\sum_{i=1}^{l} r_i \mu(A_i)\right| \le \sum_{i=1}^{l} |r_i|\mu(A_i) = \int |f|\,d\mu.$$

□

2.2 The simple to measurable pipeline goes brazy

An observation of notation: we will notate the **extended real line** as $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$.

Proposition 2.7

Let $X \subset \mathbb{R}$ (or $[0,1]$, it doesn't matter) and $f : X \to \overline{\mathbb{R}}$. The following are equivalent for all $x \in \mathbb{R}$:

1. $f^{-1}([-\infty, a]) \in M$,
2. $f^{-1}([-\infty, a)) \in M$,
3. $f^{-1}([a, \infty]) \in M$,
4. $f^{-1}((a, \infty]) \in M$.

Proof. Assume statement (1) holds for all a. Note that $[-\infty, a) = \bigcup_n [-\infty, a - \frac{1}{n}]$, so

$$f^{-1}([-\infty, a]) = \bigcup_n f^{-1}([-\infty, a - \frac{1}{n}]) \in M.$$

Thus $f^{-1}([-\infty, a]) \in M$. Now, assume (2) holds. Note that $f^{-1}(X^C) = (f^{-1}(x))^C$, so

$$f^{-1}([a, \infty]) = f^{-1}([-\infty, a)^C) = f^{-1}([-\infty, a))^C.$$

We've thus shown that (1) implies (2), which implies (3). The other two implications are identical arguments. □

Definition 2.8. A function $f : X \to \overline{\mathbb{R}}$ is **measurable** if any of the conditions in Proposition 2.7 hold.

> **Remark 2.9.** $f : X \to \mathbb{R}$ is measurable if $f^{-1}((a, \infty))$ is a measurable set for all $a \in \mathbb{R}$.

Proposition 2.10

Let $f, g : [a, b] \to \overline{\mathbb{R}}$ be functions. Let $A \subset [a, b]$ be a null set.

1. Suppose $f(x) = 0$ for all $x \in A$. Then f is measurable.
2. If f measurable and $f(x) = g(x)$ for all $x \notin A$, then g is measurable.

Proof. Fix $a \in \mathbb{R}$. If $a < 0$, then $f^{-1}([-\infty, a]) \subset A$. Since A is a null set, $f^{-1}([-\infty, a])$ is null, hence measurable. If $a \ge 0$, then $f^{-1}([-\infty, a]) = f^{-1}((a, \infty]^C) = (f^{-1}((a, \infty]))^C$. Note $f^{-1}((a, \infty)) \subset A$ so $f^{-1}((a, \infty])$ is a null set. Since M is closed under complements (friendly reminder that it's a σ−algebra!), $f^{-1}([-\infty, a]) \in M$. To show (2), judiciously use the σ−algebra properties of M with the observation that

$$g^{-1}([-\infty, a]) = \left(g^{-1}([-\infty, a]) \cap A\right) \cup \left(g^{-1}([-\infty, a]) \cap A^C\right).$$

□

Remark 2.11. Recall from previous discussions in real analysis that continuity, differentiability, and integrability are *not* preserved under pointwise convergence of functions. However, **pointwise convergence *does* preserve measurability!**

Theorem 2.12

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of measurable functions. The following extended real number valued functions are measurable:

1. $g_1 = \sup_n f_n$,
2. $g_2 = \inf_n f_n$,
3. $g_3 = \limsup_{n\to\infty} f_n$,
4. $g_4 = \liminf_{n\to\infty} f_n$.

Proof. For (1), fix $a \in \mathbb{R}$. If $g_1(x) > a$, then there exists $n \in \mathbb{N}$ such that $f_n(x) > a$. So $g^{-1}([a, \infty]) = \bigcup_n f^{-1}([a, \infty])$. So $g^{-1}([a, \infty]) \in M$. The proof of (2) is identical. Notice now that $\limsup_{n\to\infty} a_n = \inf_N (\sup_{n>N} a_n) = \lim_{N\to\infty} \sup_{n\geq N} a_n$. Then $g_3(x) = \inf_N \sup_{n>N} f_n(x) = \inf_N (h_N(x))$, where h_N is a measurable function, by (1). So g_3 is measurable by (2). The proof of (4) is identical. □

Remark 2.13. We now have that the pointwise limit of a measurable functions is measurable! Furthermore, if $g_1, \ldots, g_l$ are measurable, then $\max\{g_1, \ldots, g_l\}$ is measurable. Moreover, if f measurable, then $|f|$ also measurable.

Theorem 2.14

The set of measurable functions

$$\{f : [0,1] \to \mathbb{R} \quad \text{measurable}\}$$

is a vector space. Furthermore,

1. The set of bounded measurable functions is a sub-vector space of the first set, and
2. If $f, g : [0,1] \to \mathbb{R}$ measurable, so is the pointwise product $fg : [0,1] \to \mathbb{R}$ (defined as $(fg)(x) = f(x)g(x)$).

Proof. The first two statements are trivial. Now, fix $a \in \mathbb{Q}$, and enumerate $\mathbb{Q} = \{q_n\}$. Suppose now that $f(x) + g(x) > a$. Then $f(x) > a - g(x)$. Because $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists a $q_n \in \mathbb{Q}$ such that $f(x) > q_n > a - g(x)$. In other words, $f(x) > q_n$ and $g(x) > a - q_n$. Thus, $x \in f^{-1}((q_n, \infty)) \cap g^{-1}((a - q_n, \infty))$. Set

$$U_a = (f+g)^{-1}((a, \infty)) = \{x \mid f(x) + g(x) > a,$$

so

$$U_a \subset \bigcup_n f^{-1}((q_n, \infty)) \cap g^{-1}((a - q_n, \infty)).$$

On the other hand, if $y \in \bigcup_n f^{-1}((q_n, \infty)) \cap g^{-1}((a - q_n, \infty))$, then there exists $n \in \mathbb{N}$ such that $f(y) > q_n$ and $g(y) > a - q_n$. So $f(y) + g(y) > a$. Thus

$$\bigcup_n f^{-1}((q_n, \infty)) \cap g^{-1}((a - q_n, \infty)) \subset U_a.$$

So $U_a = (f+g)^{-1}(a, \infty)$ is measurable. □

3 Building the theory of integration

We'll delve into integrating by building up its theory on subsets of measurable functions:

1. Bounded measurable functions on $[0,1]$,
2. Non-negative (measurable) functions on $[0,1]$, and finally,
3. General measurable functions on $[0,1]$.

