

Measure Theory

Patrick Oare

Resources

These notes are primarily based off of Berkeley's Math 202A, taught by Professor Marc Rieffel. Here are some additional resources that the notes are based on.

- *Real and Functional Analysis* by Serge Lang.
- *Measure Theory* by Paul Halmos.

Introduction

Measure theory seeks to generalize the notion of lengths and volumes to more arbitrary spaces, and to make the notion well defined in \mathbb{R}^n . For example, what do we mean mathematically when we say a statement like “the interval $[0, 1]$ is half as small as $[0, 2]$ ”? Intuitively, the interval $[0, 2]$ has length 2, and the interval $[0, 1]$ has length 1. We can define a length with an integral, but this may not generalize nicely to more abstract spaces. The sets $[0, 1]$ and $[0, 2]$ are topologically equivalent (homeomorphic), so we cannot distinguish them in that sense. Instead, we want to define a new notion of **measure**, which we will define in a general context and then apply to subsets of Euclidean space.

Measure is an invaluable tool in probability theory, and it can actually be used to define a probability on a set. The axioms used to define a σ -ring and a measure are those that are required to have a well defined notion of probability, and one can define a probability space as a space X with measure $\mu(X) = 1$. The concept of a **measurable function** is tied into these definitions as well, because such a function will end up being the measure-theoretic analog of a random variable, i.e. an assignment of a value to an outcome in an abstract space.

We will see that the notion of a measure is intimately connected with the concept of **integration**, and we will see that a measure actually gives one all the tools necessary to define integration. Integrals on \mathbb{R} are defined as limits of Riemann or Darboux sums, which take a form similar to $\sum_j f(x_j)\Delta x_j$. In these sums the term Δx_j is a notion of length, and so it should not be surprising that a measure on a space induces an integral.

1 General Definitions

Given a set X , let \mathcal{M} be a collection of subsets of X . Our goal is to define a way to “add sizes” in the notion of adding intervals of \mathbb{R} . To do this, if we have two subsets $E, F \in \mathcal{M}$, we need their union and their difference to be in \mathcal{M} as well. To formalize this notion, we make a few definitions.

Definition 1: Ring, Field

Let \mathcal{M} be a nonempty collection of subsets of a set X . We say \mathcal{M} is a **ring** if:

1. For $E, F \in \mathcal{M}$, we have $E \cup F \in \mathcal{M}$.
2. For $E, F \in \mathcal{M}$, we have $E \setminus F \in \mathcal{M}$.

If \mathcal{M} is a ring and $X \in \mathcal{M}$, we say that \mathcal{M} is a **field**.

The definitions of a ring and of a field have nothing to do with the algebraic objects of rings and fields; they are completely separate definitions. Also, note that rings are closed with respect to intersections as well, as $E \cap F = E \setminus (E \setminus F) \in \mathcal{M}$. \mathcal{M} will also always contain the empty set \emptyset , and can never itself equal \emptyset by definition (however, we can have $\mathcal{M} = \{\emptyset\}$). Given a set X , we can always construct a field of subsets of X by taking $P(X)$, the power set of X . We can also construct a ring $\mathcal{M} = \{A \subseteq X : A \text{ is finite.}\}$. This will in fact be a field if X is a finite set, as then it will simply be $P(X)$. We can further refine our definitions of a ring and field:

Definition 2: σ -ring, σ -field

Let X be a set, and \mathcal{F} a family of subsets of X . Suppose that \mathcal{F} is a ring. Then we call \mathcal{F} a **σ -ring** if any countable union of elements of \mathcal{F} is in \mathcal{F} . We call \mathcal{F} a **σ -field** if it is a σ -ring and $X \in \mathcal{F}$.

In literature, a σ -field \mathcal{F} is also sometimes called a **σ -algebra**. We will be building up measure theory in terms of σ -fields. Before we do anything, we state a few propositions.

Proposition 1

Let X be a set, and let $\{\mathcal{F}_\alpha\}$ be a collection of rings of X . Then $\bigcap_\alpha \mathcal{F}_\alpha$ is a ring.

Proposition 2

Let X be a set, and A a family of subsets of X . Then there is a smallest ring containing A . This is said to be the ring **generated** by A .

Both of these propositions work equally well if the word “ring” is replaced by any of field, σ -ring, or σ -field. Now, we can create rings out of any collection of sets, so we can ask the question of if the ring generated by a topology has any nice properties.

Definition 3: Borel σ -field

Let (X, τ) be a topological space. Then the σ -field generated by τ is called the **Borel σ -field** for (X, τ) .

Definition 4: Borel σ -ring

Let (X, τ) be locally compact, and \mathfrak{C} be the collection of all compact subsets of X . Then the σ -ring generated by \mathfrak{C} is called the **Borel σ -ring** for (X, τ) .

When we define a measure space, we will need a σ -field, as well as a way to “measure” the subsets in the σ -field. To do this, we must study certain functions into $\mathbb{R}^+ \cup \{\infty\}$ (and later, we will generalize these functions to take values in normed vector spaces; for now, just recognize that we can replace $\mathbb{R}^+ \cup \{\infty\}$ with “normed vector space” in all the following definitions and theorems).

Definition 5

Let X be a set, and A a collection of subsets of X . Let $\mu : A \rightarrow \mathbb{R}^+ \cup \{\infty\}$. We say μ is **additive** if $\forall E, F \in A$ disjoint such that $E \oplus F \in A$, we have:

$$\mu(E \oplus F) = \mu(E) + \mu(F)$$

We say that μ is **finitely additive** if whenever $\{E_i\}_{i=1}^n$ is a finite collection of pairwise disjoint subsets of A , we have:

$$\mu\left(\bigoplus_{i=1}^n E_i\right) = \sum_{i=1}^n \mu(E_i)$$

We say that μ is **countably additive** if μ is finitely additive and whenever $\{E_i\}_{i=1}^\infty$ are pairwise disjoint subsets of A , we have:

$$\mu\left(\bigoplus_{i=1}^\infty E_i\right) = \sum_{i=1}^\infty \mu(E_i)$$

We can now define a measure, which is the main object that we are going to be studying. We will also define a pre-ring and pre-measure, which we will use later to generate measures.

Definition 6: Measure

A **measure** on a set X consists of a σ -ring \mathcal{M} of subsets of X , together with a function $\mu : \mathcal{M} \rightarrow \mathbb{R}^+ \cup \{\infty\}$ such that μ is countably additive. If this holds, we say (X, \mathcal{M}, μ) is a **measure space**.

Definition 7: Pre-ring, Pre-measure

Let X be a set, P a collection of subsets. We say that P is a **pre-ring**, or **semiring**, if:

1. $E, F \in P \implies E \cap F \in P$.
2. $E, F \in P \implies E \setminus F = \bigoplus_{j=1}^n G_j$ for $G_j \in P$.

If P is a pre-ring on X and $\mu : P \rightarrow \mathbb{R}^+ \cup \{\infty\}$ is countably additive, we say that μ is a **pre-measure**.

Note that many people often require \mathcal{M} to be a σ -field instead of a σ -ring, so be careful when reading other texts. An example of a measure is the **counting measure**. This is a measure defined on $P(X)$ for any set X , where we let $\mu(E) = |E|$. For a more complicated example, let $\alpha : \mathbb{R} \rightarrow \mathbb{R}$. Suppose that:

1. α is non-decreasing, i.e. $s \geq t \implies \alpha(s) \geq \alpha(t)$.
2. α is left-continuous, i.e. $\lim_{\epsilon \rightarrow 0, \epsilon > 0} \alpha(t - \epsilon) = \alpha(t)$.

Then let $P := \{[a, b) : a, b \in \mathbb{R}, a < b\}$. The map:

$$\mu([a, b)) = \alpha(b) - \alpha(a)$$

is countably additive as we will see in the next theorem. Furthermore, it is obvious that P is a pre-ring.

Theorem 1

Let α be a non-decreasing and left-continuous function $\alpha : \mathbb{R} \rightarrow \mathbb{R}$. Let $P = \{[a, b) : a, b \in \mathbb{R}, a < b\}$. On P , define μ_α by $\mu_\alpha([a, b)) = \alpha(b) - \alpha(a)$. Then μ_α is countably additive.

Proof. We will fill in the details of this proof later. \square

We will state a proposition that will be an important tool when dealing with semirings. This can be proved with the fact that unions distribute over set differences, and that if P is a semiring with $E, F \in P$, then we have $E \setminus F = \bigoplus_{j=1}^n G_j$ with $G_j \in P$.

Proposition 3

Let P be a semiring of subsets of X . If $F, F_1, \dots, F_n \in P$, then:

$$F \setminus \left(\bigcup_{i=1}^n F_i \right) = \bigoplus_{j=1}^m G_j$$

with $G_j \in P$.

Proof. This follows because $F \setminus (\bigoplus_{i=1}^n F_i) = \bigoplus_{i=1}^n (F \setminus F_i) = \bigoplus_{i=1}^n \bigoplus_{j=1}^{n_i} G_{ij}$, where since P is a semiring, $F \setminus F_i = \bigoplus_{j=1}^{n_i} G_{ij}$. \square

Some additional properties also fall out of the definitions of a measure and a pre-measure. Namely, we will be interested in properties that hold for arbitrary unions, not just disjoint unions. To build up this notion, we need two quick propositions, and some properties of semirings and pre-measures.

