LEBESGUE SPACES AND THE RIESZ-FISCHER THEOREM

MARK O'SHEA

ABSTRACT. This paper serves as an introduction to measure theory and functional analysis through an exploration of L_p spaces, also known as Lebesgue spaces. Our goal is to demonstrate the completeness of L_p spaces, formally known as the Riesz-Fischer Theorem. We assume the reader has a solid grasp of concepts in undergraduate-level real analysis and linear algebra, but might not have a rigorous foundation in more advanced analytical topics. Accordingly, we will slowly build up to the Riesz-Fischer Theorem by introducing prerequisite concepts such as metric spaces, normed spaces, sigma algebras, Borel sets, Minkowski's inequality, and the Monotone and Dominated Convergence Theorems.

CONTENTS

1. BASIC BANACH SPACE THEORY

The study of real analysis is characterized by the concepts of size and distance. This notion also holds practical value; economists, physicists, and biologists are all concerned with taking precise measurements of objects. One of the fundamental ways in which we formalize the idea of magnitude in mathematics is through the concept of a metric. As an exercise, the reader might ponder what properties of a metric make intuitive sense. For instance, take the number of miles between arbitrary points A and B. It is never negative, it is only ever zero if A and B are the same places, the journey there is the same length as the journey back, and detours are always longer. We can write this mathematically in the following way.

Definition 1.1. A *metric* on a nonempty set V is a function $d : V \times V \to [0, \infty)$[1] such that:
i) $d(f, g) = 0$ if and only if $f = g$, for all $f, g \in V$ (Definiteness)

[1]Some authors are explicit in requiring that a metric, and as we shall see later, a norm, is always greater or equal to zero. This condition is called non-negativity.

ii) $d(f,g) = d(g,f)$, for all $f,g \in V$ (Symmetry)
iii) $d(x,y) + d(y,z) \geq d(x,z)$ for all $x,y,z \in V$ (Triangle inequality)

We can easily take the idea of a metric and apply it to a set. Such a set can be something as simple as some arbitrary finite collection $\{\phi, \gamma, \rho\}$ to more complicated objects like $\mathbb{R}$. When we do so, we arrive at the idea of a metric space.

Definition 1.2. A *metric space* is a pair (V,d), where V is a nonempty set and d is a metric on v.

Example 1.3. The most typical example of a metric space is the Euclidean plane $(\mathbb{R}^2, d)$ where $d((x_1,y_1),(x_2,y_2)) = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$. As you may recall, this construction allows us to find the distance between two points by connecting a straight 2-D line between them. Another metric that we could've defined on this set is $d((x_1,y_1),(x_2,y_2)) = |x_2-x_1| + |y_2-y_1|$. This metric is also known as the 'taxicab metric' because of how this notion of distance only allows us to 'travel' between points using lines that are parallel to the x and y axes, which resembles the block-like path of a car driving through an urban center.

We can now make the definition of a metric space more robust. The reader might recall that some sequences in $\mathbb{Q}$ do not converge to a limit in $\mathbb{Q}$. Thus, we can refer to this as an 'incomplete' metric space. This fact is troublesome because we can think of an incomplete metric space as missing elements. Though a separate issue, it is worth pointing out that every incomplete metric space can be made complete by adding the missing elements. We now introduce the notion of a complete metric space, but first give the following definition.

Definition 1.4. A sequence a_n is called a *Cauchy sequence* if for any given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $n,m \geq N$ implies $|a_n - a_m| < \epsilon$.

We are now ready to consider the formal notion of a complete metric space.

Definition 1.5. A *complete metric space* is a metric space M such that every Cauchy sequence in M converges to a limit in M.

Example 1.6. Let $\mathrm{C}[0,1]$ represent the set of continuous real-valued functions defined on the closed interval $[0,1]$. Define a metric d on the set $\mathrm{C}[0,1] \times \mathrm{C}[0,1]$ by $d(f,g) = \sup\{|f(t)-g(t)| : t \in [0,1]\}$, where $f,g \in \mathrm{C}[0,1]$. Then d is a metric on the set $\mathrm{C}[0,1]$.

The fact that $\mathrm{C}[0,1]$ is a closed and bounded set, where each function is defined entirely over $[0,1]$, is relevant here. Applying the supremum norm to such a set will yield a complete metric space. However, if we change our set to $\mathrm{C}(0,1)$, we see that we may not have a complete metric space. The reason for this is that not all continuous functions are bounded, so the supremum norm cannot be defined. If we limit ourselves to the set of bounded and continuous real-valued functions on the interval $(0,1)$, then we have a complete metric space.

We now introduce a slightly more abstract notion of size, known as a norm, that is used to analyze vector spaces.

Definition 1.7. A *norm* on a vector space V over a field $\mathbb{F}$ is a function $\|\cdot\| : V \to [0,\infty)$ satisfying the following properties:
i) $\|v\| = 0$ if and only if $v = 0$ (Definiteness)
ii) $\|\lambda v\| = |\lambda| \|v\|$ for all $v \in V$ and $\lambda \in \mathbb{F}$ (Homogeneity)
iii) $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$ for all $v_1, v_2 \in V$ (Triangle inequality)

Comparing these properties with those of a metric, we see that we continue to require definiteness and the triangle inequality, but symmetry is not relevant as a condition and we additionally require that homogeneity is satisfied. A norm is a useful concept because it allows to measure the relative size of a vector, including the distance between vectors if we use the norm to induce the distance function $d(x,y) = \|x-y\|$.

Example 1.8. Let v be an n-tuple in the linear space $\mathbb{R}^n$ such that we can write this vector as $v = (v_1, v_2, \ldots, v_n)$. Then the following are norms we can define on this particular linear space that will be of particular interest later.
i) $\|v\|_1 = |v_1| + |v_2| + \ldots + |v_n|$
ii) $\|v\|_p = \{|v_1|^p + |v_2|^p + \ldots + |v_n|^p\}^{\frac{1}{p}}$, where $p \geq 1$
iii) $\|v\|_\infty = \sup\{|v_1|, |v_2|, \ldots, |v_n|\}$

It is fairly straightforward to verify that $\|v\|_1$ and $\|v\|_\infty$ are norms and that $\|v\|_p$ satisfies the definiteness and homogeneity properties of the norm, but showing that $\|v\|_p$ obeys the triangle inequality will require more technology. In particular, this result comes from Minkowski's inequality, which we shall prove in a later section. We can now describe a normed space as follows.

Definition 1.9. A *normed vector space* is a pair $(V, \|\cdot\|)$, where V is a vector space and $\|\cdot\|$ is a norm V.

Normed vector spaces and metric spaces are distinct mathematical objects, but a metric can be induced on every normed vector space. Our metric $d(x,y) = \|x-y\|$, the canonical metric, demonstrates this point. However, the converse is not true in general.

Example 1.10. For instance, we can define the following metric, which is not equivalent to a norm

$$d(x,y) = \sum_{n=1}^{\infty} 2^{-n} \frac{|x_n - y_n|}{1 + |x_n - y_n|}.$$

Just as there are incomplete metric spaces, there are also incomplete normed vector spaces. We introduce the following definition, which provides a 'nicer' structure on which to perform analysis.

Definition 1.11. A *Banach space* is a vector space equipped with a norm $(V, \|\cdot\|)$ such that V is complete. Banach spaces are also referred to as complete normed vector spaces.

Example 1.12. One of the most frequently encountered Banach spaces is the normed space $(\mathbb{R}^n, |\cdot|)$, where our norm is the absolute value. This structure induces the usual Euclidean topology on $\mathbb{R}$.

Example 1.13. Here is an enlightening non-example from [NLBE 2011]; we shall give a metric on $\mathbb{R}$ that is not Banach. Suppose for a metric $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, we define $d(x,y) = |\arctan(x) - \arctan(y)|$ for all $x, y \in \mathbb{R}$. We can see that this metric induces the usual Euclidean topology on $\mathbb{R}$, but this metric is not complete. To see this, consider the sequence $(1, 2, 3, \ldots)$ in $\mathbb{R}$. Then this sequence is a Cauchy sequence with our metric $d(x,y)$, but it clearly does not converge to a point in $\mathbb{R}$. Therefore, our d-Cauchy sequence cannot be a Cauchy sequence in $(\mathbb{R}, |\cdot|)$ because if it was, then the fact that $(\mathbb{R}, |\cdot|)$ is a Banach space would give rise to a contradiction.