3.1 Integration of bounded measurable functions $f : [0,1] \to \mathbb{R}$.

Theorem 3.1

Let $f : [0,1] \to \mathbb{R}$ be bounded. The following are equivalent:

1. f is measurable.
2. There is a sequence of simple functions $\{g_n\}$ such that g_n converges **uniformly** on $[0,1]$.
3. Let
$$\mathscr{U}_{\mu(f)} = \{u : [0,1] \to \mathbb{R} \mid u \text{ simple}, \ u \geq f\}$$
and
$$\mathscr{L}_\mu(f) = \{v : [0,1] \to \mathbb{R} \mid v \text{ simple}, \ v \leq f\}.$$
Then
$$\sup_{V \in \mathscr{L}(f)} \int v \, d\mu = \inf_{U \in \mathscr{U}} \int U \, d\mu.$$

Proof. This is very long. I'll come back to it □

Definition 3.2. Let $f = [0,1] \to \mathbb{R}$ be bounded and measurable. We define

$$\int f \, d\mu = \sup_{v \in \mathscr{L}(f)} \int v \, d\mu.$$

Proposition 3.3

Let $\{g_n\}$ be simple functions converging uniformly to f on $[0,1]$. (We know f is bounded and measurable.) Then

1. $\lim_{n\to\infty} \int g_n \, d\mu$ exists, and
2. $\lim_{n\to\infty} \int g_n \, d\mu = \int f \, d\mu$.

Proof. The main trick is that $g_n 9x) - \epsilon < f(x) < g_n(x) + \epsilon_n$ for some $\epsilon_n \to 0$. The full proof is in Frank's (Terse) Introduction to Lebesgue Integration, Proposition 3.2.3. □

Summing up what we've already synthesized:

Theorem 3.4 (Nothing New)
Let $f, g : [0,1] \to \mathbb{R}$ be bounded and measurable. Then

1. For all $c_1, c_2 \in \mathbb{R}$,
$$\int c_1 f + c_2 g \, d\mu = c_1 \int f \, d\mu + c_2 \int g \, d\mu.$$
2. If $f \le g$, then $\int f \le \int g$
3. If $|f|$ measurable, then $|\int f| \le \int |f|$
4. If $f(x) = g(x)$ for all x outside a null set, then
$$\int f \, d\mu = \int g \, d\mu.$$

3.2 Bounded Functions

Definition 3.5. Let $f : [0,1] \to \mathbb{R}$ be bounded and measurable and $E \subset [0,1]$ measurable. The we define
$$\int_E f \, d\mu = \int_0^1 \chi_E f \, d\mu.$$

Proposition 3.6
If $E = \bigcup_{i=1}^{\infty} E_i$ is the union of **pairwise disjoint** measurable sets, then
$$\int_E f \, d\mu = \sum_{i=1}^{N} \int_{E_i} f \, d\mu.$$

Proof. Notice $\chi_E f = \sum_{i=1}^{N} \chi_{E_i} f$. □

Proposition 3.7
If $f : [0,1] \to \mathbb{R}$ is a bounded and Riemann-integrable function, then f is Lebesgue measurable and the Riemann integral and Lebesgue integrals coincide.

Proof. Define
$$L_R(f) = \sup\{\int v(x) \, dx \mid v \le f, v \text{ is a step function}\},$$
and
$$L_{\mu(f)} = \sup\{\int v \, d\mu \mid v \le f, v \text{ is a simple function}\},$$
and $U_R(f), U_\mu(f)$ defined similarly with inf's in lieu of sup's. Then
$$L_R(f) \le L_{\mu(f)} \le U_{\mu(f)} \le U_R(f).$$
□

3.3 The Bounded Convergence Theorem

Example 3.8

Let

$$f_n(x) = \begin{cases} n & 0 < x < \frac{1}{n} \\ 0 & \text{otherwise} \end{cases}$$

Notice that $f_n \to 0$ pointwise, but $\int f_n \, d\mu = 1$ and $\int \lim f_n \, d\mu = 0$, so

$$\lim \int f_n \, d\mu = \int \lim f_n$$

need not hold!

Question 3.9. Is there a sufficient condition for pointwise convergence to imply the convergence of integrals?

Let's examine the problem with Example 3.8. Fix $\epsilon > 0$ and examine the "*bad sets*", defined as

$$B_m = \{x \mid |f_n(x) - 0| \geq \epsilon \text{ for some } n \geq m\}$$

. Then

$$\bigcap_{n=0}^{\infty} B_n = \varnothing \quad \text{and} \quad \mu(B_m) \to 0$$

on $B_m = (0, \frac{1}{m})$, so $f_m = m$. To get $\lim \int f_n = \int \lim f_n$, we need to simultaneously control

- measure of "bad sets",
- value of functions f_n

Theorem 3.10 (Bounded Convergence Theorem)

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of measurable functions be defined $f_n : [0,1] \to \mathbb{R}$. Suppose there is $M > 0$ such that $|f_n(x)| \leq M$ for all x and all n, and suppose $f_n \to f$ pointwise. Then

1. f is bounded and measurable, and
2. $\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$.

Proof. For (1), the pointwise limit of measurable function is measurable. Also,

$$\lim_{n\to\infty} |f_n(x)| \leq M$$

so $|f(x)| \leq M$. To prove (2), we want to show $\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$, or $\lim_{n\to\infty} |\int f_n \, d\mu - \int f \, d\mu| = 0$, or $\lim_{n\to\infty} \int |f_n - f| \, d\mu = 0$ since $|\int f_n \, d\mu - \int f \, d\mu| \leq \int |f_n - f| \, d\mu$. Okay, let's actually get into it: fix $\epsilon > 0$. Set

$$E_m = \{x \mid |f_n(x) - f(x)| < \frac{\epsilon}{2} \quad \forall n \geq m\}.$$

Since $f_n \to f$ pointwise,

$$\bigcup_{m=1}^{\infty} E_m = [0,1].$$

Also, $E_m \subset E_{m+1}$, so $\mu(E_m) \to \mu([0,1]) = 1$. Pick m such that

$$\mu(E) > 1 - \frac{\epsilon}{4m}.$$

(Note that $\mu([0,1] \setminus E_m) < \frac{\epsilon}{4m}$.) Then for all $n \geq m$,

$$\begin{aligned}\left|\int f_n \, d\mu - \int f \, \mu\right| &\leq \int |f_n - f| \, d\mu \\ &= \int_{E_m} |f_n - f| + \int_{E_m^C} |f_n - f| \\ &\leq \int_{E_m} \frac{\epsilon}{2} \, d\mu + \int_{E_m^C} 2m \, d\mu \\ &\leq \frac{\epsilon}{2}\mu(E_m) + 2m\mu(E_m^C) \\ &< \frac{\epsilon}{2} + \frac{2m\epsilon}{4m} \\ &= \epsilon.\end{aligned}$$

□

Okay, so what's the punchline of the Lebesgue Integral? We can *always* ignore behavior on a null set!

Example 3.11

Define $f(x) = \begin{cases} 7 & x \notin \mathbb{Q} \\ 0 & x \in \mathbb{Q}, \end{cases}$ and $g(x) = 7$. Then $g(x) = f(x)$ except on a null set $\mathbb{Q}$. Note that $f(x)$ is *not* Riemann integrable.

Definition 3.12. A property holds **almost everywhere** or for **almost every** x if it holds for all x outside a null set.