Proposition 4

Let P be a semiring, and $\mu : P \rightarrow \mathbb{R} \cup \{\infty\}$ be finitely additive. Then if $E \subseteq \bigoplus_{j=1}^n F_j$ for $E, F_j \in P$, then $\mu(E) \leq \sum_{j=1}^n \mu(F_j)$.

Proposition 5

Let the conditions be as above. If $\bigoplus_{j=1}^n F_j \subseteq E$, then $\sum_{j=1}^n \mu(F_j) \leq \mu(E)$.

Proof. We prove the first proposition. We have:

$$\bigoplus_{j=1}^n F_j = E \oplus \left(\bigoplus_{j=1}^n F_j \setminus E \right) = E \oplus \left(\bigoplus_{k=1}^m G_k \right)$$

Hence:

$$\sum_{j=1}^n \mu(F_j) = \mu(E) + \sum_{k=1}^m \mu(G_k) \geq \mu(E)$$

For the second proposition, write $E = (\bigoplus_{j=1}^n F_j) \oplus (E \setminus \bigoplus_{j=1}^n F_j)$ and apply a similar argument. \square

Definition 8: Monotone

Let \mathcal{F} be a family of subsets of X , and $\mu : \mathcal{F} \rightarrow \mathbb{R}^+ \cup \{\infty\}$. We say μ is **monotone** if whenever $E, F \subseteq \mathcal{F}$ and $E \subseteq F$, then $\mu(E) \leq \mu(F)$.

Corollary 1

Let $E \subseteq F$ be subsets of a semiring P . Then $\mu(E) \leq \mu(F)$. In other words, if (P, μ) is a pre-measure, then μ is monotone.

Definition 9: Countably Sub-Additive

Let $\mu : \mathcal{F} \rightarrow \mathbb{R}^+ \cup \{\infty\}$. We say that μ is **countably sub-additive** if whenever $E \subseteq \bigcup_{j=1}^{\infty} F_j$ with $E, F_j \in \mathcal{F}$ (and F_j possibly empty), then:

$$\mu(E) \leq \sum_{j=1}^{\infty} \mu(F_j)$$

Note that because we allow F_j to be empty, this must hold for finite unions as well. The notion of countable sub-additivity allows us to relax the notion of countable additivity and apply this concept more generally.

Theorem 2

Let (P, μ) be a pre-measure. Then μ is countably-sub additive.

Proof. It suffices to show this for the case that $E = \bigcap_{j=1}^{\infty} F_j$. We **disjointize** the collection $\{F_j\}_{j=1}^{\infty}$. Let $H_1 = F_1$, $H_2 = F_2 \setminus F_1$, and $H_j = F_j \setminus H_{j-1} = F_j \setminus (\bigcup_{k=1}^{j-1} F_k)$. Then the collection $\{H_j\}_{j=1}^{\infty}$ is a disjoint collection of subsets with the same union as $\{F_j\}_{j=1}^{\infty}$, so $E = \bigoplus_{j=1}^{\infty} H_j$. But because P is a semiring and $H_j = F_j \setminus (\bigcup_{k=1}^{j-1} F_k)$, we can write $H_j = \bigoplus_{k=1}^{n_j} G_{kj}$. Note because

μ is monotone, $\mu(H_j) \leq \mu(F_j)$, and we also have $\mu(H_j) = \mu(\bigoplus_{k=1}^{n_j} G_{kj}) = \sum_{k=1}^{n_j} \mu(G_{kj})$. Thus $E = \bigoplus_{j=1}^{\infty} \bigoplus_{k=1}^{n_j} G_{kj}$, so:

$$\mu(E) = \sum_{j=1}^{\infty} \sum_{k=1}^{n_j} \mu(G_{kj}) \leq \sum_{j=1}^{\infty} \mu(F_j)$$

□

2 Constructing Measures

We will now examine how we can construct a measure from a pre-measure. We will need to define the notion of an **outer measure**, and we will show that every pre-measure yields an outer measure, which will yield a measure.

Definition 10

Let \mathcal{F} be a family of subsets of X . Let $A \subseteq X$. We say that A is **countably covered** by \mathcal{F} if there is a sequence $\{F_j\}_{j=1}^{\infty}$ of elements of \mathcal{F} such that $A \subseteq \bigcup_{j=1}^{\infty} F_j$. Let $\mathcal{H}(\mathcal{F})$ be the collection of all subsets of X which are countably covered by \mathcal{F} .

Definition 11: Hereditary

Let \mathcal{H} be a family of subsets of X . We say that \mathcal{H} is **hereditary** if for all $A \in \mathcal{H}$, $B \subseteq A \implies B \in \mathcal{H}$.

Before we move on, we note some properties of $\mathcal{H}(\mathcal{F})$ which we will use frequently.

- $\mathcal{H}(\mathcal{F})$ is a σ -ring.
- $\mathcal{H}(\mathcal{F})$ is hereditary.

We are now in position to define an extension of any a pre-measure. Let $\mu : \mathcal{F} \rightarrow \mathbb{R}^+ \cup \{\infty\}$ be any map. Let $A \in \mathcal{H}(\mathcal{F})$. We define $\mu^* : \mathcal{H}(\mathcal{F}) \rightarrow \mathbb{R}^+ \cup \{\infty\}$ by:

$$\mu^*(A) := \inf \left\{ \sum_{j=1}^{\infty} \mu(F_j) : A \subseteq \bigcup_{j=1}^{\infty} F_j, F_j \in \mathcal{F} \right\}$$

Note that because $A \in \mathcal{H}(\mathcal{F})$, A is countably covered by at least one subset $\{F_j\}_{j=1}^{\infty}$, and so this set is non-empty and the infimum exists. The function μ^* essentially assigns A the size of the smallest cover containing it.

Definition 12: Outer Measure

Let \mathcal{H} be a hereditary σ -ring of subsets of X , and let $\nu : \mathcal{H} \rightarrow \mathbb{R}^+ \cup \{\infty\}$. We say that ν is an **outer measure** if:

1. $\nu(\emptyset) = 0$.
2. ν is monotone.

3. ν is countably subadditive.

Theorem 3

For μ^* defined on $\mathcal{H}(\mathcal{F})$ as above, μ^* is an outer measure.

Proof. To show μ^* is monotone, let $A \subseteq B \in \mathcal{H}(\mathcal{F})$. Then for every countable cover $\{F_j\}_{j=1}^\infty$ of B , this also covers A , and so the $\mu^*(A) \leq \mu^*(B)$. To show countable subadditivity, let $A \subseteq \bigcup_{j=1}^\infty B_j$ with $A, B_j \in \mathcal{H}(\mathcal{F})$. Let $\epsilon > 0$, and choose $\{\epsilon_j\}_{j=1}^\infty$ such that $\epsilon_j > 0$ and $\sum_{j=1}^\infty \epsilon_j = \epsilon$. For each j , choose a countable cover $\{F_{jk}\}_{k=1}^\infty \subseteq \mathcal{F}$ such that $\mu(F_{jk}) < \mu(B_j) + \epsilon_j$. Then $\{F_{jk}\}_{j,k=1}^\infty$ is a countable cover of A , and $\mu^*(A) \leq \sum_{j,k} \mu(F_{jk}) < \sum_j \mu(B_j) + \epsilon$. Since $\epsilon > 0$ is arbitrary, $\mu^*(A) \leq \sum_j \mu(B_j)$. \square

Theorem 4

Let (P, μ) be a pre-measure. Then μ^* defined as above agrees with μ on P , i.e. $\mu^*|_P = \mu$.

Proof. Let $E \in P$. Then a countable cover of E is $\{E\}$, so $\mu^*(E) \leq \mu(E)$. To show these are equal, let $\{F_j\}_{j \in \mathbb{N}}$ be a countable cover of E . μ^* is countably subadditive, so we necessarily have that $\mu(E) \leq \sum_{j \in \mathbb{N}} \mu(E_j)$, and because this holds for each cover $\{E_j\}_{j \in \mathbb{N}}$ we have $\mu(E) \leq \mu^*(E)$. \square

So, this is why we are interested in outer measures. Given a pre-measure, we can construct a corresponding outer measure on $\mathcal{H}(P)$ which extends μ . Now, we will show that we can restrict the domain of an outer measure to a type of sets we call “measurable” to turn the outer measure into a full measure.

Definition 13: ν -measurable

Let ν be an outer measure on a hereditary σ -ring \mathcal{H} . We say that a set $E \in \mathcal{H}$ is ν -**measurable** if $\forall A \in \mathcal{H}$, we have:

$$\nu(A) = \nu(A \cap E) + \nu(A \setminus E)$$

We denote the set of all ν -measurable subsets of \mathcal{H} by $M(\nu)$.