Banach spaces appear many times throughout the study of functional analysis. In our case, our eventual goal in proving the Riesz-Fischer Theorem is to demonstrate that a particular set with a defined norm is a Banach space.

2. Introducing Measure Theory

The concept of measure is one that is relatively intuitive to most people; the notion of measurement is widely applicable to fields such as economics, physics, and probability theory.

Thus, given its ubiquity, it is hardly surprising that the measure of an object is one of the most fundamental concepts in Euclidean geometry and is highly intuitive. Those characteristics influenced our modern definition of measure. For instance, on the real number line $\mathbb{R}$, we instinctively believe that intervals like $[a, b]$ or $(a, b]$ have length $a-b$ and that objects like $\emptyset$ or individual points have length 0. Also, measure should be translation-invariant; practically speaking, moving an object around will not change its length. Lastly, we expect that measure is additive when we combine distinct objects. After all, in the real world, we can always find the measure of an object by adding up the measurements of smaller pieces of it. Note that from this last idea, we can make explicit another intuitive idea: if one object is a subset of another, its measure is bounded by that of the larger object.

However, formalizing these intuitions proved difficult; it is difficult to prescribe a general method for measuring an arbitrary subset of $\mathbb{R}^n$. This issue is sometimes called the 'problem of measure'. For instance, the famous Banach-Tarski paradox (which requires the Axiom of Choice) demonstrates that a ball in $\mathbb{R}^3$ can be deconstructed into a finite number of pieces and reassembled into two distinct copies of this ball.

Fortunately, there is a generally accepted solution to the problem of measure: instead of trying to measure all of the subsets of $\mathbb{R}^n$, we shall confine ourselves to measuring sets that have some regularity. After all, the sets that cannot be measured are often pathological sets, such as Vitali sets, which we shall explore later. We will now demonstrate how we can construct sets that 'have some regularity', which we will call 'measurable sets'. But first, we introduce the structure of a σ-algebra, which allows us to conduct analysis on a regular enough object that we can derive useful results.

Definition 2.1. Suppose X is a set and $\mathcal{S}$ is a set of subsets of X. Then $\mathcal{S}$ is called a *σ-algebra* on X if the following three conditions hold:
i) $\emptyset \in \mathcal{S}$
ii) if $E \in \mathcal{S}$, then $X \backslash E \in \mathcal{S}$
iii) if $E_1, E_2, \ldots$ is a sequence of elements of $\mathcal{S}$, then $\bigcup_{k=1}^{\infty} E_k \in \mathcal{S}$

There are a few properties that are worth pointing out explicitly. For example, conditions i) and ii) imply that X is always in $\mathcal{S}$. Also, from de Morgan's law, we can derive that σ-algebra is closed under countable intersection, just as it is for countable unions. More explicitly, we can apply these rules to see that if we have two sets $D, E \in \mathcal{S}$, then we can deduce that $D \cup E, D \cap E,$ and $D \backslash E \in \mathcal{S}$.

The motivation for this structure is that we aim to define a collection of sets that can serve as the domain for our measure μ. Therefore, it follows that both the empty set and the entire set should be measurable and also that knowing the measure of E and X for $E \subseteq X$ implies $X \setminus E = E^c$ should have measure

$\mu(X) - \mu(E)$. Lastly, knowing the measure of each $\mu(E_i)$ implies that we should know the measure of $\cup E_i$. In other words, we extend infinitely (but still countably) the idea that knowing the measures of two objects means we should be able to find the measure of their sum.

Example 2.2. Here are a few examples of a σ-algebra.
i) Suppose X is a set. Then trivially, $\{\emptyset, X\}$ is a possible σ-algebra on X.
ii) In $\mathbb{R}$, we can construct the σ-algebra $\{\emptyset, \mathbb{R}, \mathbb{Q}, \mathbb{R}\backslash\mathbb{Q}\}$.
iii) Suppose X is a set. Then the set of all subsets E of X such that E is countable or $X\backslash E$ is countable is a σ-algebra on X.
iv) In $\mathbb{R}$, the set of all subsets of $\mathbb{R}$ is a σ-algebra on $\mathbb{R}$.
v) In $\mathbb{R}$, the collection of Borel sets, which we are about to define, is a σ-algebra on $\mathbb{R}$. It is important enough that it deserves its own definition.

The following two definitions naturally follow our discussion of σ-algebras and will be relevant terms later on.

Definition 2.3. A *measurable space* is an ordered pair $(X, \mathcal{S})$, where X is a set and $\mathcal{S}$ is a σ-algebra on X.

Definition 2.4. Any element of $\mathcal{S}$ is called a $\mathcal{S}$*-measurable* set.

Example 2.5. For instance, if we have the ordered pair $(\mathbb{R}, \mathcal{S})$, where $\mathcal{S}$ is the set of all subsets of $\mathbb{R}$ that are countable or have a countable complement, then it follows that the set of rational numbers is $\mathcal{S}$-measurable but the set of positive real numbers is not $\mathcal{S}$-measurable.

We now formally introduce Borel sets, a key example of a σ-algebra.

Definition 2.6. The smallest σ-algebra on $\mathbb{R}$ containing all open subsets of $\mathbb{R}$ is called the collection of *Borel subsets of* $\mathbb{R}$. An element of this σ-algebra is called a *Borel set*.

Example 2.7. The reader can verify that every closed subset, countable subset, and half-open interval of $\mathbb{R}$ is a Borel set.

Definition 2.8. Suppose that $(X, \mathcal{S})$ is a measurable space. A function $f : X \to \mathbb{R}$ is called a $\mathcal{S}$*-measurable function*, or sometimes just a measurable function if the context is clear, if for every Borel set $B \subset \mathbb{R}$ we have that $f^{-1}(B) \in \mathcal{S}$.

There are many other ways in which we could define a measurable function. The symmetric structure of a σ-algebra allows for this. In particular, many analysts define a measurable function as satisfying the criteria $f^{-1}((a, \infty)) \in \mathcal{S}$ for all $a \in \mathbb{R}$. This definition will be useful going forward because it is far more practical. Sources such as [Axler2020] offer proof on the equivalences of these definitions, but it is not directly relevant to our purpose, so we shall leave it as an exercise to the reader. Intuitively, however, we can see that taking complements and then countable unions and intersections with (a, ∞) will eventually generate every Borel set, assuming $f^{-1}((a, \infty))$ is in a σ-algebra.

3. Properties of Measure

Now that we have clarified the places where our measure can live, it makes sense to change our attention to defining a measure and its properties. We give credit to [Axler2020] as the source for proofs in the following section.

Definition 3.1. Suppose X is a set and $\mathcal{S}$ is a σ-algebra on X. A *measure* on $(X, \mathcal{S})$ is a function $\mu : \mathcal{S} \to [0, \infty)$ such that, for every disjoint sequence $E_1, E_2, \dots$ of sets in $\mathcal{S}$, we have that $\mu(\emptyset) = 0$ and

$$\mu(\bigcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} \mu(E_k).$$

Note that the way we have defined measure is sometimes referred to as a 'positive' or 'nonnegative' measure. The motivation for this definition comes from attempting to extend the notion of the length of an interval. However, as the definition demonstrates, our notion is measure is generalized and can encompass notions such as length, area, volume, or more abstract quantities.