Example 3.13

Let $f : [0,1] \to \mathbb{R}$, $f(x) = \begin{cases} x & x \notin \mathbb{Q} \\ 1/x & x \in \mathbb{Q} \setminus \{0\} \\ 0 & x = 0 \end{cases}$ f is not bounded, but for *almost* every x, $|f(x)| \leq 2$ so we can say f is **essentially bounded.**

Theorem 3.14

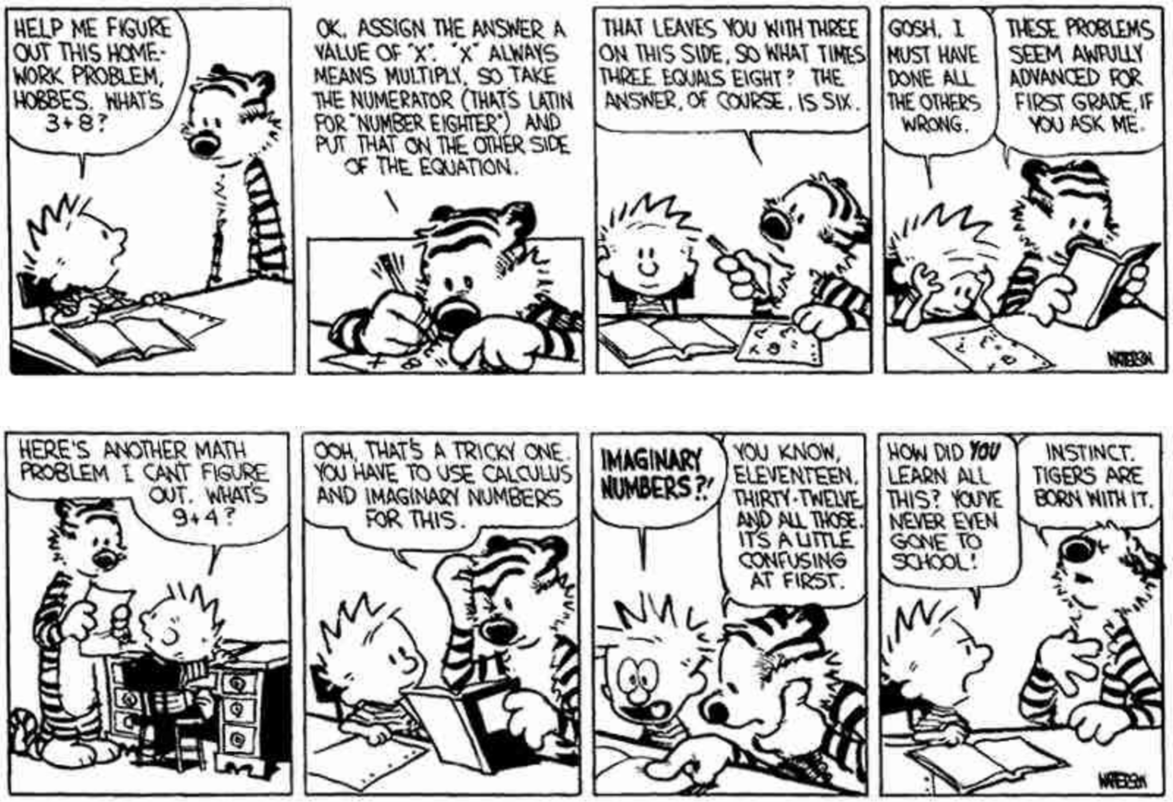
Let $f_n : [0,1] \to \mathbb{R}$ be measurable functions, suppose there exists M such that $|f_n(x)| \leq M$ for all n and almost every x. Suppose $f_n(x) \to f(x)$ for almost every x. Then f is

1. measurable,
2. bounded on a null set, and
3. $\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$.

Proof. Set $A = \{x \mid \lim |f_n(x) - f(x)| \neq 0$, $B = \{x \mid |f_n(x)| \geq M\}$, $E = A \cup \bigcup_{n=1}^{\infty} B_n$, and notice $\mu(E) = 0$. Letting $g_n = \chi_{E^C} f_n$ and $g = \chi_{E^C} f$, notice g_n, g are bounded and measurable, and $g_n(x) \to g(x)$ for all x. So

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

□


HELP ME FIGURE OUT THIS HOME-WORK PROBLEM, HOBBES. WHAT'S 3+8?
OK, ASSIGN THE ANSWER A VALUE OF "X". "X" ALWAYS MEANS MULTIPLY, SO TAKE THE NUMERATOR (THAT'S LATIN FOR "NUMBER EIGHTER") AND PUT THAT ON THE OTHER SIDE OF THE EQUATION.
THAT LEAVES YOU WITH THREE ON THIS SIDE, SO WHAT TIMES THREE EQUALS EIGHT? THE ANSWER, OF COURSE, IS SIX.
GOSH, I MUST HAVE DONE ALL THE OTHERS WRONG.
THESE PROBLEMS SEEM AWFULLY ADVANCED FOR FIRST GRADE, IF YOU ASK ME.
WATTERSON
HERE'S ANOTHER MATH PROBLEM I CAN'T FIGURE OUT. WHAT'S 9+4?
OOH, THAT'S A TRICKY ONE. YOU HAVE TO USE CALCULUS AND IMAGINARY NUMBERS FOR THIS.
IMAGINARY NUMBERS?!
YOU KNOW, ELEVENTEEN, THIRTY-TWELVE AND ALL THOSE. IT'S A LITTLE CONFUSING AT FIRST.
HOW DID YOU LEARN ALL THIS? YOU'VE NEVER EVEN GONE TO SCHOOL!
INSTINCT. TIGERS ARE BORN WITH IT.
WATTERSON

1 Chapter 2, Appendix B

1.1 Properties and construction of the Lebesgue measure on $\mathbb{R}$

1.2 Properties of Lebesgue outer measure

1.3 Definition of Measurable Sets using Outer measure

1.4 Properties of Lebesgue Measure

2 Chapter 3

2.1 Measurable Functions

We wish to define the Lebesgue integral in a fashion similar to that of the Riemann (and the regulated) integral. However, instead of using step functions to approximate a given function, we use a much more general class of functions called *simple functions.*

Definition 2.1. If $A \subset [0,1]$, its **characteristic function** $\mathbb{1}_A(x)$ is defined by

$$\mathbb{1}_A(x) = \begin{cases} 1 & x \in A, \\ 0 & \text{otherwise.} \end{cases}$$

Definition 2.2. A finite **measurable partition** of $[0,1]$ is a collection $\{A_i\}_{i=1}^n$ of measurable subsets which are pairwise disjoint and whose union is $[0,1]$.

Definition 2.3. A function $f : [0,1] \to \mathbb{R}$ is called **Lebesgue simple**, or **simple**, provided there exist a finite measurable partition $A_{i}{}_{i=1}^{n}$ and real numbers $\{r_i\}$ such that

$$f(x) = \sum_{i=1}^{n} r_i \mathbb{1}_{A_i}.$$

The **Lebesgue integral** of a simple function is defined by

$$\int f \, d\mu = \sum_{i=1}^{n} r_i \mu(A_i).$$

Lemma 2.4 (Properties of simple functions)
The set of simple functions is a vector space and the Lebesgue integral of simple functions satisfies the following properties:

1. Linearity: If f, g are simple functions and $c_1, c_2 \in \mathbb{R}$, then
$$\int c_1 f + c_2 g \, d\mu = c \int f \, d\mu + c_2 \int g \, d\mu.$$
2. Monotonicity: If f, g are simple and $f(x) \leq g(x)$ for all x, then
$$\int f \, d\mu \leq \int g \, d\mu.$$
3. Absolute value: If f is simple, then $|f|$ also simple and
$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu.$$

If $f : [0,1] \to \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$ is called an extended real valued function. For $a \in \mathbb{R}$ we denote the set $(-\infty, a] \cup \{-\infty\}$ by $[-\infty, a]$ and the set $[a, \infty) \cup \{\infty\}$ by $[a, \infty]$.