Theorem 5

Let ν be an outer measure on \mathcal{H} . Then $M(\nu)$ is a σ -ring, and ν restricted to $M(\nu)$ is a measure.

Proof. TODO. \square

Proposition 6

Each element of P is μ^* measurable, i.e. $P \subseteq M(\mu^*)$.

We have showed what we wanted to show: given a pre-measure, we now have a natural construction to create a measure which naturally extends the pre-measure! In practice, it is much easier to specify a pre-measure and then construct the corresponding measure. It is usually easier to specify all the sets in a semiring than to specify the sets in a σ -ring, and so we will usually define measures by defining their action on a semiring. For the rest of this section, if we refer to μ^* , we assume it was constructed from a premeasure P .

It is often difficult to determine which sets are ν measurable for a measure ν . However, the following proposition is very helpful.

Proposition 7

Let (\mathcal{H}, ν) be any outer measure. If $E \in \mathcal{H}$ and $\nu(E) = 0$, then E is ν -measurable.

Proof. Suppose that $\nu(E) = 0$, and $A \in \mathcal{H}$. Then $\nu(A \setminus E) + \nu(A \cap E) = \nu(A \setminus E) + 0 = \nu(A \setminus E) + \nu(E) = \nu(A)$ because $\nu(A \cap E) \leq \nu(E) = 0$. \square

Definition 14: Complete, null set

A measure (S, μ) is **complete** if whenever $E \in S$ and $\mu(E) = 0$, then for all $A \subseteq E$ we have $A \in S$. Sets in S with $\mu(E) = 0$ are called **null sets**.

Note that for a complete measure, null sets form a hereditary σ -ring. By the previous proposition, any outer measure (\mathcal{H}, ν) with the corresponding measure on $M(\nu)$ is complete. However, for a premeasure (P, μ) , the extended measure μ^* on $S(P)$ (the σ -ring generated by P) may not be complete because $S(P) \subseteq M(\mu^*)$. To ensure the measure μ^* is complete, we need it to be defined on the entire $M(\mu^*)$. Now, we will focus on the uniqueness of extended measures.

Definition 15: σ -finite

Let \mathcal{F} be a family of subsets of X , and $\nu : \mathcal{F} \rightarrow \mathbb{R}^+ \cup \{\infty\}$. We say that a set $A \in \mathcal{F}$ is **σ -finite** if there is $\{B_j\}_{j=1}^\infty \subseteq \mathcal{F}$ such that $A \subseteq \bigcup_{j=1}^\infty B_j$ and $\nu(B_j) < \infty$ for each j . We say (\mathcal{F}, ν) is **σ -finite** if every $A \in \mathcal{F}$ is σ -finite. If $X \in \mathcal{F}$ and X is σ -finite, we say that (\mathcal{F}, ν) is **totally σ -finite**.

The σ -finite sets of a measure are those which are covered by a countable union of sets of finite measure, and these definitions will have much to do with how we can extend measures uniquely.

Proposition 8

Let (P, μ) be a premeasure. Construct $\mathcal{H}(P)$, μ^* , and $M(\mu^*)$. Let S be any σ -ring with $P \subseteq S \subseteq M(\mu^*)$, and let $\nu : S \rightarrow \mathbb{R}^+ \cup \{\infty\}$ be countably subadditive and monotone such that $\nu|_P = \mu$. Then $\nu(A) \leq \mu^*(A)$ for each $A \in S$.

Proof. Let $A \in S$. Note that $S \subseteq M(\mu^*) \subseteq \mathcal{H}(P)$, so we can find $\{E_j\}_{j \in \mathbb{N}}$ such that $A \subseteq \bigcup_{j \in \mathbb{N}} E_j$ and $\nu(A) \subseteq \nu(\bigcup_{j \in \mathbb{N}} E_j) \leq \sum_{j=1}^\infty \nu(E_j) = \sum_{j=1}^\infty \mu(E_j)$, and because $\mu^*(A)$ is defined to be the infimum over all covers, we have $\nu(A) \leq \mu^*(A)$. \square

Theorem 6

Let (P, μ) be σ -finite. Then for any S with $P \subseteq S \subseteq M(\mu^*)$ and any measure ν on S such that $\nu|_P = \mu$, we have $\nu = \mu^*|_S$.

Proof. TODO. \square

So, if our premeasure is σ -finite, then there is only one unique extension μ^* of μ to any σ -ring containing P . We will now define an important measure on \mathbb{R} .

Definition 16: Lebesgue Measure

Let $\alpha : \mathbb{R} \rightarrow \mathbb{R}$, $\alpha(r) = r$ with premeasure $\mu_\alpha([a, b)) = b - a$. Let $\mu := \mu_\alpha^*|_{M(\mu_\alpha^*)}$. Then μ on $M(\mu_\alpha^*)$ is called the **Lebesgue measure** on \mathbb{R} .

The Lebesgue measure on \mathbb{R} is the canonical example of what one thinks of as a measure. Intuitively, a measure is a way to define a length or volume on a set. In \mathbb{R} , the Lebesgue measure does precisely that. One important property of the Lebesgue measure is that it is **translation invariant**, i.e. if A is Lebesgue-measurable and we define its translate $A + r_0 := \{x + r_0 : x \in A\}$ for $r_0 \in \mathbb{R}$, then $\mu(A) = \mu(A + r_0)$. There are also different variants of the Lebesgue measure. If we restrict μ to Borel subsets of \mathbb{R} , we get the **Lebesgue-Borel measure**. Furthermore, for any left continuous non-decreasing function α on \mathbb{R} , the corresponding measure μ_α^* on $M(\mu_\alpha^*)$ is called the **Lebesgue-Stieltjes measure** on \mathbb{R} .

Finally, we end this section with a few comments on sequences of sets. Suppose we have a sequence of sets $\{E_j\}_{j=1}^\infty$. If $E_{j+1} \subseteq E_j$ is a decreasing tower, we define $E := \bigcap_{j=1}^\infty E_j$ and write $E_j \downarrow E$. If $E_j \subseteq E_{j+1}$ is an increasing tower, then we define $E := \bigcup_{j=1}^\infty E_j$ and write $E_j \uparrow E$.

Proposition 9

Let (S, μ) be a measure. If $\{E_j\}_{j=1}^\infty \subseteq S$ such that $E_j \uparrow E$, then $\mu(E_j) \uparrow \mu(E)$, i.e. $\lim \mu(E_j) = \mu(E)$ and $\mu(E_j) \leq \mu(E)$ for each j .

On the other hand, if $E_j \downarrow E$, then it is not necessarily true that $\lim \mu(E_j) = \mu(E)$. Consider $X = \mathbb{R}$ equipped with the Lebesgue measure, and $E_j = [j, \infty)$. Then $\mu(E_j) = \infty$ and $\bigcap_{j=1}^\infty E_j = \emptyset$, so $\mu(E) = 0 \neq \lim \mu(E_j)$. However, if for at least 1 j we have $\mu(E_j) < \infty$, then $E_j \downarrow E$ will imply $\mu(E) = \lim \mu(E_j)$.

3 Integration of Simple Functions

We will discuss integration on arbitrary measure spaces. Let (X, S, μ) be a measure space, and B a Banach space. Up to now, we have generally been using B to be the real numbers, but we can generalize our constructions in the previous two sections to allow measures to take values in an arbitrary B . We will integrate B -valued functions. Some examples of Banach spaces include $B(X, \mathbb{R})$, $C_b(X)$, and $C(X)$ with X compact.

Definition 17: Simple Measurable Function

A function $f : X \rightarrow B$ is a **simple S-measurable function** (SMF) if the range of f is finite, and for every $b \in \text{im}(f)$, $b \neq 0$, we have $\{x : f(x) = b\} \in S$. Then f can be written:

$$f = \sum_{j=1}^n b_j \chi_{E_j}$$

with the b_j 's distinct and the E_j 's disjoint and measurable (or $f \equiv 0$ everywhere, but this is an edge case).

Proposition 10

If f, g are SMFs, then $f + g$ is a SMF.

Definition 18: Simple Integrable Function

We say f is a **simple integrable function (SIF)** if $f = \sum_{j=1}^n b_j \chi_{E_j}$ is a SMF with $\mu(E_j) < \infty$ for each j . If f is a SIF, we define:

$$\int f d\mu := \sum_{j=1}^n b_j \mu(E_j)$$

The entire idea of integration on measure spaces hinges on these simple integrable functions. We will eventually build up enough tools to show that we can uniquely define the integral of a function by approximating it with SIFs. Note that $\int f d\mu \in B$ takes values in the Banach space B . SMFs and SIFs are both step functions, but the measure of each step in an SIF must be finite. The function $f = \chi_{\mathbb{R}}$ is a SMF, but not a SIF because the measure of $f^{-1}(1) = \infty$ is not finite. For many of the things we will do following this, we only require B to be a normed vector space and not a Banach space. If B must be a Banach space for any propositions, we will specify that.

The set of SIFs form a vector space under pointwise operations for the same reason that the set of SMFs do. Furthermore, we also have that:

$$f \mapsto \int f d\mu$$

is a linear map $SIF(X, B) \rightarrow B$. We now state some properties of SIFs.