Example 3.2. Here are noteworthy examples of measures that we can define on a space.
i) The *Dirac measure* δ_c is defined on $(X, \mathcal{S})$ for $E \subset \mathcal{S}$ by

$$\delta_c(E) \begin{cases} 1, & \text{if } c \in E, \\ 0, & \text{if } c \notin E. \end{cases}$$

ii) The *counting measure* is the measure μ defined on the σ-algebra of all subsets of X such that $\mu(E) = n$ if E is a finite set with n discrete elements. If E is not finite, then $\mu(E) = \infty$.
iii) Consider the measurable space $(X, \mathcal{S})$ and $E \in \mathcal{S}$. Also, let D range over all finite subsets of E. Then consider a function $f : X \to [0, \infty)$. We can define a measure μ as

$$\mu(E) = \sup\Big\{ \sum_{x \in D} f(x) : D \subseteq E \text{ is finite} \Big\}.$$

iv) Consider the measurable space $(\mathbb{R}, \mathcal{B})$, where $\mathcal{B}$ is the σ-algebra generated by the set of Borel subsets of $\mathbb{R}$. Then we can define the *outer measure* $|A|$ of a set $A \subset \mathbb{R}$ by

$$|A| = \inf\Big\{ \sum_{k=1}^{\infty} \ell(I_k) : I_1, I_2, \dots \text{ are open intervals such that } A \subset \bigcup_{k=1}^{\infty} I_k \Big\}.$$

In this case, we define the length of an open interval $\ell(I)$ as

$$\ell(I) = \begin{cases} b - a, & \text{for some } a, b \in \mathbb{R} \text{ with } a < b, \\ 0, & \text{if } I = \emptyset, \\ \infty, & \text{if } I = (-\infty, a) \text{ or } I = (a, \infty) \text{ for some } a \in \mathbb{R}, \\ \infty, & \text{if } I = (-\infty, \infty). \end{cases} \tag{3.3}$$

Example 3.4. We now give an important non-example. Consider the measurable space $(\mathbb{R}, \mathcal{S})$, where $\mathcal{S}$ represents the set of all subsets of $\mathbb{R}$. If we apply the outer measure again as defined by (3.3), we see that we will not have a measure because, for general sets in $\mathbb{R}$, the outer measure does not satisfy countable additivity on disjoint sets, a core property of measure. We will elaborate on this point with Example 3.14 when we review Vitali sets, a classic example of a non-measurable set. It was not a problem that we used the outer measure in our previous example because we restricted our analysis to measurable sets: specifically, the Borel sets.

The concept of outer measure is foundational to measure theory; for instance, we will briefly use it to build up to the notion of a Lebesgue measurable function, which itself is relevant to deeper results in functional analysis.

Definition 3.5. A set $A \subset \mathbb{R}$ is called a *Lebesgue measurable set* if there exists a Borel set $B \subset A$ such that $|A \backslash B| = 0$. Here, we apply the outer measure as defined by (3.3).

Using this definition, we can introduce the following key concept.

Definition 3.6. A function $f : A \to \mathbb{R}$, where $A \subset \mathbb{R}$, is called a *Lebesgue measurable function* if $f^{-1}(B)$ is a Lebesgue measurable set for every Borel set $B \subset \mathbb{R}$.

We note that there are a number of equivalent definitions for a Lebesgue measurable set. However, familiarity with each of them is not central to the focus of this paper, so we do not explore them in detail here.

As we've explored, the specific measure that we choose is relevant to our analysis. If we are considering a particular measure alongside a measurable space $(X, \mathcal{S})$, then we shall call it a measure space and use the following notation.

Definition 3.7. A *measure space* is an ordered triple $(X, \mathcal{S}, \mu)$, where X is a set, $\mathcal{S}$ is a σ-algebra on X, and μ is a measure on $(X, \mathcal{S})$.

We now introduce a few foundational properties of a measure.

Corollary 3.8. *(measure preserves order). Consider the measure space $(X, \mathcal{S}, \mu)$ and $D, E \in \mathcal{S}$ are such that $D \subset E$. Then $\mu(D) \leq \mu(E)$.*

Proof. Observe that $E = D \cup (E \backslash D)$ and this is a disjoint union. Then because measure is additive, we have

$$\mu(E) = \mu(D) + \mu(E \backslash D) \geq \mu(D).$$

□

Corollary 3.9. *(measure of set difference). Consider the measure space $(X, \mathcal{S}, \mu)$. If $D, E \in \mathcal{S}$ are such that $D \subset E$ and $\mu(D) < \infty$, then $\mu(E \backslash D) = \mu(E) - \mu(D)$.*

Proof. Suppose $\mu(D) < \infty$. Then subtracting $\mu(D)$ from both sides leads to our result. □

Corollary 3.10. *(countable subadditivity). Suppose then that we have the measure space $(X, \mathcal{S}, \mu)$ and $E_1, E_2, \ldots \in \mathcal{S}$. Then*

$$\mu(\bigcup_{k=1}^{\infty} E_k) \leq \sum_{k=1}^{\infty} \mu(E_k).$$

Proof. Let $D_1 = \emptyset$ and $D_k = E_1 \cup \cdots \cup E_{k-1}$ for $k \geq 2$. Then $E_1 \backslash D_1, E_2 \backslash D_2, \ldots$ is a disjoint sequence of subsets of X whose union equals $\cup_{k=1}^{\infty} E_k$. Then, due to the countable additivity of μ, we have

$$\begin{aligned} \mu(\bigcup_{k=1}^{\infty} E_k) &= \mu(\bigcup_{k=1}^{\infty} (E_k \backslash D_k)) \\ &= \sum_{k=1}^{\infty} \mu((E_k \backslash D_k)). \end{aligned}$$

Finally, because we know that measure preserves order, we conclude

$$\sum_{k=1}^{\infty} \mu((E_k \backslash D_k)) \leq \sum_{k=1}^{\infty} \mu(E_k),$$

and the result is proven. □

Corollary 3.11. *(measure of increasing union). Suppose $(X, \mathcal{S}, \mu)$ is a measure space and the sets $E_1 \subset E_2 \subset \cdots$ are in $\mathcal{S}$. Then*

$$\mu(\bigcup_{k=1}^{\infty} E_k) = \lim_{k=1} \mu(E_k).$$

Proof. If $\mu(E_k) = \infty$ for some $k \in \mathbb{Z}^+$, then our result holds because both sides will equal ∞. Therefore, we only need to consider the case where $\mu(E_k) < \infty$ for all $k \in \mathbb{Z}^+$.
Suppose that $E_0 = \emptyset$. Then we have that

$$\bigcup_{k=1}^{\infty} E_k = \bigcup_{k=1}^{\infty} (E_j \backslash E_{j-1}).$$

Therefore, we can see that

$$\mu(\bigcup_{k=1}^{\infty} E_k = \sum_{j=1}^{\infty} \mu(E_j \backslash E_{j-1})$$

$$= \lim_{k \to \infty} \sum_{j=1}^{k} \mu(E_j \backslash E_{j-1})$$

$$= \lim_{k \to \infty} \sum_{j=1}^{k} (\mu(E_j) - \mu(E_{j-1}))$$

$$= \lim_{k \to \infty} \mu(E_k).$$

□

Corollary 3.12. *(measure of decreasing intersection). Suppose $(X, \mathcal{S}, \mu)$ is a measure space and the sets $E_1 \supset E_2 \supset \cdots$ are in $\mathcal{S}$, with $\mu(E_1) < \infty$. Then*

$$\mu(\bigcap_{k=1}^{\infty} E_k) = \lim_{k \to \infty} \mu(E_k).$$

Proof. From de Morgan's laws, we know that

$$E_1 \backslash \bigcap_{k=1}^{\infty} E_k = \bigcup_{k=1}^{\infty} (E_1 \backslash E_k).$$

Then consider an increasing sequence of sets in $\mathcal{S}$ such that

$$E_1 \backslash E_1 \subset E_1 \backslash E_2 \subset E_1 \backslash E_3$$

is an increasing sequence of sets in $\mathcal{S}$. A quick application of Corollary 3.11 tells us that

$$\mu(E_1 \backslash \bigcap_{k=1}^{\infty} E_k) = \lim_{k \to \infty} \mu(E_1 \backslash E_k).$$

Applying Corollary 3.9, we can rewrite this as

$$\mu(E_1) - \mu(\bigcap_{k=1}^{\infty} E_k) = \mu(E_1) - \lim_{k\to\infty} \mu(E_k),$$

and the result follows. □

Corollary 3.13. *(measure of union). Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $D, E \in \mathcal{S}$, with $\mu(D \cap E) < \infty$. Then*

$$\mu(D \cup E) = \mu(D) + \mu(E) - \mu(D \cap E).$$

Proof.

$$D \cup E = (D\backslash(D \cap E)) \cup (E\backslash(D \cap E)) \cup (D \cap E).$$

Then because the right hand side of this equation is a disjoint union, we have

$$\begin{aligned}\mu(D \cup E) &= \mu(D\backslash(D \cap E)) + \mu(E\backslash(D \cap E)) + \mu(D \cap E) \\ &= (\mu(D) - \mu(D \cap E)) + (\mu(E) - \mu(D \cap E)) + \mu(D \cap E) \\ &= \mu(D) + \mu(E) - \mu(D \cap E).\end{aligned}$$

□

Example 3.14. Here, we offer an example of a non-measurable set in the form of Vitali sets. We begin by showing how to construct such a set. Let us start by considering the set of real numbers in $[0, 1]$. We then partition this set using equivalence classes. Our rule for determining whether two numbers are equivalent is the following: for $x, y \in [0, 1] \subset \mathbb{R}$, we call $x \sim y$ if $x - y \in \mathbb{Q}$. Notice that this implies that all rational numbers in $[0, 1]$ are equivalent to each other.