Proposition 2.5
If $f : [0,1] \to \mathbb{R}$ is an extended real valued function, then the following are equivalent:

1. For any $a \in [-\infty, \infty]$, the set $f^{-1}([-\infty, a])$ is Lebesgue measurable.
2. For any $a \in [-\infty, \infty]$, the set $f^{-1}([\infty, a))$ is Lebesgue measurable.
3. For any $a \in [-\infty, \infty]$, the set $f^{-1}([a, \infty])$ is Lebesgue measurable.
4. For any $a \in [-\infty, \infty]$, the set $f^{-1}((a, \infty])$ is Lebesgue measurable.‘

Definition 2.6. An extended real valued function f is called Lebesgue measurable if it satisfies one (and hence all) of the properties of Proposition 2.5.

Proposition 2.7
If $f(x)$ is a function which has the value 0 except on a set of measure 0, then $f(x)$ is measurable.

Theorem 2.8

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of measurable functions. Then the extended real valued functions

$$\begin{aligned} g_1(x) &= \sup_{n\in\mathbb{N}} f_n(x) \\ g_2(x) &= \inf_{n\in\mathbb{N}} f_n(x) \\ g_3(x) &= \limsup_{n\to\infty} f_n(x) \\ g_4(x) &= \liminf_{n\to\infty} f_n(x) \end{aligned}$$

are all measurable.

Theorem 2.9

The set of Lebesgue measurable functions from $[0,1]$ to $\mathbb{R}$ is a vector space. The set of bounded Lebesgue measurable functions is a vector subspace.

2.2 Lebesgue integration of Bounded measurable functions

Theorem 2.10

If $f : [0,1] \to \mathbb{R}$ is a bounded function, then the following are equivalent:

1. The function f is Lebesgue measurable.
2. There is a sequence of simple functions $\{f_n\}_{n=1}^{\infty}$ which converges uniformly to f.
3. If $\mathscr{U}_\mu(f)$ denotes the set of all simple functions $u(x)$ such that $f(x) \le u(x)$ for all x and if $\mathscr{L}_\mu(f)$ denotes the set of all simple functions $v(x)$ such that $v(x) \le f(x)$ for all x, then

$$\sup_{v\in\mathscr{L}_\mu(f)} \left\{\int v\,d\mu\right\} = \inf_{u\in\mathscr{U}_\mu(f)} \left\{\int u\,d\mu\right\}.$$

Definition 2.11. If $f : [0,1] \to \mathbb{R}$ is a bounded measurable function, then we define its Lebesgue integral by

$$\int f\,d\mu = \inf_{u\in\mathscr{U}_\mu(f)} \left\{\int u\,d\mu\right\},$$

or equivalently,

$$\int f\,d\mu = \sup_{v\in\mathscr{L}_\mu(f)} \left\{\int u\,d\mu\right\},$$

Proposition 2.12

If $\{g_n\}_{n=1}^{\infty}$ is any sequence of simple functions converging uniformly to a bounded measurable function f, then $\lim_{n\to\infty} \int g_n\,d\mu$ exists and is equal to $\int f\,d\mu$.

Theorem 2.13

The Lebesgue integral, defined on the vector space of bounded Lebesgue measurable functions on $[0,1]$ satisfies the following properties:

1. Linearity: If f, g are Lebesgue measurable functions and $c_1, c_2 \in \mathbb{R}$, then
$$\int c_1 f + c_2 g \, d\mu = c_1 \int f \, d\mu + c_2 \int g \, d\mu.$$
2. Monotonicity: If f, g are Lebesgue measurable and $f(x) \leq g(x)$ for all x, then $\int f \, d\mu \leq \int g \, d\mu$.
3. Absolute Value: If f is Lebesgue measurable, then $|f|$ is as well and $|\int f \, d\mu| \leq \int |f| \, d\mu$.
4. Null Sets: If f, g are bounded functions and $f(x) = g(x)$ except on set of measure zero, then f is measurable if, and only if, g is measurable. If they are measurable, then $\int f \, d\mu = \int g \, d\mu$.

Definition 2.14. If $E \subset [0,1]$ is a measurable set and f is a bounded mesaurable function we define the Lebesgue integral of f over E by
$$\int_E f \, d\mu = \int f \mathbb{1}_E \, d\mu.$$

Proposition 2.15

If E, F are disjoint measurable subsets of $[0,1]$, then
$$\int_{E \cup F} f \, d\mu = \int_E f \, d\mu + \int_F f \, d\mu.$$

Proposition 2.16 (Riemann integrable functions are Lebesgue integrable.)

Every bounded Riemann integral function $f : [0,1] \to \mathbb{R}$ is measurable and hence Lebesgue integrable. The values of the Riemann and Lebesgue integrals coincide.

2.3 The Bounded Convergence Theorem

Example 2.17

Let
$$f_n(x) = \begin{cases} n & x \in [\frac{1}{n}, \frac{2}{n}], \\ 0 & \text{otherwise} \end{cases}$$

Then f is a step function equal to n on an interval of length $\frac{1}{n}$ and 0 elsewhere. Thus
$$\int f_n \, d\mu = n \frac{1}{n} = 1.$$

But for any $x \in [0,1]$, we have $f_n(x) = 0$ for all sufficiently large n. Thus the sequence $\{f_n\}_1^\infty$ converges pointwise to the 0 function. Hence
$$\int \left(\lim_{n \to \infty} f_n(x) \right) d\mu = 0$$
but
$$\lim_{n \to \infty} \int f_n \, d\mu = 1.$$

Theorem 2.18 (The Bounded Convergence Theorem)
Suppose that $\{f_n\}_1^\infty$ is a sequence of measurable functions which converges pointwise to a function f and there is a constant $M > 0$ such that $|f_n(x)| \le M$ for all n and all $x \in [0,1]$. Then f is a bounded measurable function and

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

Definition 2.19. If a property holds for all x except on a set of measure zero, we say that it holds **almost everywhere** or for **almost all** values of x.

Theorem 2.20 (The *Better* Bounded Convergence Theorem)
Suppose $\{f_n\}_1^\infty$ is a sequence of bounded measurable functions and f is a bounded function such that

$$\lim_{n\to\infty} f_n(x) = f(x)$$

for almost all x. Suppose also that there exists a constant $M > 0$ such that for each $n > 0$,

$$|f_n(x)| \le M$$

for almost all $x \in [0,1]$. Then f is a measurable function, satisfying $|f(x)| \le M$ for almost all $x \in [0,1]$, and

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

3 Chapter 4

3.1 Integration of non-negative functions

Definition 3.1. If $f : [0,1] \to \mathbb{R}$ is a non-negative Lebesgue measurable function, we let $f_n(x) = \min\{f(x), n\}$. Then f_n is a bounded measurable function and we define

$$\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu.$$

If $\int f \, d\mu < \infty$, we say f is integrable.