1. If f is a SMF, then $x \mapsto \|f(x)\|_B$ is a SMF, because if $f = \sum_{j=1}^n b_j \chi_{E_j}$, then we can simply take the Banach space norm of each b_j to form $x \mapsto \|f(x)\|_B$. Likewise, if f is a SIF, then so is $x \mapsto \|f(x)\|_B$.
2. If f is a \mathbb{R} -valued SIF and $f \geq 0$ (i.e. $f(x) \geq 0$ for each x) then:

$$\int f d\mu \geq 0$$

3. If f, g are \mathbb{R} -valued SIFs with $f \geq g$, then:

$$\int f d\mu \geq \int g d\mu$$

For f a B -valued SIF, define:

$$\|f\|_1 = \int (x \mapsto \|f(x)\|_B) d\mu = \int \|f(x)\|_B d\mu(x)$$

Proposition 11

Let $f, g \in SIF(X, B)$. Then:

$$\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$$

This implies that $\|\cdot\|_1$ is a seminorm on the vector space $SIF(X, B)$. We will turn this into a norm on a quotient space shortly.

Definition 19: Carrier

Let $f : X \rightarrow B$. We define the **carrier** of f to be:

$$\text{carrier}(f) := \{x \in X : f(x) \neq 0\}$$

If f is a SIF, then $\text{carrier}(f) \in S$ is measurable. In topology, the support $\text{supp}(f) = \overline{\text{carrier}(f)}$, so we have already developed some intuition for this. Now, define:

$$\eta := \{f : f \in SIF(X, B) \text{ and } \mu(\text{carrier}(f)) = 0\}$$

η is a vector subspace of $SIF(X, B)$ because $\text{carrier}(f+g) \leq \text{carrier}(f) + \text{carrier}(g)$. Intuitively, η is all the SIFs whose carrier is a null set. In English, that means that if $f \in \eta$, then f is equal to 0 except on a null set, i.e. $f = 0$ almost everywhere. On \mathbb{R} with the Lebesgue measure, η is the set of functions which are nonzero at only a countable number of points.

If $f \in \eta$, then $\int f d\mu = 0$, and so $\|f\|_1 = 0$. Conversely, suppose f is a SIF (let $f = \sum_{j=1}^n b_j \chi_{E_j}$) with $\|f\|_1 = 0$. Then $\sum_{j=1}^n \|b_j\| \mu(E_j) = 0 \implies \mu(E_j) = 0$ if b_j is nonzero, hence $f \in \eta$. So, we see that if we quotient $SIF(X, B)$ by η , we will set each element f with $\|f\|_1 = 0$ to 0. Thus, consider:

$$SIF(X, \mu)/\eta(S, \mu)$$

(changing notation a bit). Then we see that $\|\cdot\|_1$ is a norm on this space, and $f \mapsto \int f d\mu$ is well defined because $\int f d\mu = 0$ if $f \in \eta$. So, $(SIF(S, \mu)/\eta(S, \mu), \|\cdot\|_1)$ is a normed vector space. Let:

$$L^1(X, S, \mu)$$

be the abstract completion of $SIF(S, \mu)/\eta(S, \mu)$. We will eventually determine what this looks like and how to work with this space, and this will allow us to integrate arbitrary functions. However, before we continue with integration we must make a digression into measurable functions and modes of convergence.

4 Measurable Functions

We now study measurable functions, which will give us a setting to define the general integral. Let $\{b_n\}$ be a Cauchy sequence in B and $E \in S$ with $0 < \mu(E) < \infty$. Then $\{b_n \chi_E\}$ is a Cauchy sequence in $SIF(S, \mu)/\eta(S, \mu)$ (recall when we say a sequence in this space is Cauchy we mean with respect to $\|\cdot\|_1$), but we need B to be complete for this to converge.

Suppose that $\{E_j\}$ is a sequence of disjoint subsets of S with $0 < \mu(E_j) < \frac{1}{2^j}$. Form $E = \bigoplus_{j=1}^{\infty} E_j$, and let $F_n = \bigoplus_{j=1}^n E_j$. Then $F_n \uparrow E$. Let $b \in B$. Then $\{b \chi_{F_n}\}$ is Cauchy in $SIF(S, \mu)/\eta(S, \mu)$. We have:

$$\int b \chi_{F_n} d\mu = b \mu(F_n)$$

Our motivation for the general integral is that we wish to define the integral of $b \chi_E$ such that:

$$\int b \chi_E d\mu = \lim \int b \chi_{F_n} d\mu$$

To do this, S must be closed under countable unions so that $E \in S$. Furthermore, if S is actually a σ -ring, then we can disjointize subsets and apply this process to a family of subsets which is not disjoint, hence it is important that S be a σ -ring. We also need μ to be countably additive so that $\sum_{j=1}^n \mu(E_j) = \mu(\bigoplus_{j=1}^n E_j) = \mu(F_n)$, and so we see that this procedure will only work if (X, S, μ) is a measure space. Let (X, S, μ) be a measure space.

Definition 20: S -measurable function

Let B be a Banach space, and let $f : X \rightarrow B$. We say f is **S-measurable** if there is a sequence $\{f_n\}$ of SMFs that converge pointwise to f .

Definition 21: Null Set

A subset $A \subseteq X$ is a **set of measure 0** for μ (a **null set** for μ) if $\exists E \in S$ such that $A \subseteq E$ and $\mu(E) = 0$.

Definition 22: μ -measurable function

Let B be a Banach space, and let $f : X \rightarrow B$. We say f is **μ -measurable** if there is a sequence (f_n) of B -valued SMFs such that $f_n \rightarrow f$ pointwise off a μ -null set ($f_n \rightarrow f$ almost everywhere).

Definition 23: Almost everywhere

If a property P depends on $x \in X$, we say that it holds **almost everywhere** (almost surely) if it holds for all points of X except for those of a μ -null set.

If f, g are S or μ -measurable, then $f + g$ is as well. Similarly, for $r \in \mathbb{R}$, rf is S or μ -measurable too. So, the set of S -measurable (μ -measurable) functions form a vector space. Also, the map $x \mapsto \|f(x)\|_B$ is S or μ -measurable as well. If $B = \mathbb{R}$, then $f \wedge g$ and $f \vee g$ (where $(f \vee g)(x) = \max\{f(x), g(x)\}$, and vice versa for $f \wedge g$) are S or μ -measurable. Finally, if f is B -valued and g is \mathbb{R} -valued, then gf is S or μ -measurable too.

We care about μ -measurability because we don't care what the function f does on null sets; it could be infinite on a null set, but we don't care because that set has measure 0. Now, let $\mathbf{m}(X, S, B)$ be the vector space of S -measurable B -valued functions, and let $\mathbf{m}(X, S, \mu, B)$ be the vector space of μ -measurable B -valued functions. We want a rigorous notion of convergence of sequences in $\mathbf{m}(X, S, B)$ (and in $\mathbf{m}(X, S, \mu, B)$) so if (f_n) is a sequence in $\mathbf{m}(X, S, B)$, and $f_n \rightarrow f$ a.e., then $f \in \mathbf{m}(X, S, B)$ as well. To easily prove this, we define the notion of a separable space.

Definition 24: Separable space

Let (X, τ) be a topological space. A subset $A \subseteq X$ is **separable** if it contains a countably dense subset, or its closure contains a countably dense subset.

We will show that if $f \in \mathbf{m}(X, S, B)$, then $\text{im}(f) \subseteq B$ is separable. First, if (f_n) is a sequence of functions with $\text{im}(f_n)$ separable for each n and $f_n \rightarrow f$ pointwise, then $\text{im}(f)$ is separable as

well. This is because:

$$E = \overline{\bigcup_{n=1}^{\infty} \text{im}(f_n)}$$

is a countable union of separable sets and hence separable, and because $\text{im}(f) \subseteq E$, $\text{im}(f)$ is separable as well. This leads to the following corollary:

Corollary 2

If (f_n) is a sequence in $\mathfrak{m}(X, S, B)$ and if $f_n \rightarrow f$ pointwise, then $\text{im}(f)$ is separable.

This follows immediately because if f_n is a SMF, then its range is finite and hence separable. Now, let f be a SMF with values in B . If \mathcal{O} is an open set in B such that $0 \notin \mathcal{O}$, then $f^{-1}(\mathcal{O}) \in S$. This lends motivation to the next lemma.

Lemma 1

Let (f_n) be a sequence of B -valued functions, and assume that each f_n has the property that $f_n^{-1}(\mathcal{O}) \in S$ for every open $\mathcal{O} \subseteq B \setminus \{0\}$. If $f_n \rightarrow f$ pointwise, then f also has this property.

Proof. TODO. □

Corollary 3

If (f_n) is a sequence of SMFs and $f_n \rightarrow f$ pointwise, then f has the property that if $\mathcal{O} \subseteq B$ is open and $0_B \notin \mathcal{O}$, then $f^{-1}(\mathcal{O}) \in S$.

So, if f is S -measurable, then it pulls back open sets not containing 0 into measurable sets.