Having dealt with the rationals, consider some irrational number $n \in [0, 1]$. The difference of an irrational and a rational is always irrational, so we can also see that no irrational number is in the same class as the rational numbers. We must then create a new equivalence class for n. Then notice that while the difference between n and another irrational may be either rational or irrational, the difference between n and any rational translate of n is always rational: $n - (n + r) \in \mathbb{Q}$. Thus, our equivalence class for n must also contain every rational translate of this number that exists in $[0, 1]$. Informally speaking, we can think of the numbers in the equivalence class for a given n as having the same irrational 'ending'. Every real number in $[0, 1]$ belongs to exactly one of the countably infinite equivalence classes that we can inductively construct by this process. Using the Axiom of Choice, we can choose one element from each equivalence class to form a Vitali set V.

We now prove by contradiction that V is non-measurable because it violates our intuitively-derived properties of measure. To begin, suppose V is measurable and recall that we can enumerate the rational numbers in $[-1, 1]$ as $r_1, r_2, \ldots$. We can then use these numbers to create rationally shifted copies of this particular Vitali set, generating the disjoint sets $V + r_1, V + r_2, \ldots$. Then consider the union of these sets. It must be a subset of $[-1, 2]$ because we are only shifting V by at most one unit in the positive and negative directions. Thus, we have

$$\bigcup_{n=1}^{\infty} (V + r_n) \subseteq [-1, 2].$$

From the fact that measure preserves order, formally defined by Corollary 3.8, we deduce that
$$\mu(\bigcup_{n=1}^{\infty}(V+r_n)) \leq \mu([-1,2]).$$
We can also see that every real number in $[0,1]$ must appear in our union, so
$$[0,1] \subseteq \bigcup_{n=1}^{\infty}(V+r_n) \subseteq [-1,2].$$
This implies that
$$\mu([0,1]) \leq \mu(\bigcup_{n=1}^{\infty}(V+r_n)) \leq \mu([-1,2]).$$
If we let our measure be the outer measure defined by (3.3) such that $\mu([a,b]) = b-a$, then we can impose bounds on the measure of the union of our Vitali sets and rewrite this as
$$1 \leq \mu(\bigcup_{n=1}^{\infty}(V+r_n)) \leq 3.$$
We know that each of our Vitali sets has the same length because they are copies of one another, so $(V+r_n)$ is identical for all $n \in \mathbb{N}$. Then the last piece of the proof is to recognize that measure is additive for disjoint subsets, so by Definition 3.1, we know the sum of the measures of the Vitali sets is bounded such that
$$1 \leq \sum_{n=1}^{\infty}(\mu(V+r_n)) \leq 3.$$
Of course, it is impossible to add a finite value infinitely many times to itself and obtain a finite value. Therefore, our result is a contradiction.

As mentioned earlier, one could argue that this set is a pathological example of a non-measurable set. Among such cases, Vitali sets are among the easiest to construct. Intuitively, though we point to flaws, this idea should build confidence in the robustness of our definition of measurability. In fact, Robert Solovay demonstrated in 1970 that the existence of a non-Lebesgue measurable set cannot be proved under the axioms of Zermelo-Frankel set theory without the Axiom of Choice [Solovay1970]. It follows that proving the outer measure–which we used to measure our Vitali sets–satisfies additivity, in general, for any countable number of disjoint sets in $\mathbb{R}$ would need to contradict the Axiom of Choice. However, by limiting our attention to measurable sets, we can avoid such issues.

4. Measure-Theoretic Convergence Theorems

We begin our study by briefly reviewing two fundamental types of convergence.

Definition 4.1. Suppose X is a set, $f_1, f_2, \ldots$ is a sequence of functions from X to $\mathbb{R}$, and f is a function also from X to $\mathbb{R}$. Then the sequence $f_1, f_2, \ldots$ *converges pointwise* on X to f if
$$\lim_{k\to\infty} f_k(x) = f(x)$$
for each $x \in X$.

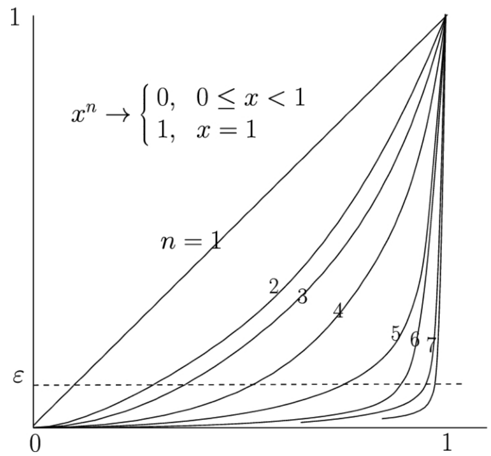

FIGURE 1. Graph of $f_n(x) = x^n$, from [Magnus2020].

Alternatively, we could say that a sequence $f_1, f_2, \dots$ converges pointwise on X to f if for each $x \in X$ and every $\epsilon > 0$, there exists $n \in \mathbb{Z}^+$ such that $|f_k(x) - f(x)| < \epsilon$ for all integers $k \geq n$.

Definition 4.2. Suppose X is a set, $f_1, f_2, \dots$ is a sequence of functions from X to $\mathbb{R}$, and f is a function from X to $\mathbb{R}$. Then the sequence $f_1, f_2, \dots$ is said to *converge uniformly* on X to f if for every $\epsilon > 0$, there exists $n \in \mathbb{Z}^+$ such that $|f_k(x) - f(x)| < \epsilon$ for all integers $k \geq n$ and all $x \in X$.

The distinction between pointwise and uniform convergence is subtle. Intuitively, if a sequence of functions will eventually converge to f for every point in the domain as $n \to \infty$, then we have pointwise convergence. However, the essence of the distinction is that a sequence of functions may converge to f at different rates for different values of the domain. In particular, if our sequence of functions converges at a uniform rate, we arrive at uniform convergence, a stronger form of pointwise convergence.

For instance, consider the sequence of functions $f_n(x) = x^n$ on the domain $[0, 1]$. The graph of these functions is given by (fig. 1). It is clear that as $n \to \infty$, all $f_n(x)$ approach 0 on the domain $[0, 1)$. However, at the point $x = 1$ itself, we see that the value of $f_n(1) = 1$ for all $n \in \mathbb{N}$ because multiplying 1 by itself any number of times is always equal to 1. Thus, it is clear that at different points along the domain, we have different rates of convergence. If a function is uniformly continuous, we should be able to construct an ϵ-radius around a given $f_n(x)$ that encapsulates the entire function, without worrying about where we are along the domain. As (fig. 1) demonstrates, it is possible to construct such a radius that does not include the point where $f_n(1) = 1$, so we cannot have uniform convergence. Equivalently stated, if the limit function is discontinuous, then we cannot have uniform convergence. In fact, as sources like [Pugh2015] explain, the uniform limit of continuous functions is continuous. Evidently, pointwise convergence is not a particularly strong concept.

Now that we have built a rigorous foundation in our understanding of convergence, we will begin to consider more complicated theorems that build on these ideas. After all, the convergence theorems in Lebesgue theory are integral to the study of functional analysis. One of the advantages of Lebesgue theory over the theory of Riemann integration is that many features are preserved under limits. The convergence theorems that we explore in this section are built on this idea. In contrast, Riemann integrable functions are not closed under pointwise limits, a fact that is well demonstrated by sources like [Axler2020], which makes them more difficult to generalize and apply. For example, if we have some sequence of Riemann integrable functions $\{f_n\} \to f$, then the limit is not necessarily Riemann integrable. Assuming that f_n for all n is measurable, however, allows us to conclude that f is measurable. Therefore, these results will be key to performing mathematical analysis, where we often consider the limits of functions or sets.

As mentioned, one of the major advantages of Lebesgue integration over Riemann integration is that the pointwise limit of $\mathcal{S}$-measurable functions is itself $\mathcal{S}$-measurable. We shall now prove that this is true.