Proposition 3.2
If f is a non-negative integrable function and $A = \{x \mid f(x) = +\infty\}$, then $\mu(A) = 0$.

Proposition 3.3
Suppose f, g are non-negative measurable functions with $g(x) \le f(x)$ for almost every x. If f is integrable, then g is integrable and

$$\int g \, d\mu \le \int f \, d\mu.$$

In particular, if $g = 0$ almost everywhere, then $\int g \, d\mu = 0$.

Corollary 3.4

If $f : [0,1] \to \mathbb{R}$ is a non-negative integrable function and $\int f \, d\mu = 0$, then $f(x) = 0$ for almost all x.

Theorem 3.5 (Absolute Continuity)

Suppose f is a non-negative integrable function. Then for any $\epsilon > 0$, there exists a $\delta > 0$ such that $\int_A f \, d\mu < \epsilon$ for every measurable $A \subset [0,1]$ with $\mu(A) < \delta$.

Corollary 3.6 (Continuity of the Integral)

If $f : [0,1] \to \mathbb{R}$ is a non-negative integrable function and we define $F(x) = \int_{[0,x]} f \, d\mu$, then $F(x)$ is continuous.

3.2 Convergence Theorems

We can generalize the aforementioned Bounded Convergence Theorem to the following results, where instead of having a constant bound on the functions f_n, we bound them by an integrable function g. (We can do this because of absolute continuity!)

Theorem 3.7 (Lebesgue Convergence for Non-negative functions)

Suppose f_n is a sequence of non-negative measurable functions and g is a non-negative integrable function such that $f_n(x) \le g(x)$ for all n and almost all x. If $\lim f_n(x) = f(x)$ for almost all x, then f is integrable and

$$\int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu.$$

Theorem 3.8 (Fatou's Lemma)

Suppose g_n is a sequence of non-negative mueasruable functions defined on $[0,1]$. If $\lim g_n(x) = f(x)$ for almost all x, then

$$\int f \, d\mu \le \liminf_{n\to\infty} \int g_n \, d\mu.$$

In particular, if $\liminf \int g_n \, d\mu < +\infty$, then f is integrable.

Theorem 3.9 (Monotone Convergence Theorem)

Suppose g_n is an increasing sequence of non-negative measurable functions. If $\lim g_n(x) = f(x)$ for almost all x, then

$$\int f \, d\mu = \lim_{n\to\infty} \int g_n \, d\mu.$$

In particular, f is integrable if, and only if, $\lim \int g_n \, d\mu < +\infty$.

Corollary 3.10 (Integral of infinite series)

Suppose u_n is a non-negative measurable function and f is a non-negative function such that $\sum_{n=1}^{\infty} u_n(x) = f(x)$ for almost all x. Then

$$\int f\,d\mu = \sum_{n=1}^{\infty} u_n\,d\mu.$$

3.3 General Integrable Functions

Recall that a collection $\mathscr{A}$ of subsets of I is called a σ−algebra provided it contains the set I and is closed under taking complements, countable unions, and countable intersections.

Definition 3.11. If $\mathscr{A}$ is a σ−algebra of subsets of I, then a function $v : \mathscr{A} \to \mathbb{R}$ is called a **finite measure** provided

- $v(A) \geq 0$ for every $A \in \mathscr{A}$,
- $v(\varnothing) = 0$, $v(I) < \infty$,, and
- v is countably additive, i.e. if $\{A_n\}_{n=1}^{\infty}$ are pairwise disjoint sets in A, then
$$v\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} v(A_n).$$

Definition 3.12. Let v be a finite measure defined on the σ−algebra $M(I)$. If $f(x) = \sum_{i=1}^{n} r_i \mathbb{1}_{A_i}$ is a simple function then its **integral with respect to** v is defined by

$$\int f\,dv = \sum_{i=1}^{n} r_i v(A_i).$$

If $g : [0,1] \to \mathbb{R}$ is a bounded measurable function, then we define its **integral with respect to** v by

$$\int g\,dv = \inf_{u \in \mathscr{U}_\mu(g)} \left\{\int u\,dv\right\}.$$

If h is a non-negative extended measurable function we define

$$\int h\,dv = \lim_{n\to\infty} \int \min\{h, n\}\,dv.$$

Definition 3.13. If v is a measure defined on $M(I)$, the Lebesgue measurable subsets of I, then we say v is absolutely continuous with respect to Lebesgue measure μ if $\mu(A) = 0$ implies $v(A) = 0$.

Theorem 3.14

If v is a measure defined on $M(I)$ which is absolutely continuous with respect to Lebesgue measure, then for any $\epsilon > 0$, there exists a $\delta > 0$ such that $v(A) < \epsilon$ whenever $\mu(A) = \delta$.

Proposition 3.15

If f is a non-negative integrable function on I and we define

$$v_f(A) = \int_A f\,d\mu,$$

then v_f is a measure with σ−algebra $M(I)$ which is absolutely continuous with respect to Lebesgue measure μ.

Theorem 3.16 (Radon-Nikodym)

If v is a measure with σ−algebra $M(I)$ which is absolutely continuous with respect to Lebesgue measure μ, then there is a non-negative integrable function f on $[0,1]$ such that

$$v(A) = \int_A f\,d\mu.$$

The function f is unique up to measure zero, i.e. if g is another function with these properties, then $f = g$ almost everywhere.

Remark 3.17. The function f is called the *Radon-Nikodym derivate* of v with respect to μ. In fact, the Radon-Nikodym Theorem is more general than stated, as it applies to any two finite measures v and μ defined on a σ−algebra $\mathscr{A}$ with v absolutely continuous with respect to μ.

We will now consider extended measurable functions which may be unbounded both above and below. Define

$$f^+(x) = \max\{f(x), 0\} \text{ and } f^-9x) = -\min\{f(x), 0\}.$$

Definition 3.18. if $f : [0,1] \to \mathbb{R}$ is a measurable function, then we say f is **Lebesgue integrable** provided both f^+ and f^- are integrable (as non-negative functions). If f is integrable, we define

$$\int f\,d\mu = \int f^+\,d\mu - \int f^-\,d\mu.$$

Proposition 3.19

Suppose f, g are measurable functions on $[0,1]$ and $f = g$ almost everywhere. Then if f is integrable, so is g and $\int f\,d\mu = \int g\,d\mu$. In particular, if $f = 0$ almost everywhere $\int f\,d\mu = 0$.

Proposition 3.20

The measurable function $f : [0,1] \to \mathbb{R}$ is integrable if, and only if, the function $|f|$ is integrable.

Theorem 3.21 (Lebesgue Convergence Theorem)

Suppose f_n is a sequence of measurable functions and g is a non-negative integrable function such that $|f_n(x)| \le g(x)$ for all n and almost all x. if $\lim f_n(x) = f(x)$ for almost all x, then f is integrable and

$$\int f\,d\mu = \lim_{n\to\infty} \int f_n\,d\mu.$$

Theorem 3.22

If $f : [0,1] \to \mathbb{R}$ is an integrable function, then given $\epsilon > 0$, there is a step function $g : [0,1] \to \mathbb{R}$ and a measurable subset $A \subset [0,1]$ such that $\mu(A) < \epsilon$ and

$$|f(x) - g(x)| < \epsilon$$

for all $x \notin A$. Moreover, if $|f(x)| \le M$ for all x, then we may choose g with this same bound.