Theorem 7

Let $f : X \rightarrow B$. Then f satisfies the following two properties:

1. $\text{im}(f)$ is separable.
2. $\forall \mathcal{O} \subseteq B$ with \mathcal{O} open, $0_B \notin \mathcal{O}$, we have $f^{-1}(\mathcal{O}) \in S$.

if and only if f is S -measurable.

Proof. TODO. □

Corollary 4

If (f_n) is a sequence of S -measurable (also holds for μ -measurable) functions and $f_n \rightarrow f$ pointwise, then f is S -measurable.

This ends our brief discussion of measurable functions. We will eventually use these to define the general integral, but first we must discuss the different modes of convergence in measure spaces. In the future, if we write that a function is measurable, we mean either S -measurable or μ -measurable.

5 Convergence

We now need to consider different ways that functions can converge. Our ultimate goal will be to rigorously define the integral of a function by constructing a sequence of SIFs which converge to that function. We begin with a famous theorem of Egoroff's.

Theorem 8: Egoroff

Let (f_n) be a sequence of μ -measurable functions. Let $E \in S$ with $\mu(E) < \infty$. Suppose that $f_n \rightarrow f$ pointwise on E (in fact, let $f_n \rightarrow f$ a.e.). Then for each $\epsilon > 0$, there is $F \subseteq E$ such that $f_n \rightarrow f$ uniformly on F and $\mu(E \setminus F) < \epsilon$.

Proof. TODO. □

Essentially, Egoroff's theorem says that we can make the region which $f_n \not\rightarrow f$ uniformly inside of E arbitrarily small. Note the precondition that $\mu(E) < \infty$. Indeed, let us examine an example where $\mu(E) = \infty$. Take \mathbb{R} with the Lebesgue measure, and $f_n = \chi_{[n, n+1]}$. Then this function converges pointwise to 0, but we cannot find a set for $\epsilon = 1$ such that $f_n \rightarrow 0$ uniformly on F and $\mu(\mathbb{R} \setminus F) < 1$.

Definition 25: Almost uniform convergence

Let (f_n) be a sequence of measurable functions converging to f , and $E \in S$. We say that $f_n \rightarrow f$ **almost uniformly** if for every $\epsilon > 0$, there is $F \subseteq E$ such that $f_n \rightarrow f$ uniformly on F and $\mu(E \setminus F) < \epsilon$.

Egoroff's theorem tells us that if $f_n \rightarrow f$ pointwise on a set of finite measure, then this convergence is almost uniform as well. There is a stronger converse to this.

Proposition 12

If (f_n) converges to f almost uniformly on E , then $f_n \rightarrow f$ almost everywhere (pointwise except off a null set).

Proof. Suppose $f_n \rightarrow f$ almost uniformly. For each n , we can find $F_n \subseteq E$ such that $\mu(E \setminus F_n) < \frac{1}{n}$ and $f_n \rightarrow f$ uniformly on F_n . Form $F = \bigcup_{n=1}^{\infty} F_n$. Then $f_n \rightarrow f$ pointwise on F , because for each $x \in F$, we have $x \in F_n$ for some n . We claim that $\mu(E \setminus F) = 0$. Then because each $F_n \subseteq F$, we have $\mu(E \setminus F) \leq \mu(E \setminus F_n) < \frac{1}{n}$ for each n , thus $\mu(E \setminus F) \leq 0$ and is 0. So, $f_n \rightarrow f$ pointwise except on the null set $E \setminus F$. □

Just as we have Cauchy sequences of real numbers, we also can consider different types of Cauchy sequences of functions.

Definition 26: Almost uniformly Cauchy

Let (f_n) be a sequence of μ -measurable functions and $E \in S$. We say that (f_n) is **almost uniformly Cauchy** on E if for each $\epsilon > 0$, there is $F \subseteq E$ such that $\mu(E \setminus F) < \epsilon$ and (f_n) is uniformly Cauchy on F , i.e. for each $\delta > 0$, we can find N such that for $n, m \geq N$, we have $\|f_n(x) - f_m(x)\|_B < \delta$ for all $x \in F$.

Proposition 13

Let B be a Banach space. If (f_n) is almost uniformly Cauchy on E , then there is $f : E \rightarrow B$ such that $f_n \rightarrow f$ almost uniformly.

Proof. Let $\epsilon > 0$. Then we can find $F \subseteq E$ such that $\mu(E \setminus F) < \epsilon$ and (f_n) is uniformly Cauchy on F . But this implies that $(f_n(x))$ converges because B is complete and this is Cauchy, so let this converge to $f(x) \in B$. At this point, we can pick F_n with $\mu(E \setminus F_n) < \frac{1}{n}$ to define f on all of $\bigcup_{n \in \mathbb{N}} F_n$ by defining the function f on $F_{n+1} \setminus \bigcup_{k=1}^n F_k$ recursively, since each step will assign f a value on F_n , and once we assign f a value, we wish to keep it assigned. \square

In other words, a sequence of almost uniformly Cauchy will converge almost uniformly to a function if B is complete (just like the notion of convergence in metric spaces). Now, Cauchy sequences of SIFs with respect to $\|\cdot\|_1$ do not always converge with respect to other modes of convergence. For example, consider $[0, 1]$ with the Lebesgue measure. We can define $E_1 = [0, \frac{1}{2}]$, $E_2 = [\frac{1}{2}, 1]$, $E_3 = [0, \frac{1}{3}]$, ... and let $f_n = \chi_{E_n}$. Then this is Cauchy w.r.t. $\|\cdot\|_1$, yet this does not converge almost everywhere or almost uniformly.

Instead, we will aim to show that we can find a convergent subsequence of a Cauchy sequence. To formalize this, we need another notion of convergence.

Definition 27: Convergence in measure

Let (f_n) be a sequence of measurable functions, and f measurable. Then we say that f_n **converges to f in measure** on $E \subseteq X$ if for each $\epsilon > 0$, we have:

$$\lim_{n \rightarrow \infty} \mu(\{x \in E : \|f(x) - f_n(x)\|_B > \epsilon\}) = 0$$

Intuitively, given any tolerance $\epsilon > 0$, $f_n \rightarrow f$ in measure if the measure of the set of $x \in E$ that are outside the ϵ tube eventually gets arbitrarily small. To prove this, we will often have to introduce a $\delta > 0$ to show that this limit goes to 0 for each ϵ . In the sequence above, the sequence f_n converges to 0 in measure; no matter the tolerance, the measure of the set on which f_n is nonzero will eventually be arbitrarily small. We will discuss how the types of convergence are related, and prove some basic propositions about convergence in measure.

Proposition 14

If (f_n) converges to f almost uniformly, then $f_n \rightarrow f$ in measure.

Proof. Let $\epsilon > 0$. Let $G_n := \{x \in E : \|f(x) - f_n(x)\| > \epsilon\}$. We need to show that $\lim \mu(G_n) = 0$. So, let $\delta > 0$. Because $f_n \rightarrow f$ a.u., we can pick $F \subseteq E$ such that $f_n \rightarrow f$ uniformly on F and $\mu(E \setminus F) < \delta$. Thus we can find $N \in \mathbb{N}$ such that $n \geq N$ implies $\|f(x) - f_n(x)\| < \epsilon$ for $x \in F$, and thus for $n \geq N$ we have $G_n \subseteq E \setminus F$. Therefore, we have $\mu(G_n) \leq \mu(E \setminus F) < \delta$ for $n \geq N$, and hence $\lim \mu(G_n) = 0$. \square

Proposition 15

Let (f_n) be a sequence of measurable functions. Suppose that $f_n \rightarrow f$ in measure and $f_n \rightarrow g$ in measure. Then $f = g$ almost everywhere.

Proof. Let $\epsilon > 0$. Then $\|f(x) - g(x)\| \geq \|f(x) - f_n(x)\| + \|f_n(x) - g(x)\|$, so if $\|f(x) - g(x)\| > 0$, then one of these summands must be $> \frac{\epsilon}{2}$. Thus, $F := \{x \in E : \|f(x) - g(x)\| > \epsilon\} \subseteq \{x \in E : \|f(x) - f_n(x)\| > \frac{\epsilon}{2}\} \cup \{x \in E : \|f_n(x) - g(x)\| > \frac{\epsilon}{2}\} =: F_1^n \cup F_2^n$. Hence we have $\mu(F) \leq \mu(F_1^n) + \mu(F_2^n)$. Thus for $\delta > 0$, we can find $N \in \mathbb{N}$ such that for $n \geq N$, we have $\mu(F_1^n), \mu(F_2^n) < \frac{\delta}{2}$ because f_n converges in measure to f and to g , so $\mu(F) < \delta$. Therefore, $\mu(F) = 0$, so $f = g$ a.e. \square

Definition 28: Cauchy in measure

Let (f_n) be a sequence of measurable functions. We say (f_n) is **Cauchy in measure** on $E \subseteq X$ if for each $\epsilon > 0$, we have:

$$\mu(\{x \in E : \|f_n(x) - f_m(x)\| > \epsilon\}) \rightarrow 0 \text{ as } m, n \rightarrow \infty$$

i.e. given $\delta > 0$, we can find $N \in \mathbb{N}$ such that for $n, m \geq N$, we have $\mu(\{x \in E : \|f_n(x) - f_m(x)\| > \epsilon\}) < \delta$.