Theorem 4.3. *(measure space is closed under pointwise limits). Suppose that $(X, \mathcal{S})$ is a measurable space and we have a sequence of $\mathcal{S}$-measurable functions $f_1, f_2, \ldots$ from $X \to \mathbb{R}$. Also, assume that $\lim_{k\to\infty} f_k(x)$ exists for all $x \in X$. We define $f : X \to \mathbb{R}$ by*

$$f(x) = \lim_{k\to\infty} f_k(x).$$

Then f is an $\mathcal{S}$-measurable function.

Proof. We follow [Axler2020] and demonstrate that

$$f^{-1}((a,\infty)) = \bigcup_{j=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{k=m}^{\infty} f_k^{-1}((a+\frac{1}{j},\infty)). \tag{4.4}$$

In particular, to prove that these sets are equal, we shall demonstrate that each is a subset of the other. We begin by assuming $x \in f^{-1}((a,\infty))$; in other words, we have $\{x \in X : f(x) > a\}$. Also, recall that $\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$. These facts allow us to conclude that there is some $j \in \mathbb{Z}^+$ such that $f(x) > a + \frac{1}{j}$. From the definition of a limit, we can now see that there must be some $m \in \mathbb{Z}^+$ such that $f_k(x) > a + \frac{1}{j}$ for all $k \geq m$. Thus, it follows that x is in the right hand side of (4.4). We now assume x is a member of the right hand side of (4.4). Therefore, we have that there must exist $j, m \in \mathbb{Z}^+$ such that $f_k(x) > a + \frac{1}{j}$ for all $k \geq m$. We can then apply the limit as $k \to \infty$ to see that $f(x) \geq a + \frac{1}{j} > a$. Thus, x is a member of the left hand side of our equation. We've thus proven that each side is a subset of the other, so we must have equality. That (4.4) is true implies that $f^{-1}((a,\infty)) \in \mathcal{S}$, meaning it is $\mathcal{S}$-measurable. □

Now that we have established the closure of a measure space under pointwise limits, the below corollary readily follows; we credit [Bartle1966] for the proofs in the remainder of this section.

Corollary 4.5. *If (f_n) is a sequence in $(X, \mathcal{S}, \mu)$ which converges to f on X, then $f \in (X, \mathcal{S}, \mu)$.*

Proof. Our result follows from Theorem 4.3. In this case, $f(x) = \lim f_n(x) = \liminf f_n(x)$. □

Before we continue our discussion of limits, we need to carefully reintroduce the functions on which we will apply these theorems, so we shall start on a basic level. We will introduce characteristic functions and use them to construct simple functions. When we consider many results that might apply for L_p spaces, we will often find ourselves asking whether the result is true for simple functions. In fact, as we will demonstrate by later, we can express every measurable function as a limit of simple functions, so this simplification will enhance our study.

Another simplification that we often make is to require our functions to be nonnegative. This assumption allows us to apply powerful results, like the Monotone Convergence Theorem, which require that functions are nonnegative. We can then extend our result to include negative functions because every real-valued function can be interpreted as the difference of two nonnegative functions, as justified by the definition we will explore next. Thus, results that hold for nonnegative functions also tend to hold for all real-valued functions.

Definition 4.6. If we have some function f on a set X to $\mathbb{R}$, we can define the nonnegative functions

$$f^+(x) = \sup\{f(x), 0\}, \text{ and } f^-(x) = \sup\{-f(x), 0\}, \forall x \in X,$$

where we call f^+ is called the *positive part of* f and f^- is called the *negative part of* f.

This definition is relatively straightforward, but it is quite powerful. In particular, it readily follows that $f = f^+ - f^-$ and $|f| = f^+ + f^-$. The integrals of these functions shall follow a similar structure. These results imply that a function f is measurable if and only if f^+ and f^- are measurable.

We now introduce another elementary function: the characteristic function.

Definition 4.7. The *characteristic function* χ_E, sometimes called an indicator function, is defined by

$$\chi_E(x) \begin{cases} = 1, & x \in E, \\ = 0, & x \notin E. \end{cases}$$

We can use this idea to develop the following function. In the following definitions let $c_1, \ldots, c_n$ denote an arbitrary sequence of arbitrary real numbers.

Definition 4.8. A *step function* is a real-valued function $\Psi(x)$ which is a finite linear combination of characteristic functions of intervals, defined by

$$\Psi(x) = \sum_{j=1}^{n} c_j \chi_{E_j},$$

where E_j represents an interval.

Definition 4.9. We can then use the prior construction to define the *integral of a step function* $\Psi(x)$ to be

$$\int \Psi(x) = \sum_{j=1}^{n} c_j (b_j - a_j),$$

where E_j represents an interval with endpoints a_j and b_j.

The prior definition is relevant because we can use step functions to define the Riemann integral as the limit of the integrals of the step functions that approximate our target function. For instance, the lower Riemann integral of a function $f(x)$ would be defined as the supremum of the integrals of all step functions $\Psi(x)$ such that $\Psi(x) \leq f(x)$ for all $x \in [a, b]$; we would also let $\Psi(x) = 0$ for $x \notin [a, b]$.

In constructing the Lebesgue integral, we follow a parallel process, but with some slight differences. Namely, instead of using step functions, which are defined to work with intervals of $\mathbb{R}$, we shall generalize our sets to be any $E \in \mathcal{S}$. We then arrive at the following definition.

Definition 4.10. A *simple function* is a real-valued function ψ which is a finite linear combination of characteristic functions of sets belonging to $\mathcal{S}$. That is, a simple function can be represented as

$$\psi(x) = \sum_{j=1}^{n} c_j \chi_{E_j}$$

where χ_{E_j} is the characteristic function of a set E_j in $\mathcal{S}$. We can define a *unique standard representation* for a simple measurable function such that the a_j are distinct and the E_j are distinct nonempty subsets of X and $X = \bigcup_{j=1}^{n} E_j$.

Generalizations of step functions, simple functions are a foundation for the rest of our study. Some authors may equivalently state that step and simple functions are defined by having a finite number of output values.

Definition 4.11. We can now define the *integral of a simple function* $\psi(x)$ to be

$$\int \psi(x) = \sum_{j=1}^{n} c_j \mu(E_j),$$

where $\mu(E_j)$ denotes the measure of a set in $\mathcal{S}$. Canonically, we now consider the values for c_j to be distinct and the sets E_i to be disjoint.

Note that our definition does not explicitly employ a characteristic function for the set E_j. Therefore, the curious reader may look to [Royden2018] for evidence of the fact that if we have disjoint sets E_i, then

$$\psi = \sum_{i=1}^{n} c_i \chi_{E_i} \text{ on } E \text{ implies that } \int \psi(x) = \sum_{i=1}^{n} c_i \mu(E_i).$$

We shall apply this fact in the following pages.

As we mentioned, to simplify our analysis, we can limit our attention to nonnegative functions. When relevant, we will adjust our notation to reflect the fact that we are working with nonnegative measurable functions.

Definition 4.12. In particular, we shall denote that collection of all nonnegative $\mathcal{S}$-measurable functions mapping X to the extended real number line by $(X, \mathcal{S}, \mu)^+$. A function $f \in (X, \mathcal{S}, \mu)^+$ is a nonnegative $\mathcal{S}$-measurable function.

Now that we have laid out our terms, we shall introduce the properties of functions in a measure space, starting with those of simple functions.

Theorem 4.13. *If $\psi \in (X, \mathcal{S}, \mu)^+$, then there exists a sequence (ψ_n) in $(X, \mathcal{S}, \mu)^+$ such that*
a) Each ψ_n is a simple measurable function,

b) $0 \leq \psi_n(x) \leq \psi_{n+1}(x)$ *for* $x \in X, n \in \mathbb{N}$*, and*
c) $\psi(x) = \lim \psi_n(x)$ *for each* $x \in X$*.*

Proof. Let n be a fixed natural number. If $k = 0, 1, \ldots, n2^n - 1$, let E_{kn} be the set

$$E_{kn} = \{x \in X : k2^{-n} \leq \psi(x) < (k+1)2^{-n}\},$$

and if $k = n2^n$, we can let E_{kn} be the set $x \in X : \psi(x) \geq n$. Notice that the sets $E_{kn} : k = 0, 1, \ldots, n2^n$ are disjoint, in $\mathcal{S}$, and have union equal to X. If we let $\psi_n = k2^{-n}$ on E_{kn}, then ψ_n belongs to $(X, \mathcal{S}, \mu)^+$. Then properties (a), (b), and (c) can quickly be demonstrated. □

We now demonstrate that simple measurable functions obey a few basic properties.