Theorem 3.23

The Lebesgue integral satisfies the following properties:

1. Linearity: If f, g are Lebesgue measurable functions and $c_1, c_2 \in \mathbb{R}$, then
$$\int c_1 f + c_2 g \, d\mu = c_1 \int f \, d\mu + c_2 \int g \, d\mu.$$
2. Montonicity: If f, g are Lebesgue measurable and $f(x) \leq g(x)$ for all x, then $\int f \, d\mu \leq \int g \, d\mu$.
3. Absolute value: If f is Lebesgue measurable, then $|f|$ is also and $|\int f \, d\mu| \leq \int |f| \, d\mu$.
4. Null sets: If f, g bounded functions and $f(x) = g(x)$ except on a set of measure zero, then ff is measurable if and only if g is measurable. If they are measurable, then $\int f \, d\mu = \int g \, d\mu$.

4 Chapter 5

Let X be a finite set with n elements, like $X = \{1, 2, 3, \ldots, n\}$ and we define a measure v on X which is called the *counting measure*. More precisely, we take a $\sigma-$algebra the family of all subsets of X and for any $A \subset X$, we define $v(A)$ to be the number of elements in the set A. Clearly, *any* function $f : X \to \mathbb{R}$ is measurable. Since there is a partition of X given by $A_i = \{i\}$, and f is constant on each A_i (thus $f = \sum_{i=1}^n r_i \mathbb{1}_{A_i}$, $r_i = f(i)$), any function is a simple function. Thus

$$\int f \, dv = \sum_{i=1}^n r_i v(A_i) = \sum_{i=1}^n f(i).$$

We will denote the collection $\{f \mid f : X \to \mathbb{R}\}$ by $L^2(X)$. More formally, there is a vector space isomorphism of $L^2(X)$ and $\mathbb{R}^n$ given by $f \longleftrightarrow (x_1, x_2, \ldots, x_n)$, where $x_i = f(i)$. If $f, g \in L^2(X)$ correspond to the vectors x, y, respectively, then $x_i = f(i), y_i = g(i)$, so

$$< x, y >= \sum_{i=1}^n x_i y_i = \sum_{i=1}^n f(i) g(i) = \int f g \, dv$$

and

$$\|x\|^2 =< x, x >= \sum_{i=1}^n x_i^2 = \sum_{i=1}^n f(i)^2 = \int f^2 \, dv.$$

Definition 4.1. a measurable function $f : [-1, 1] \to \mathbb{R}$ is called **square integrable** if $f(x)^2$ is integrable. We denote the set of all square integrable functions by $L^2[-1, 1]$, and the norm of $f \in L^2[-1, 1]$ by

$$\|f\| = \left(\int f^2 d\,\mu \right)^{1/2}.$$

Proposition 4.2

The norm $\| \, \|$ on $L^2[-1, 1]$ satisfies $\|af\| = |a| \|f\|$ for all $a \in \mathbb{R}$ and all $f \in L^2[-1, 1]$. Moreover, for all f, $\|f\| \geq 0$ with equality only if $f = 0$ almost everywhere.

Lemma 4.3
If $f, g \in L^2[-1,1]$, then fg is integrable and

$$2\int |fg|\,d\mu \le \|f\|^2 + \|g\|^2$$

Equality holds if, and only if, $|f| = |g|$ almost everywhere.

Theorem 4.4
$L^2[-1,1]$ is a vector space.

Theorem 4.5 (Holder Inequality)
If $f, g \in L^2[-1,1]$, then

$$\int |fg|\,d\mu \le \|f\|\|g\|.$$

Equality holds if, and only if, there is a constant c such that $|f(x)| = c|g(x)|$ or $|g(x)| = c|f(x)|$ almost everywhere.

Corollary 4.6
If $f, g \in L^2[-1,1]$, then

$$\left|\int fg\,d\mu\right| \le \|f\|\|g\|.$$

Theorem 4.7 (Minkowski's Inequality)
If $f, g \in L^2[-1,1]$, then

$$\|f+g\| \le \|f\| + \|g\|.$$

Definition 4.8. If $f, g \in L^2[-1,1]$, then we define their **inner product** by

$$\langle f, g\rangle = \int fg\,d\mu.$$

Theorem 4.9
For any $f_1, f_2, g \in L^2[-1,1]$ and any $c_1, c_2 \in \mathbb{R}$, the inner product on $L^2[-1,1]$ satisfies the following properties:

1. Commutativity: $\langle f, g\rangle = \langle g, f\rangle$.
2. Bilinearity: $\langle c_1 f_1 + c_2 f_2, g\rangle = c_1\langle f_1, g\rangle + c_2\langle f_2, g\rangle$.
3. Positive definiteness: $\langle g, g\rangle = \|g\|^2 \ge 0$ with equality if, and only if, $g = 0$ almost everywhere.

4.1 Convergence in L^2

Note that $\operatorname{dist}(f, g) = 0$ if, and only if, $f = g$ almost everywhere, so if we wish to be pedantic, the metric space $L^2[-1,1]$ is really just the equivalence classes of functions which are equal almost everywhere.

Definition 4.10. If $\{f_n\}_{n=1}^{\infty}$ is a sequence in $L^2[-1,1]$, then it is said to converge to in measure of order 2 or to converge in $L^2[-1,1]$ if there is a function $f \in L^2[-1,1]$ such that

$$\lim_{n\to\infty} \|f - f_n\| = 0.$$

Lemma 4.11 (Density of Bounded Functions)

If we define

$$f_n(x) = \begin{cases} n & f(x) > n \\ f(x) & -n \le f(x) \le n \\ -nf(x) < -n \end{cases}$$

then

$$\lim_{n\to\infty} \|f - f_n\| = 0.$$

Proposition 4.12 (Density of Step Functions and Continuous Functions)

The step functions are dense in $L^2[-1,1]$. That is, for any $\epsilon > 0$ and for any $f \in L^2[-1,1]$, there is a step function $g : [-1,1] \to \mathbb{R}$ such that $\|f - g\| < \epsilon$. Likewise, there is a continuous function $h : [-1,1] \to \mathbb{R}$ such that $\|f - h\| < \epsilon$. The function h may be chosen so $h(-1) = h(1)$.

Definition 4.13. An inner product space $(\mathscr{V}, \|,\|)$ which is complete (i.e. in which Cauchy sequences converge) is called a **Hilbert space**.

Theorem 4.14

$L^2[-1,1]$ is a Hilbert space.

4.2 Real Hilbert Space

If $\mathscr{H}$ is a Hilbert spcae and $\{x_n\}$ is a sequence, then $\lim_{n\to\infty} x_n = x$ means that for any $\epsilon > 0$, there is an $N > 0$ such that $\|x - x_n\| < \epsilon$ whenever $n \ge N$. If $\{x_n\}$ is a sequence in $\mathscr{H}$, then $\sum_{m=1}^{\infty} u_m = s$ means that the limit of partial sums $s_n = \sum_{m=1}^{n} u_m$ converges to s. As expected, a series $\sum_{m=1}^{\infty} u_m$ converges absolutely provided that $\sum_{m=1}^{\infty} \|u_m\|$ converges.