Definition 29: Cauchy in mean, equivalent

Let (f_n) be a sequence of SIFs. We say that (f_n) is **Cauchy in mean** (mean Cauchy) if (f_n) is Cauchy with respect to $\|\cdot\|_1$. We say that (f_n) and (g_n) are **equivalent Cauchy sequences** if $\|f_n - g_n\|_1 \rightarrow 0$ as $n \rightarrow \infty$.

Proposition 16

Let (f_n) be a sequence of SIFs that is mean Cauchy. Then (f_n) is Cauchy in measure.

Proof. Let $\epsilon > 0$, and let $\delta > 0$. Choose $N \in \mathbb{N}$ such that $n, m \geq N$ implies $\|f_n - f_m\|_1 < \delta\epsilon$. Consider $\mu(\{x \in E : \|f_n(x) - f_m(x)\| > \epsilon\})$. Note that if $\|f_n(x) - f_m(x)\| > \epsilon$, then we have:

$$\frac{1}{\epsilon} \|f_n(x) - f_m(x)\| \geq \chi_{\{x \in E : \|f_n(x) - f_m(x)\| > \epsilon\}}(x)$$

Then we have:

$$\begin{aligned} \mu(\{x \in E : \|f_n(x) - f_m(x)\| > \epsilon\}) &= \int \chi_{\{x \in E : \|f_n(x) - f_m(x)\| > \epsilon\}} d\mu \leq \frac{1}{\epsilon} \int \|f_n(x) - f_m(x)\| d\mu(x) \\ &< \frac{\delta\epsilon}{\epsilon} = \delta \end{aligned}$$

for $n, m \geq N$, hence (f_n) is Cauchy in measure. \square

Proposition 17

Let $f_n \rightarrow f$ and $g_n \rightarrow g$ in measure. Then:

1. $f_n + g_n \rightarrow f + g$ in measure.
2. For $r \in \mathbb{R}$, $rf_n \rightarrow rf$ in measure.

3. $(\|f_n(\cdot)\|)$ converges to $\|f(\cdot)\|$.

This last proposition will be necessary to prove the next theorem, which will allow us to find a convergent subsequence of a sequence which is Cauchy in measure. Our main goal will be to relate each mode of convergence to one another.

Theorem 9: Riez-Weyl

Let (f_n) be a sequence of measurable functions. If (f_n) is Cauchy in measure, then there is a subsequence of (f_n) which is almost uniformly Cauchy.

Proof. TODO. □

Proposition 18

Let (f_n) be a sequence which is Cauchy in measure, and suppose it has a subsequence (f_{n_k}) such that $f_{n_k} \rightarrow f$ almost uniformly. Then (f_n) converges to f in measure.

Proof. TODO. □

To sum up, we have seen the following sequence of implications: (f_n) mean Cauchy $\implies (f_n)$ Cauchy in measure $\implies (f_n)$ has a convergent subsequence (f_{n_k}) which is almost uniformly Cauchy $\implies (f_{n_k})$ converges a.u. to some function $f \implies (f_n)$ converges to f in measure, and f is unique almost everywhere.

Thus, if (f_n) is a mean Cauchy sequence or (f_n) is Cauchy in measure, then it converges to a function in measure. We are almost there! We just need a few more tools, then we will have built up enough machinery with convergence to define the integral of a function which is not step.

Proposition 19

Let $(f_n), (g_n)$ be mean Cauchy sequences of SIFs which are equivalent, i.e. $\|f_n - g_n\|_1 \rightarrow 0$ as $n \rightarrow \infty$. Then if $f_n \rightarrow f$ in measure, so does (g_n) .

Proof. Let (h_n) be the sequence $(f_1, g_1, f_2, g_2, \dots)$. Then (h_n) is mean Cauchy and it has a subsequence (f_n) which converges to f in measure, hence (h_n) converges to f in measure. Therefore $g_n \rightarrow f$ in measure as well. □

Now, consider $\mathfrak{m}(X, S, \mu, B)$. We can define an equivalence relation by letting $f \sim g$ iff $f = g$ almost everywhere. Then define:

$$\mathcal{M}(X, S, \mu, B) := \mathfrak{m}(X, S, \mu, B) / \sim$$

So, we get a map:

$$\Theta : \{\text{equivalence classes of mean Cauchy sequences of } B\text{-valued SIFs}\} \rightarrow \mathcal{M}(X, S, \mu, B)$$

This map Θ is injective, as shown in the following proposition.

Proposition 20

If (f_n) and (g_n) are mean Cauchy sequences and if they both converge in measure to the same

function f , then (f_n) and (g_n) are equivalent Cauchy sequences.

Proof. TODO. □

This allows us to relate all our notions of convergence for mean Cauchy sequences. We need to do this to define the general integral, because if a mean Cauchy sequence converges to f with respect to one mode of convergence, we would like it if there was a mean Cauchy sequence which converged to f with respect to all the modes of convergence.

Theorem 10

Let $f \in \mathfrak{m}(X, S, \mu, B)$. TFAE:

1. There is a mean Cauchy sequence (f_n) which converges to f in measure.
2. There is a mean Cauchy sequence (f_n) which converges to f almost uniformly.
3. There is a mean Cauchy sequence (f_n) which converges to f almost everywhere.

Proposition 21

TODO.

6 Integration of General Functions

We finally have enough machinery at our disposal so that we can define the general integral.

Definition 30: μ -integrable

Let $f \in \mathfrak{m}(X, S, \mu, B)$. Then f is **μ -integrable** if it satisfies any of the equivalent conditions in the previous theorem. Let $\mathcal{L}^1(X, S, \mu, B)$ be the set of μ -integrable functions. For $f \in \mathcal{L}^1(X, S, \mu, B)$, there is a mean Cauchy sequence (f_n) of SIFs converging to f in measure, almost uniformly, or almost everywhere. So, we define:

$$\int f d\mu := \lim \int f_n d\mu$$

For $E \in S$, define:

$$\int_E f d\mu = \int (\chi_E \cdot f) d\mu$$

First note that this definition of the integral is well defined because if $g_n \rightarrow f$ in measure / a.u. / a.e., then $g_n \sim f_n$, and thus $\lim \int f_n d\mu = \lim \int g_n d\mu$. We will now discuss some properties of this set. If we write \mathcal{L}^1 , assume we mean $\mathcal{L}^1(X, S, \mu, B)$.

1. If $f, g \in \mathcal{L}^1$, then $f + g \in \mathcal{L}^1$, and $\int (f + g) d\mu = \int f d\mu + \int g d\mu$. If $\alpha \in \mathbb{R}$, then $\alpha f \in \mathcal{L}^1$ and $\int \alpha f d\mu = \alpha \int f d\mu$.
2. If $f \in \mathcal{L}^1$, then $(x \mapsto \|f(x)\|_B) \in \mathcal{L}^1$. So, define:

$$\|f\|_1 := \int \|f(x)\| d\mu(x)$$

3. If $f \in \mathcal{L}^1(X, \mathbb{R})$ and $f \geq 0$ a.e., then:

$$\int f d\mu \geq 0$$

If (f_n) is mean Cauchy and \mathbb{R} -valued and converges to f , then $(\max\{f_n, 0\})$ is also mean Cauchy and converges to f .

4. If $f, g \in \mathcal{L}^1(X, \mathbb{R})$ and $f \geq g$, then:

$$\int f d\mu \geq \int g d\mu$$

5. If $f, g \in \mathcal{L}^1$, then $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$.

6. If $f \in \mathcal{L}^1$ and (f_n) is a mean Cauchy sequence of SIFs converging to f , then $\|f_n\|_1 \rightarrow \|f\|_1$, so:

$$\|f\|_1 = \int \|f(x)\|_B d\mu(x) = \lim \int \|f_n(x)\|_B d\mu(x)$$

Therefore we also have:

$$\|f - g\|_1 = \lim \|f_n - g_n\|_1$$

which appears to be very similar to the metric we defined on the completion of our metric space.

7. If $E, F \in S$ and $E \cap F = \emptyset$, then:

$$\int_{E \oplus F} f d\mu = \int_E f d\mu + \int_F f d\mu$$

8. Let $f \in \mathcal{L}^1(X, \mathbb{R})$ and $f \geq 0$. If $E, F \in S$ and $F \subseteq E$, then:

$$\int_F f d\mu \leq \int_E f d\mu$$

All of the above implies that $\|\cdot\|_1$ is a seminorm on $\mathcal{L}^1(X, S, \mu, B)$, and the space $\mathcal{L}^1(X, S, \mu, B)$ is complete with respect to $\|\cdot\|_1$ (by analogy with the work we have done previously with the completion of a metric space). However, this is not yet a full norm on the space, because if a function equals 0 almost everywhere, then it will have norm 0. So, we make a definition.