Theorem 4.14. *(linearity for bounded measurable functions). If* ϕ *and* ψ *are simple functions in* $(X, \mathcal{S}, \mu)^+$ *and we have a scalar* $a \geq 0$*, then*

$$\int a\phi \; d\mu = a \int \phi \; d\mu.$$

Also,

$$\int (\phi + \psi) \; d\mu = \int \phi \; d\mu + \int \psi \; d\mu.$$

Proof. We begin by demonstrating that our integral is linear under scalar multiplication. If $a = 0$, then $a\phi$ disappears and our proof is trivial. If $a > 0$, then $a\phi$ is in $(X, \mathcal{S}, \mu)^+$ and has the standard representation

$$a\phi = \sum_{j=1}^{n} ab_j \chi_{E_j}.$$

Thus, we can say that

$$\begin{aligned}\int a\phi \; d\mu &= \sum_{j=1}^{n} ab_j \mu(E_j)\\ &= a \sum_{j=1}^{n} b_j \mu(E_j)\\ &= a \int \phi d\mu.\end{aligned}$$

Next, we demonstrate that our integral is linear under addition. We shall let our simple functions assume the following standard representations:

$$\phi = \sum_{j=1}^{n} a_j \chi_{E_j}, \text{ and } \psi = \sum_{k=1}^{m} b_k \chi_{F_k},$$

where we have disjoint nonempty sets and unique scalars a_j and b_k. Therefore, it follows that the function $\phi + \psi$ can be represented as

$$\phi + \psi = \sum_{j=1}^{n} \sum_{k=1}^{m} (a_j + b_k) \chi_{E_j \cap F_k}.$$

However, we should note that this particular representation of $\phi + \psi$, as a linear combination of the disjoint sets $E_j \cap F_k$, is actually not necessarily the standard representation for $\phi + \psi$; in particular, it is possible that the scalar values $(a_j + b_k)$ are not unique, which is clearly incompatible with our definition of standard representation. Thus, let us take the distinct numbers in the set of $\{a_j + b_k\}$ and represent these as c_h for some $h \in \mathbb{N}$. Then let G_h be the union of those sets $E_j \cap F_k \neq \emptyset$ such that $a_j + b_k = c_h$. Therefore, we have the summation over all j and k

$$\mu(G_h) = \sum_h \mu(E_j \cap F_k).$$

Then, we have that the standard representation of $\phi + \psi$ can be written as

$$\phi + \psi = \sum_{h=1}^{p} c_h \chi_{G_h}.$$

Thus, we can deduce that

$$\int (\phi + \psi)\ d\mu = \sum_{h=1}^{p} c_h \mu(G_h)$$

$$= \sum_{h=1}^{p} \sum_h c_h \mu(E_j \cap F_k)$$

$$= \sum_{h=1}^{p} \sum_h (a_j + b_k) \mu(E_j \cap F_k)$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{m} (a_j + b_k) \mu(E_j \cap F_k)$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{m} a_j \mu(E_j \cap F_k) + \sum_{j=1}^{n} \sum_{k=1}^{m} b_k \mu(E_j \cap F_k).$$

Next, we can apply the fact that X is the union of both of the disjoint families $\{E_j\}$ and $\{F_k\}$ to see that

$$\mu(E_j) = \sum_{k=1}^{m} \mu(E_j \cap F_k), \text{ and that } \mu(F_k) = \sum_{j=1}^{n} \mu(E_j \cap F_K).$$

It follows that

$$\int (\phi + \psi)\ d\mu = \sum_{j=1}^{n} a_j \mu(E_j) + \sum_{k=1}^{m} b_k \mu(F_k).$$

Thus, our result holds. □

Now that we have demonstrated that integration satisfies linearity, we can see that our earlier requirement that the sets E_i are disjoint is not necessary any further.

Theorem 4.15. *If ψ is a simple function in $(X, \mathcal{S}, \mu)^+$ and λ is defined for E in $\mathcal{S}$, then*

$$\lambda(E) = \int \psi \chi_E\ d\mu,$$

so λ is a measure on $(X, \mathcal{S}, \mu)^+$.

Proof. We begin by noticing that

$$\psi\chi_E = \sum_{j=1}^{n} a_j \chi_{E_j \cap E}.$$

Therefore, from Theorem 4.14, we can apply induction to demonstrate that

$$\lambda(E) = \int \psi\chi_E \, d\mu = \sum_{j=1}^{n} a_j \int \chi_{E_j \cap E} \, d\mu$$
$$= \sum_{j=1}^{n} a_j \mu(E_j \cap E).$$

As the reader can verify, the mapping $E \to \mu(E_j \cap E)$ is a measure, so we can say that we have expressed λ as a linear combination of measures on $\mathcal{S}$. Then it can also be verified that λ is a measure on $\mathcal{S}$. □

Then we use the integral of a simple function to define the following.

Definition 4.16. We define the *Lebesgue integral* of a nonnegative function f as the common value of the supremum of the integrals of all simple functions that are less than or equal to the function for a given x and the infimum of the integrals of all simple functions that are greater or equal to the function for a given x. Symbolically, we have

$$\int f \, d\mu = \sup\Big\{\int_X \psi \, d\mu : \psi \le f\Big\},$$
$$\text{and} \quad \int f \, d\mu = \inf\Big\{\int_X \phi \, d\mu : f \le \phi\Big\},$$

where ψ, ϕ are simple functions and f is a nonnegative function on X.

Here are two key properties of the Lebesgue integral.

Theorem 4.17. *(homogeneity of nonnegative measurable functions). (a) If $f, g \in (X, \mathcal{S}, \mu)^+$ and $f \le g$, then*

$$\int f \, d\mu \le \int g \, d\mu. \tag{4.18}$$

(b) If $f \in (X, \mathcal{S}, \mu)^+$, we have $E, F \in \mathcal{S}$, and $E \subseteq F$, then

$$\int_E f \, d\mu \le \int_F f \, d\mu.$$

Proof. (a) If ψ is a simple function in $(X, \mathcal{S}, \mu)^+$ such that $0 \le \psi \le f$, then $0 \le \psi \le g$. Thus, (4.18) must hold.

(b) We have $f_{\chi_E} \le f_{\chi_F}$, so (b) follows from (a). □

The following, together with Fatou's Lemma and the Dominated Convergence Theorem, are celebrated results in Lebesgue theory. It is natural to wonder whether operators such as $\lim_{n\to\infty}$ and $\int$ commute. Though they do not in general, the Monotone and Dominated Convergence Theorems allow us to interchange these operators given a certain criteria. As we will see later, Fatou's Lemma answers this question if we cannot make any additional assumptions about the functions in question.

Theorem 4.19. *(Monotone Convergence Theorem).*[2] *Suppose $(X, \mathcal{S}, \mu)^+$ is a measure space and $0 \leq f_2 \leq f_2 \leq \dots$ is an increasing sequence of $\mathcal{S}$-measurable functions in this space. Define $f : X \to [0, \infty)$ by*

$$f(x) = \lim_{k \to \infty} f_k(x).$$

Then

$$\lim_{k \to \infty} \int f_k \, d\mu = \int f \, d\mu \tag{4.20}$$

Proof. From Corollary 4.5, we know that f is a measurable function. We can apply Theorem 4.17 to the fact that $f_n \leq f_{n+1} \leq f$ to demonstrate that

$$\int f_n \, d\mu \leq \int f_{n+1} \, d\mu \leq f \, d\mu$$

for all $n \in \mathbb{N}$. Therefore, we have that

$$\lim \int f_n \, d\mu \leq \int f \, d\mu.$$

We then aim to establish the opposite inequality. Let a be a real number that satisfies $0 < a < 1$ and let ψ be a simple measurable function that satisfies $0 \leq \psi \leq f$. Also, let

$$A_n = \{x \in X : f_n(x) \geq a\psi(x)\}$$

such that $A_n \in \mathcal{S}, A_n \subseteq A_{n+1}$ and $X = \cup A_n$. From Theorem 4.17, we have that

$$\int_{A_n} a\psi \, d\mu \leq \int_{A_n} f_n \, d\mu \leq \int f_n \, d\mu. \tag{4.21}$$

The sequence (A_n) is monotone increasing and has union X, so it follows from Theorem 4.15 and Corollary (3.11 that

$$\int \psi \, d\mu = \lim \int_{A_n} \psi \, d\mu.$$

Thus, when we take the limit of (4.21) with respect to n, we derive

$$a \int \psi \, d\mu \leq \lim \int f_n \, d\mu.$$

This holds for all a with $0 < a < 1$, so we deduce that

$$\int \psi \, d\mu \leq \lim \int f_n \, d\mu.$$

We know that ψ is a simple function in $(X, \mathcal{S}, \mu)^+$ that satisfies $0 \leq \psi \leq f$, thus we conclude that

$$\int f \, d\mu = \sup_{\psi} \int \psi \, d\mu \leq \lim \int f_n \, d\mu.$$

When we combine this result with the opposite inequality, we obtain (4.20). □

[2] In many ways, continuity and measurability are analogous notions. The curious reader may study Dini's Theorem, a result that is quite similar to the Monotone Convergence Theorem. As given by [Rudin1976], it holds that if a monotone sequence of continuous functions converges pointwise on a compact space and if the limit function is also continuous, then the convergence is uniform.