Proposition 4.15

If a series in a Hilbert space converges absolutely, then it converges.

We say $x, y \in \mathscr{H}$ is perpendicular (but fr let's just say **orthogonal**) if $\langle x, y\rangle = 0$.

Theorem 4.16 (Pythagorean Theorem)

If $x_1, x_2, \ldots, x_n$ are mutually orthogonal elements of a Hilbert space, then

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2.$$

Definition 4.17. If $\mathscr{H}$ is a Hilbert space, a **bounded linear functional** on $\mathscr{H}$ is a function $L : \mathscr{H} \to \mathbb{R}$ such that for all $v, w \in \mathscr{H}$ and $c_1, c_2 \in \mathbb{R}$, $L(c_1 v + c_2 w) = c_1 L(v) + c_2 L(w)$ and such that there is a constant M satisfying $|L(v)| \leq M\|v\|$ for all $v \in \mathscr{H}$.

Proposition 4.18 (Cauchy-Schwartz Inequality)
If $(\mathscr{H}, \langle\, ,\, \rangle)$ is a Hilbert space and $v, w \in \mathscr{H}$, then

$$|\langle v, w\rangle| \leq \|v\|\|w\|,$$

with equality if, and only if, v and w are multiples of a single vector.

Lemma 4.19
Suppose $\mathscr{H}$ is a Hilbert space and $L : \mathscr{H} \to \mathbb{R}$ is a bounded linear functional which is not identically 0. If $\mathscr{V} = L^{-1}(1)$, then there is a unique $x \in \mathscr{V}$ such that

$$\|x\| = \inf_{v \in \mathscr{V}} \|v\|.$$

That is, there is a unique vector in $\mathscr{V}$ closest to 0. Moreover, the vector x is orthogonal to every element of $L^{-1}(0)$, i.e. if $v \in \mathscr{H}$ and $L(v) = 0$, then $\langle x, v\rangle = 0$.

Theorem 4.20
If $\mathscr{H}$ is a Hilbert space and $L : \mathscr{H} \to \mathbb{R}$ is a bounded linear functional, then there is a unique $x \in \mathscr{H}$ such that $L(v) = \langle v, x\rangle$.

4.3 Abstract Fourier Series

It is not generally possible to find vectors $\{u_n\}_{n=1}^{\infty}$ in a Hilbert space $\mathscr{H}$ such that any $v \in \mathscr{H}$ can be expressed as a *finite* linear combination of the u_n's. Instead we want to write $v \in \mathscr{H}$ as an infinite series

$$v = \sum_{i=1}^{\infty} a_i u_i.$$

Definition 4.21. A family of vectors $\{u_n\}$ in a Hilbert space $\mathscr{H}$ is called **orthonormal** if for each n, $\|u_n\| = 1$ and $\langle u_n, u_m\rangle = 0$ if $n \neq m$.

Theorem 4.22
If $\{u_n\}_{n=0}^{N}$ is a finite orthonormal family of vectors in a Hilbert space $\mathscr{H}$ and $w \in \mathscr{H}$, then the minimum value of

$$\left\| w - \sum_{n=1}^{N} c_n u_n \right\|$$

for all choices of $c_n\mathbb{R}$ occurs when $c_n = \langle w, u_n\rangle$.

Definition 4.23. If $\{u_n\}_{n=0}^{\infty}$ is an orthonormal family of vectors in a Hilbert space $\mathscr{H}$, it is called **complete** if every $w \in \mathscr{H}$ can be written as an infinite series

$$w = \sum_{n=0}^{\infty} c_n u_n$$

for some chloice of the numbers $c_n \in \mathbb{R}$.

Definition 4.24. The n^{th} Fourier coefficient of w with respect to an orthonormal family $\{u_n\}_{n=0}^{\infty}$ is the number $\langle w, u_n \rangle$. The infinite series

$$\sum_{n=0}^{\infty} \langle w, u_n \rangle u_n$$

is called the **Fourier series** of w.

Theorem 4.25 (Bessel's Inequality)

If $\{u_i\}_{i=0}^{\infty}$ is an orthonormal family of vectors in a Hilbert space $\mathscr{H}$ and w is any element of $\mathscr{H}$, then

$$\sum_{i=0}^{\infty} |\langle w, u_i \rangle|^2 \leq \|w\|^2.$$

In particular, this series converges.

Proposition 4.26 (Fourier series converge)

If $\{u_n\}_{n=0}^{\infty}$ is an orthonormal family of vectors in a Hilbert space $\mathscr{H}$ and $w \in \mathscr{H}$, then the Fourier series

$$\sum_{i=0}^{\infty} \langle w, u_i \rangle u_i$$

with respect to $\{u_i\}_{i=0}^{\infty}$ converges. If the orthonormal family is complete, then it converges to w. Morever, it is unique in the sense that if $\sum_{i=0}^{\infty} c_i u_i = w$, then $c_i = \langle w, u_i \rangle$.

Theorem 4.27 (Parseval's Theorem)

If $\{u_n\}_{n=0}^{\infty}$ is an orthonormal family of vectors in a Hilbert space $\mathscr{H}$ and $w \in \mathscr{H}$, then

$$\sum_{i=0}^{\infty} |\langle w, u_i \rangle|^2 = \|w\|^2$$

if, and only if, the Fourier series with respect to $\{u_n\}_{n=0}^{\infty}$ converges to w, i.e.

$$\sum_{i=0}^{\infty} \langle w, u_i \rangle u_i = w.$$

5 Chapter 6

5.1 Pointwise convergence of classical Fourier Series

Definition 5.1. We define the inner product $\langle\, ,\, \rangle$ on $L^2[-\pi, \pi]$, the vector space of square integrable functions on $[-\pi, \pi]$ by

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f g \, d\mu.$$

Theorem 5.2

The family of functions

$$\mathscr{F} = \left\{ \frac{1}{\sqrt{2}}, \cos(nx), \sin(nx) \right\}_{n=1}^{\infty}$$

is an orthonormal family on $L^2[-\pi, \pi]$.

Definition 5.3. If f is an element of $L^2[-\pi, \pi]$, then its classical Fourier coefficients are

$$\begin{aligned} A_0 &= \frac{1}{2\pi} \int f(x)\, d\mu, \\ A_n &= \frac{1}{\pi} \int f(x) \cos(nx)\, d\mu \\ B_n &= \frac{1}{\pi} \int f(x) \sin(nx)\, d\mu \end{aligned}$$

for $n > 0$. The Fourier series of f is

$$A_0 + \sum_{n=1}^{\infty} A_n \cos(nx) + \sum_{n=1}^{\infty} B_n \sin(nx).$$

Remark 5.4. The orthonormal family $\mathscr{F}$ is complete.