Definition 31

Let \sim be the equivalence relation $f \sim g$ iff $f = g$ a.e. on $\mathcal{L}^1(X, S, \mu, B)$. Define:

$$L^1(X, S, \mu, B) := \mathcal{L}^1(X, S, \mu, B) / \sim$$

On $L^1(X, S, \mu, B)$, $\|\cdot\|_1$ is indeed a valid norm. Therefore:

$$(L^1(X, S, \mu, B), \|\cdot\|_1)$$

is a normed vector space which is complete for its norm, i.e. it is a Banach space. Now that we have an integral, we can also define a notion of an indefinite integral.

Definition 32: Indefinite integral

The **indefinite integral** of f , denoted μ_f , is the function $\mu_f : S \rightarrow B$ such that:

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

Property 7 shows that μ_f is finitely additive. In fact, μ_f is a B -valued measure on S , i.e. it is countably additive. It is also monotone by property 8. We will now briefly discuss the notion of absolute continuity of measures, and then prove a few helpful convergence theorems.

Proposition 22

If $f \in \mathcal{L}^1$, then $C_f := \text{carrier}(f)$ is σ -finite.

Proof. Let (f_n) be a mean Cauchy sequence converging to f almost everywhere. Then $\mu(C_{f_n}) < \infty$, so let $D = \bigcup_{n=1}^{\infty} C_{f_n}$. Then D is σ -finite by definition, so because $f_n \rightarrow f$ a.e., $C_f \subseteq D$, and thus C_f is σ -finite. \square

Proposition 23

Let $f \in \mathcal{L}^1(X, S, \mu, B)$. Then for any $\epsilon > 0$, we can find $E \in S$ such that $\mu(E) < \infty$, and:

$$\left\| \int_{X \setminus E} f d\mu \right\| < \epsilon$$

Proof. Find a SIF g such that $\|f - g\|_1 < \epsilon$. Let $E = \text{carrier}(g)$. Then we are done because g is zero on $X \setminus E$, so:

$$\left\| \int_{X \setminus E} f d\mu \right\| = \left\| \int_{X \setminus E} (f - g) d\mu \right\| \leq \int_{X \setminus E} \|f(x) - g(x)\| d\mu(x) \leq \|f - g\|_1 < \epsilon$$

\square

We say that \mathcal{L}^1 functions are **almost supported on sets of finite measure**. Intuitively, this means that the functions in \mathcal{L}^1 must die off sufficiently fast as we look toward infinity. Given any tolerance, we can choose a finite set large enough to make the integral of f on this set smaller than the given tolerance. Next, we define the notion of absolute continuity.

Definition 33: Absolutely continuous

Let μ, ν be measures on a σ -ring (X, S) . We say that ν is **absolutely continuous** with respect to μ if given $\epsilon > 0$, we can find $\delta > 0$ such that $\nu(E) < \delta$ implies $\mu(E) < \epsilon$. We write $\nu \ll \mu$.

Proposition 24

If $f \in \mathcal{L}^1(X, S, \mu, B)$, then for all $\epsilon > 0$, there is $\delta > 0$ such that if $\mu(E) < \delta$, then:

$$\mu_f(E) < \epsilon$$

i.e. $\mu \ll \mu_f$.

Proof. Let $\epsilon > 0$. Choose g an SIF with $\|f - g\|_1 < \frac{\epsilon}{2}$. Then $\|f\|_1 \leq \|f - g\|_1 + \|g\|_1$, so for $E \in S$, we have:

$$\int_E \|f(x)\| d\mu(x) \leq \frac{\epsilon}{2} + \left\| \int_E g d\mu \right\|$$

But, g is bounded, so $\left\| \int_E g d\mu \right\| \leq \|g\|_\infty \mu(E)$, so set $\delta = \frac{\epsilon}{2}$. Then $\mu(E) < \delta$ implies that $\mu_f(E) = \left\| \int_E f d\mu \right\| < \epsilon$. \square

6.1 Convergence Theorems

To end this section, we state two important theorems which will be useful for proving that functions are integrable. We begin with the Dominated Convergence Theorem.

Theorem 11: Lebesgue Dominated Convergence Theorem

Let (f_n) be a sequence of functions in $\mathcal{L}^1(X, S, \mu, B)$, and suppose that $f_n \rightarrow f$ a.e. (note this implies f is measurable). If there is a \mathbb{R} -valued function $g \in \mathcal{L}^1(X, S, \mu, B)$ such that $\|f_n(x)\| \leq g(x)$ for each $x \in X$ (i.e. each f_n is **dominated** by g) then (f_n) is a mean Cauchy sequence.

Proof. TODO. \square

In particular, the DCT implies that the f given in the assumption is integrable. This is because (f_n) mean Cauchy implies that (f_n) is Cauchy in measure, so a subsequence converges a.u. to some function which must equal f a.e., hence (f_n) itself converges to f a.u., and hence $f \in \mathcal{L}^1(X, S, \mu, B)$. This is reflected in the following corollary.

Corollary 5

Let f be a B -valued μ -measurable function. If there is $g \in \mathcal{L}^1(X, S, \mu, \mathbb{R})$ such that $\|f(x)\| \leq g(x)$ a.e., then $f \in \mathcal{L}^1(X, S, \mu, B)$.

Proof. TODO. \square

Now, we can use this to prove another convergence theorem which is quite useful as well.

Theorem 12: Monotone Convergence

Let (f_n) be a sequence of \mathbb{R} -valued functions, measurable, with $f_n \geq 0$, and assume that $f_{n+1} \geq f_n$ (the sequence is monotone). If there is a constant C such that $\int f_n d\mu \leq C$ for each $n \in \mathbb{N}$, then $f_n \in \mathcal{L}^1$ and (f_n) is mean Cauchy. Furthermore, $f_n \rightarrow f$ a.e. for some $f \in \mathcal{L}^1(X, S, \mu, \mathbb{R})$.

Proof. The sequence $(\int f_n d\mu)$ is a sequence of increasing positive numbers which is bounded above by C . For $n \geq m$, this implies that $\|f_n - f_m\|_1 = \int f_n d\mu - \int f_m d\mu \rightarrow 0$ as $n, m \rightarrow \infty$, so (f_n) is mean Cauchy. We can then use what we've already seen to build the chain of implications to show that (f_n) converges to an integrable function f . \square

7 Normed Vector Spaces

7.1 Bounded Operators

We begin our brief study of normed vector spaces by considering bounded operators. Let V and W be normed vector spaces, and let $T : V \rightarrow W$ be a linear map.

Definition 34: Bounded

We say T is a **bounded operator** if there is some $C \geq 0$ such that for each $v \in V$, we have:

$$\|Tv\| \leq C\|v\|$$

If T is bounded, we set:

$$\|T\|_\infty := \sup\{\|Tv\| : v \in V, \|v\| \leq 1\} = \inf\{C \in \mathbb{R} : \|Tv\| \leq C\|v\|, \forall v \in V\}$$

Theorem 13

The following are equivalent.

1. T is continuous.
2. T is continuous at 0.
3. T is bounded.
4. T is Lipschitz, and hence uniformly continuous.

Proof. TODO. □

Let $B(V, W)$ be the set of all bounded linear maps from V to W . For $S, T \in B(V, W)$, we have:

$$\|(S + T)v\|_W \leq \|Sv\|_W + \|Tv\|_W \leq \|S\| \cdot \|v\| + \|T\| \cdot \|v\| = (\|S\| + \|T\|)\|v\|$$

so we see that $\|S + T\| \leq \|S\| + \|T\|$. Furthermore, $\|\cdot\|$ also respects scaling, hence we have the following proposition.

Theorem 14

Let V and W be normed vector spaces. Then $B(V, W)$ is itself a normed vector space with the sup norm $\|\cdot\|$.

Let V, W be Banach spaces with $T \in B(V, W)$. Let (X, S, μ) be a measure space. Then we get an operator:

$$\Phi_T : \mathcal{L}^1(X, S, \mu, V) \rightarrow \mathcal{L}^1(X, S, \mu, W)$$

To define this, let f be an SMF with values in V . Then we can define $(Tf)(v) := Tf(v)$, which is a SMF with values in W . If (f_n) is a sequence of SMFs converging pointwise to a function f (so f is measurable), then (Tf_n) converges pointwise to (Tf) . Hence we need to show that (Tf_n) is mean Cauchy. We have:

$$\|Tf_n - Tf_m\|_1 = \int \|Tf_n(x) - Tf_m(x)\| d\mu(x) \leq \|T\| \int \|f_n(x) - f_m(x)\| d\mu(x) = \|T\| \cdot \|f_n - f_m\|_1$$

but we can make $\|f_n - f_m\|_1$ arbitrarily small since (f_n) is mean Cauchy, hence (Tf_n) is mean Cauchy and so $Tf \in \mathcal{L}^1(X, S, \mu, W)$. Furthermore, if $f_n \rightarrow f$ a.e. with (f_n) mean Cauchy, then:

$$\int Tf d\mu = \lim \int Tf_n d\mu = \lim T \int f_n d\mu = T(\lim \int f_n d\mu) = T(\int f d\mu)$$

where we were able to pull T out of the integral because it is simply a sum, as each f_n is a SIF. So, we can make the following proposition.