We now introduce Fatou's Lemma. Notice that, in contrast to the Monotone and Dominated Convergence Theorems, we do not impose any special requirements on our functions; it is as close as we'll be able to get to interchange the limit and integration operations without such requirements.

Theorem 4.22. *(Fatou's Lemma). Suppose that* $(X, \mathcal{S}, \mu)^+$ *is a measure space and* $f_1, f_2, \ldots$ *is a sequence of functions in this space. Define a function* $f : X \to (-\infty, \infty)$ *by* $f(x) = \liminf_{k\to\infty} f_k(x)$*. Then*

$$\int f \; d\mu \leq \liminf_{k\to\infty} \int f_k \; d\mu.$$

Proof. Let $g_m = \inf\{f_m, f_{m+1}, \ldots\}$ so that $g_m \leq f_n$ whenever $m \leq n$. Therefore,

$$\int g_m \; d\mu \leq \int f_n \; d\mu, \text{ where } m \leq n,$$

so that

$$\int g_m \; d\mu \leq \liminf \int f_n \; d\mu.$$

The sequence (g_m) is increasing and converges to $\liminf f_n$, so the Monotone Convergence Theorem implies

$$\begin{aligned} \int (\liminf f_n) \; d\mu &= \lim \int g_m \; d\mu \\ &\leq \liminf \int f_n \; d\mu. \end{aligned}$$

□

In addition to Fatou's Lemma, the Monotone Convergence Theorem allows us to derive other fundamental results. For instance, we have already proven that integration is linear and monotonic for bounded measurable functions Theorem 4.14, but we are now equipped to demonstrate that these properties can be extended to any nonnegative measurable functions. We have the following theorem.

Theorem 4.23. *(linearity of scalar multiplication for nonnegative measurable functions). If* $f \in (X, \mathcal{S}, \mu)^+$ *and* $a \geq 0$*, then* $af \in (X, \mathcal{S}, \mu)^+$ *and*

$$\int af \; d\mu = a \int f \; d\mu.$$

Proof. The proof is trivial for $a = 0$. When $a > 0$, we can let (f_n) be a sequence of simple functions in $(X, \mathcal{S}, \mu)^+$ that is monotone increasing and converging to f on X. Then (af_n) is a monotone sequence converging to af. We can then apply Theorem 4.14 and the Monotone Convergence Theorem to obtain

$$\begin{aligned} \int af \; d\mu &= \lim \int af_n \; d\mu \\ &= a \lim \int f_n \; d\mu \\ &= a \int f \; d\mu. \end{aligned}$$

□

Theorem 4.24. *(linearity of addition for nonnegative functions). If $f, g \in (X, \mathcal{S}, \mu)^+$, then $f+g$ belongs to $(X, \mathcal{S}, \mu)^+$ and*

$$\int (f+g)\ d\mu = \int f\ d\mu + \int g\ d\mu.$$

Proof. Consider the monotone increasing sequences (ψ_n) and (ϕ_n), converging to f and g, respectively. Then $(\psi_n + \phi_n)$ is a monotone increasing sequence that converges to $f+g$. It then follows from Theorem 4.14 and the Monotone Convergence Theorem that

$$\begin{aligned}\int (f+g)\ d\mu &= \lim \int (\psi_n + \phi_n)\ d\mu \\ &= \lim \int \psi_n\ d\mu + \lim \int \phi\ d\mu \\ &= \int f\ d\mu + \int g\ d\mu.\end{aligned}$$

□

Finally, we come to another strength of Lebesgue theory: absolute integrability. Recall that in evaluating a Riemann integral, we may not be able to compute the improper Riemann integral of a function. However, this is not an issue here, as we shall see.

Theorem 4.25. *(absolute integrability). A measurable function f is integrable if and only if $|f|$ is integrable. In particular,*

$$|\int f\ d\mu| \leq \int |f|\ d\mu.$$

Proof. By definition, f is integrable if and only if f^+ and f^- belong to $(X, \mathcal{S}, \mu)^+$ and have finite integrals. Since $|f|^+ = |f| = f^+ + f^-$ and $|f|^- = 0$, our statement follows from Theorems 4.17 and 4.24. Furthermore,

$$\begin{aligned}|\int f\ d\mu| &= |\int f^+\ d\mu - \int f^-\ d\mu| \\ &\leq \int f^+\ d\mu + \int f^-\ d\mu \\ &= \int (f^+ + f^-)\ d\mu \\ &= \int |f|\ d\mu.\end{aligned}$$

□

We have demonstrated before that homogeneity and linearity are fundamental properties that hold for more special types of functions. We now generalize these to all measurable functions.

Theorem 4.26. *(homogeneity of measurable functions). If f is measurable, g is integrable, and $|f| \leq |g|$, then f is integrable, and*

$$\int |f|\ d\mu \leq \int |g|\ d\mu.$$

Proof. This result follows from Theorems 4.17 and 4.25. □

Theorem 4.27. *(linearity of addition for measurable functions). The sum $f + g$ of integrable functions is also integrable. That is,*

$$\int (f + g)\ d\mu = \int f\ d\mu + \int g\ d\mu.$$

Proof. We follow [Bartle1966]. First, note that since we take f and g as functions in our measure space, it follows that $|f|$ and $|g|$ are also in this set. Then by our definition of a metric space, $|f + g| \leq |f| + |g|$, so we deduce from Theorems 4.24 and 4.26 that $f + g$ is in the measure space. Next, we see that

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

This implies that

$$\begin{aligned}\int (f + g)\ d\mu &= \int (f^+ + g^+)\ d\mu - \int (f^- + g^-)\ d\mu \\ &= \int f^+\ d\mu - \int f^-\ d\mu + \int g^+\ d\mu - \int g^-\ d\mu \\ &= \int f\ d\mu + \int g\ d\mu.\end{aligned}$$

□

Theorem 4.28. *(linearity of scalar multiplication for measurable functions). A constant multiple of an integrable function af is also integrable. That is,*

$$\int af\ d\mu = a \int f\ d\mu.$$

Proof. If $a = 0$, then $af = 0$ everywhere, such that

$$\int af\ d\mu = 0 = a \int f\ d\mu.$$

For $a > 0$, we know that $(af)^+ = af^+$ and $(af)^- = af^-$. Then

$$\begin{aligned}\int af\ d\mu &= \int af^+\ d\mu - \int af^-\ d\mu \\ &= a(\int f^+\ d\mu - \int f^-\ d\mu) \\ &= a \int f\ d\mu.\end{aligned}$$

Lastly, suppose consider the case when we have a negative scalar $-a$ for $a > 0$. then $(-af)^+ = -af^-$ and $(-af)^- = -af^+$. Then

$$\begin{aligned}\int -af\ d\mu &= \int -af^+\ d\mu - \int -af^-\ d\mu \\ &= -a(\int f^+\ d\mu - \int f^-\ d\mu) \\ &= -a \int f\ d\mu.\end{aligned}$$

□

Before we introduce the Dominated Convergence Theorem, we should clarify the meaning of 'almost every', often abbreviated in mathematical literature as a.e. We have the following definition.

Definition 4.29. Consider the measure space $(X, \mathcal{S}, \mu)$. A set $E \in \mathcal{S}$ is said to contain *almost every* element of X if $\mu(X \backslash E) = 0$. We can also say that two functions f and g are *equal μ-almost everywhere* if the disparity between them has measure zero.