We will be particularly interested in the set which is the unit circle in $\mathbb{R}^2$ and which we denote by $\mathbb{T}$. Futhermore, we consider $C(\mathbb{T})$ as the set of continuous functions $h : [-\pi, \pi] \to \mathbb{R}$ which satisfy $h(-\pi) = h(\pi)$.

Theorem 5.5 (Stone-Weierstrass)

Suppose $A \subset C(\mathbb{T})$ is an algebra (vector space with multiplicative closure) satisfying

1. the constant function 1 is in A, and
2. A separates points, i.e. for any distinct θ_0 and θ_1 in $\mathbb{T}$, there is $p \in A$ such that $p(\theta_0) \neq p(\theta_1)$.

Then given any $\epsilon > 0$ and any $f \in C(\mathbb{T})$, there is $p \in A$ such that $|f(\theta) - p(\theta)| < \epsilon$ for all $\theta \in \mathbb{T}$.

What we use from this theorem is that any $f \in C(\mathbb{T})$ can be approximated by a "trigonometric polynomial".

Corollary 5.6

Suppose that g is a continuous function defined on $[-\pi, \pi]$ such that $g(-\pi) = g(\pi)$. If $\epsilon > 0$, then there are $N > 0$ and $a_n, b_n \in \mathbb{R}$, $1 \leq n \leq N$ such that $|g(x) - p(x)| < \epsilon$ for all x, where

$$p(x) = a_0 + \sum_{n=1}^{N} a_n \cos(nx) + \sum_{n=1}^{N} b_n \sin(nx).$$

Theorem 5.7 (Fourier Series converge in L^2)

Suppose that $f \in L^2[-\pi, \pi]$. Then the Fourier series for f with respect to the orthonormal family $\mathscr{F}$ converges to f in $L^2[-\pi, \pi]$. In particular, the orthonormal family $\mathscr{F}$ is complete.

Theorem 5.8 (Carleson's theorem)
Suppose $f \in L^2[-\pi,\pi]$ and

$$A_0 + \sum_{n=1}^{\infty} A_n \cos(nx) + \sum_{n=1}^{\infty} B_n \sin(nx)$$

is its classical Fourier series. Then this series converges to $f(x)$ for almost all values of $x \in [-\pi,\pi]$.

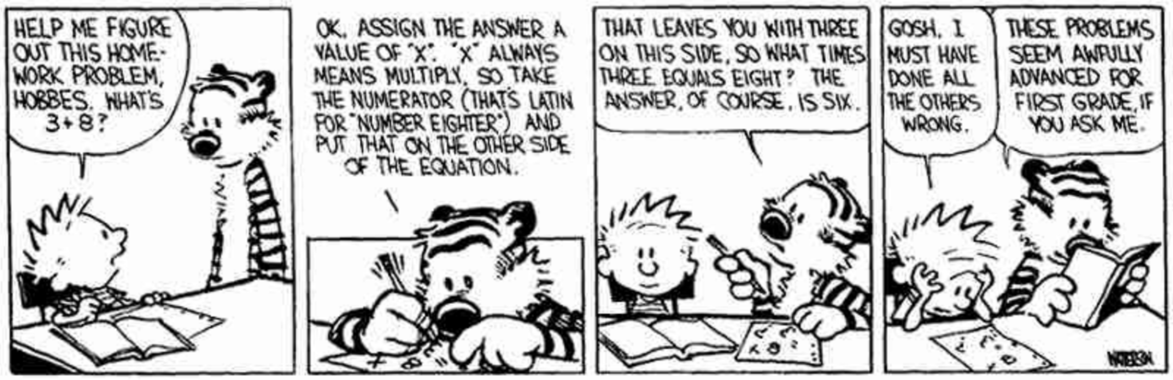
Theorem 5.9
If $f : [-\pi,\pi] \to \mathbb{R}$ is differentiable at $x_0 \in (-\pi,\pi)$, then the Fourier series of f at x_0,

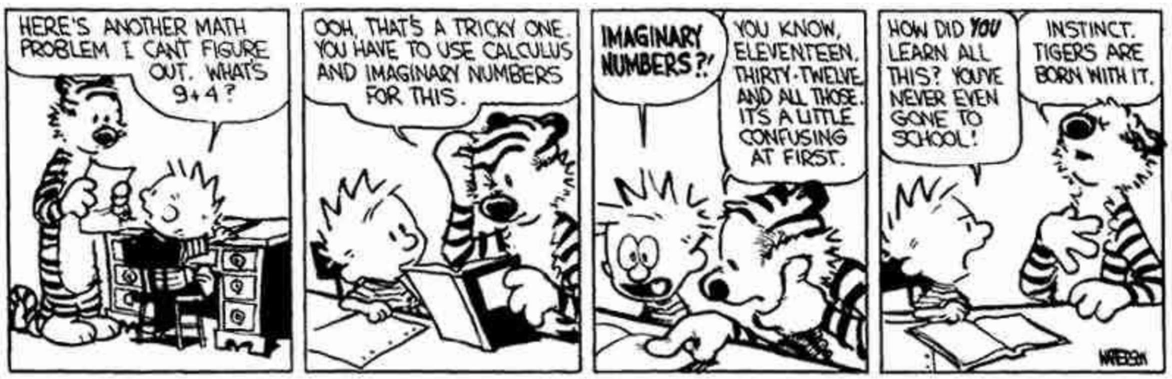
$$A_0 + \sum_{n=1}^{\infty} A_n \cos(nx_0) + \sum_{n=1}^{\infty} B_n \sin(nx_0),$$

converges to $f(x_0)$. IF the right and left derivatives of f exist at $-\pi$ and π respectively, then the Fourier series evaluated at either $-\pi$ or π converges to $\frac{f(-\pi)+f(\pi)}{2}$.

5.2 Using Fourier Coefficients/Series to Evaluate Infinite Series

5.3 Complex Hilbert Space


HELP ME FIGURE OUT THIS HOME-WORK PROBLEM, HOBBES. WHAT'S 3+8?
OK. ASSIGN THE ANSWER A VALUE OF "X". "X" ALWAYS MEANS MULTIPLY, SO TAKE THE NUMERATOR (THAT'S LATIN FOR "NUMBER EIGHTER") AND PUT THAT ON THE OTHER SIDE OF THE EQUATION.
THAT LEAVES YOU WITH THREE ON THIS SIDE, SO WHAT TIMES THREE EQUALS EIGHT? THE ANSWER, OF COURSE, IS SIX.
GOSH, I MUST HAVE DONE ALL THE OTHERS WRONG.
THESE PROBLEMS SEEM AWFULLY ADVANCED FOR FIRST GRADE, IF YOU ASK ME.


HERE'S ANOTHER MATH PROBLEM I CAN'T FIGURE OUT. WHAT'S 9+4?
OOH, THAT'S A TRICKY ONE. YOU HAVE TO USE CALCULUS AND IMAGINARY NUMBERS FOR THIS.
IMAGINARY NUMBERS?!
YOU KNOW, ELEVENTEEN, THIRTY-TWELVE AND ALL THOSE. IT'S A LITTLE CONFUSING AT FIRST.
HOW DID YOU LEARN ALL THIS? YOU'VE NEVER EVEN GONE TO SCHOOL!
INSTINCT. TIGERS ARE BORN WITH IT.