Proposition 25

Let $f \in \mathcal{L}^1(X, S, \mu, V)$ and $T : V \rightarrow W$. Then $Tf \in \mathcal{L}^1(X, S, \mu, W)$ and:

$$\int (Tf) d\mu = T \left(\int f d\mu \right).$$

Note that $\|Tf\|_1 \leq \|T\| \cdot \|f\|_1$ by definition of $\|T\|$, and $T(f + g) = Tf + Tg$. So define:

$$\Phi_T(f) := Tf.$$

Then $\Phi_T \in B(L^1(X, S, \mu, V), L^1(X, S, \mu, W))$, and:

$$\|\Phi_T\| \leq \|T\|.$$

7.2 The Continuous Dual

Let $K = \mathbb{R}$ or \mathbb{C} . Given a K -vector space V , let V^* be its dual space, i.e. $V^* = \text{Hom}_K(V, \mathbb{R})$.

Definition 35: Continuous dual

The **continuous dual** of V is the set of all bounded linear functions on V and denoted by V' .

So, $V' \subseteq V^*$, and V' is a Banach space because K is complete. Now, let $V = L^1(X, S, \mu, \mathbb{R})$. How can we determine V' ? Let $h \in V$ be a bounded measurable function, so $\|h\|_\infty < \infty$. Then, define ϕ_h on V by:

$$\phi_h(f) := \int_X hf d\mu$$

Note hf is integrable because it is dominated by $\|h\|_\infty f$ and hf is measurable. So, we have:

$$|\phi_h(f)| \leq \int_X |h(x)| \cdot |f(x)| d\mu(x) \leq \|h\|_\infty \int_X |f(x)| d\mu(x) = \|h\|_\infty \cdot \|f\|_1$$

In particular, this implies that:

$$\|\phi_h\| \leq \|h\|_\infty$$

as we see that $\|h\|_\infty$ is an upper bound on the scaling constant.

Definition 36: Essentially bounded

Let $g : X \rightarrow V$ be measurable. We say that g is **essentially bounded** if there is a constant c such that:

$$\mu(\{x \in X : \|g(x)\| > c\}) = 0.$$

If g is essentially bounded, set $\|g\|_\infty$ to be the infimum over all c for which this holds. Let $\mathcal{L}^\infty(X, S, \mu, V)$ be the vector space of all essentially bounded functions equipped with the

seminorm $\|\cdot\|_\infty$.

An essentially bounded function is bounded almost everywhere. On \mathbb{R} with the Lebesgue measure, this means that f is essentially bounded if f is only unbounded at a countable number of points. Note that $\|\cdot\|_\infty$ is not a valid norm because a function which equals 0 almost everywhere will be essentially bounded. So, in the spirit of how we defined L^1 , we can define L^∞ . We let \sim be the equivalence relation on \mathcal{L}^∞ such that $f \sim g$ if and only if $f = g$ a.e., and define:

$$L^\infty(X, S, \mu, V) := \mathcal{L}^\infty(X, S, \mu, V) / \sim$$

Then $L^\infty(X, S, \mu, V)$ is a normed vector space, and in fact a Banach space. Given $h \in \mathcal{L}^\infty(X, \mathbb{R})$ and $f \in \mathcal{L}^1(X, S, \mu, B)$, then $hf \in \mathcal{L}^1(X, S, \mu, B)$ because hf is dominated by the function $\|h\|_\infty f \in \mathcal{L}^1(X, S, \mu, B)$. If $B = \mathbb{R}$, then we have:

$$\int h f d\mu \in \mathbb{R}$$

so define $\Phi_h : L^1(\mathbb{R}) \rightarrow \mathbb{R}$ as $f \mapsto \int h f d\mu$. Then Φ_h is a bounded linear functional and $\|\Phi_h\| \leq \|h\|_\infty$ (they are in fact equal).

Theorem 15

Let (X, S, μ) be a σ -finite measure space. Then we have:

$$L^1(X, S, \mu, K)^* \cong L^\infty(X, S, \mu, \mathbb{C})$$

$$\Phi_h \xrightarrow{\sim} h$$

If (X, S, μ) is not σ -finite, we need to use essentially bounded functions which are locally measurable.

Now, let U, V, W be normed vector spaces, and consider the chain of maps:

$$U \xrightarrow{S} V \xrightarrow{T} W$$

with S, T bounded and linear. Then TS (the composition of T and S) is a bounded linear operator, and $\|TS\| \leq \|T\| \cdot \|S\|$. Let:

$$B(V) := B(V, V)$$

Then $B(V)$ is an algebra for its usual addition and product $(S, T) \mapsto ST$, and pointwise scaling.

Definition 37: Normed algebra

A **normed algebra** (over K) \mathcal{A} is an algebra with a norm $\|\cdot\|$ such that for $a, b \in \mathcal{A}$, we have:

$$\|a \cdot b\| \leq \|a\| \cdot \|b\|$$

If \mathcal{A} is complete with respect to $\|\cdot\|$, we call it a **Banach algebra**.

So $B(V)$ is a normed algebra, and $C(X)$ is as well.

Proposition 26

If W is complete, then $B(V, W)$ is complete and hence is a Banach space.

Proof. TODO. □

We see that if V is a Banach space, then $B(V)$ is a Banach algebra. Let $h \in L^\infty(X, S, \mu, K)$. Then for $f \in L^1(X, S, \mu, B)$, we know that $hf \in L^1(X, S, \mu, B)$. Thus:

$$\|hf\|_1 = \int \|h(x)f(x)\| d\mu(x) \leq \|h\|_\infty \|f\|_1$$

We can thus define $T_h \in B(L^1(X, B))$ by $T_h f := hf$, and we see that $\|T_h\| \leq \|h\|_\infty$ (these are in fact equal). The map $h \mapsto T_h$ preserves multiplication $T_h T_k = T_{hk}$. So, the set:

$$\{T_h : h \in L^\infty(X, S, \mu, K)\}$$

is a commutative subalgebra of $B(L^1(X, B))$. We will study these further next semester.

8 Examples

In this section, I will enumerate the multiple different spaces that we have studied and their properties.

- Let (X, τ) be a topological space. Then:

$$C(X) = \{f : X \rightarrow \mathbb{R} \mid f \text{ is continuous}\}$$

is the space of all continuous functions $X \rightarrow \mathbb{R}$. We can define a norm $\|\cdot\|_\infty$ on $C(X)$:

$$\|f\|_\infty = \sup_{x \in X} |f(x)|$$

If X is compact, then $\|\cdot\|_\infty$ makes $C(X)$ into a complete metric space.

- The space $X = [0, 1]$. Define $C([0, 1])$ as above. As discussed, $C([0, 1])$ with the sup norm $\|\cdot\|_\infty$ is complete (however, it is not a compact space). We can also put additional norms on this space:

$$\begin{aligned} \|f\|_1 &= \int_0^1 |f(x)| d\mu(x) \\ \|f\|_p &= \left(\int_0^1 |f(x)|^p d\mu(x) \right)^{\frac{1}{p}} \end{aligned}$$

Under these norms, $C(X)$ is not complete.

- Let X be any set and V be a normed vector space. Then:

$$B(X, V) = \{f : X \rightarrow V \mid f \text{ is bounded}\}$$

where by bounded we mean $\exists R > 0$ such that $f(X) \subseteq B_R(0)$. We define:

$$\|f\|_\infty = \sup_{x \in X} \|f(x)\|$$

which exists because f is bounded. $\|\cdot\|_\infty$ is a norm on this vector space, and $B(X, V)$ is a Banach space with respect to this norm.

- Let X be a set and Y a metric space. Then $B(X, Y)$ is the collection of bounded functions $f : X \rightarrow Y$, and for $f, g \in B(X, Y)$, we can define:

$$d_\infty(f, g) = \sup_{x \in X} d(f(x), g(x))$$

Then $(B(X, Y), d_\infty)$ is a metric space. If Y is complete, so is $B(X, Y)$.

- Let (X, τ) be a topological space and V a Banach space. Then:

$$C_b(X, V) = \{f : X \rightarrow V \mid f \text{ is bounded and continuous}\}$$

$C_b(X, V)$ is a closed subspace of $C(X, V)$, and is a Banach space as well.

- For X a topological space and Y a metric space, let $C(X, Y)$ be the set of all continuous functions $X \rightarrow Y$. Refer to the section on compact metric spaces for a discussion of the properties of this space.
- Let X be locally compact and Hausdorff, and let $K = \mathbb{R}$ or \mathbb{C} . Then define:

$$C_C(X) = \{f : X \rightarrow K \mid f \text{ is continuous with compact support}\}$$

(recall $\text{supp}(f) = \overline{\{x \in X : f(x) \neq 0\}}$). This is an algebra with pointwise multiplication, and $C_C(X) \subseteq C_b(X)$. The sup norm $\|\cdot\|_\infty$ is a valid norm, so $C_C(X)$ is a Banach algebra. The closure of $C_C(X)$ in $C(X)$ is continuous functions on X that vanish at infinity.