A classic example is that the outer measure of the rational numbers $|\mathbb{Q}| = 0$, so almost every real number is irrational; to be clear, the measure of the rational numbers is zero because, on the real line, each rational is an isolated point–these are defined to have zero measure.

We are now equipped to interpret the following result.

Theorem 4.30. *(Dominated Convergence Theorem). Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f : X \to (-\infty, \infty)$ is $\mathcal{S}$-measurable, and $f_1, f_2, \ldots$ are $\mathcal{S}$-measurable functions from X to $(-\infty, \infty)$ such that*

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

for almost every $x \in X$. If there exists an $\mathcal{S}$-measurable function $g : X \to (-\infty, \infty)$ such that

$$\int g \, d\mu < \infty \text{ and } |f_n(x)| \leq g(x)$$

for every $n \in \mathbb{Z}^+$ and almost every $x \in X$, then

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

Proof. We can redefine the functions f_n, f on a set of measure 0. Then we can assume that convergence occurs on the entirety of X. From Theorem 4.26, we deduce that f is integrable. We have that $g + f_n \geq 0$, so we apply Fatou's Lemma and Theorem 4.27 to obtain

$$\begin{aligned}\int g \, d\mu + \int f \, d\mu = \int (g + f) \, d\mu &\leq \liminf \int (g + f_n) \, d\mu \\ &= \liminf (\int g \, d\mu + \int f_n \, d\mu) \\ &= \int g \, d\mu + \liminf \int f_n \, d\mu.\end{aligned}$$

Thus,

$$\int f \, d\mu \leq \liminf \int f_n \, d\mu. \tag{4.31}$$

Since $g - f_n \geq 0$, we can apply Fatou's Lemma and Theorem 4.27 again to see that

$$\begin{aligned}\int g \, d\mu - \int f \, d\mu = \int (g - f) \, d\mu &\leq \liminf \int (g - f_n) \, d\mu \\ &= \int g \, d\mu - \limsup \int f_n \, d\mu,\end{aligned}$$

and it follows that

$$\limsup \int f_n \, d\mu \leq \int f \, d\mu. \tag{4.32}$$

Combining (4.31) and (4.32), we conclude that

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

□

5. Lebesgue Spaces

In constructing Lebesgue spaces, our eventual goal is to incorporate the structure of a Banach space on the set of all integrable functions on a measure space $(X, \mathcal{S}, \mu)$. We begin by introducing an important preliminary definition.

Definition 5.1. Suppose that $(X, \mathcal{S}, \mu)$ is a measure space, $0 < p < \infty$, and $f : X \to \mathbb{F}$ is $\mathcal{S}$-measurable. Then the *p-norm* of f is denoted by $\|f\|_p$ and is defined by

$$\|f\|_p = (\int |f|^p \; d\mu)^{\frac{1}{p}}.$$

We are now equipped to define a general Lebesgue space.

Definition 5.2. Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p \leq \infty$. The *Lebesgue space* $L^p(\mu)$ (or L^p) is defined to be the set of $\mathcal{S}$-measurable functions $f : X \to \mathbb{F}$ such that $\|f\|_p < \infty$.

The cases where we let $p = 1$ or $p = \infty$ are usually given special attention. The definition of L^1 follows from our prior definitions. Importantly, $\|f\|_1$ may not be not a true norm, by some standards. It might be called a semi-norm because it may not satisfy definiteness: this comes as a result of the fact that the Lebesgue measure can be zero for nonempty sets. For instance, we might have $\|\chi_E\|_1 = 0$ for $\chi_E \neq 0$. We shall leave further exploration of this as an exercise. For our purposes, we shall satisfy definiteness by claiming that our norm is zero if and only if $f(x) = 0$ almost everywhere. Therefore, it is important to note than any element of this space is actually an equivalence class of functions that are equal μ-almost everywhere.

To rigorously define the case when $p = \infty$, we will need to revisit our definition for the p-norm.

Definition 5.3. The *essential supremum* of f, $\|f\|_\infty$, is defined by

$$\|f\|_\infty = \inf\{t > 0 : \mu(\{x \in X : |f(x)| > t\} = 0\}.$$

It remains to show that $\|f\|_p$ satisfies the properties of a norm on L^∞. Now that we have clarified how the p-norm would be defined, the construction for L^∞ follows from Definition 5.2.

Example 5.4. We also have special notation for L_p spaces when μ is counting measure on $\mathbb{Z}^+$, the set L^p is denoted by ℓ^p. Therefore, if $0 < p < \infty$, then

$$\ell^p = \{(a_1, a_2, \ldots) : \text{ each } a_k \in \mathbb{F} \text{ and } \sum_{k=1}^{\infty} |a_k|^p < \infty\}.$$

Also,

$$\ell^\infty = \{(a_1, a_2, \ldots) : \text{ each } a_k \in \mathbb{F} \text{ and } \sup_{k \in \mathbb{Z}^+} |a_k| < \infty\}.$$

It is relatively manageable to prove that all of the properties of a norm hold for the p-norm, with the notable exception of the triangle inequality. In fact, the next section is entirely devoted to proving this holds, and thus the p-norm is a true norm.

Our eventual goal is to demonstrate that L_p is a Banach space, so in addition to showing the norm holds, we will need to demonstrate that L_p is a complete vector space.

We will now give short proofs that demonstrate L^p is closed under addition and multiplication, which will help us show that L_p is a vector space. Soon after, we will prove Minkowski's inequality, the triangle inequality for our norm, which is a stronger condition to the following result when $p \geq 1$.

Theorem 5.5. *(L^p is closed under addition). Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p < \infty$. Then*

$$\|f + g||_p^p \leq 2^p(\|f||_p^p + \|g||_p^p)$$

for all $f, g \in L^p$ and all $a \in \mathbb{F}$.

Proof. Suppose $f, g \in L^p$. If $x \in X$, then

$$\begin{aligned} |f(x) + g(x)|^p &\leq (|f(x)| + |g(x)|)^p \\ &\leq (2\max\{|f(x)|, |g(x)|\})^p \\ &\leq 2^p(|f(x)|^p + |g(x)|^p). \end{aligned}$$

We can then integrate both sides of the above inequality with respect to μ, which gives the desired inequality. This result implies that if $\|f||_p < \infty$ and $\|g||_p < \infty$, then $\|f + g||_p < \infty$. □

Theorem 5.6. *(L^p is closed under scalar multiplication). Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $0 < p < \infty$. Then*

$$\|af||_p = |a|\|f||_p$$

for all $f, g \in L^p$ and all $a \in \mathbb{F}$.

Proof. The result follows from the definition of $\| \cdot ||_p$. □

Theorem 5.7. *(L^p is a vector space). For all $f \in L^p$ and all $a \in \mathbb{F}$, with the usual operations of addition and scalar multiplication of functions, L^p is a vector space.*

Proof. L^p contains the constant function 0, is closed under addition and scalar multiplication, and is a subspace of $\mathbb{F}^X$. Thus it is a vector space. □

6. Constructing Minkowski's Inequality

Definition 6.1. The *conjugate* of a number $p \in (1, \infty)$ is the number $q = \frac{p}{(p-1)}$, which is the unique number $q \in (1, \infty)$ for which

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Note that the conjugate of 1 is defined to be ∞ and the conjugate of ∞ is defined to be 1.

Theorem 6.2. *(Young's inequality). Suppose $1 \leq p \leq \infty$. Then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

for all $a, b \geq 0$.

Proof. We follow [Axler2020]. Let $b > 0$. Then we define some function $f : (0, \infty) \to \mathbb{R}$ by

$$f(a) = \frac{a^p}{p} + \frac{b^q}{q} - ab.$$

We can then differentiate this function with respect to a to obtain that $f'(a) = a^{p-1} - b$. Then, using the fact that $pq = p + q$ by the definition of conjugate, we can optimize this function to find that f attains its global minimum when $a = b^{\frac{1}{p-1}}$ and is therefore increasing on $(b^{\frac{1}{p-1}}, \infty)$. We see that at the point $f(b^{\frac{1}{p-1}}) = 0$, so it must be true that $f(a) \geq 0$ when $a \in (0, \infty)$ and the result follows. □

Theorem 6.3. *(Hölder's inequality). Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $1 \leq p \leq \infty$, and $f, g : X \to \mathbb{F}$ are $\mathcal{S}$-measurable. Then*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Proof. Suppose that $1 < p < \infty$. We leave the cases where $p = 1$ and $p = \infty$ as exercises for the reader.

First, consider the special case where $\|f\|_p = \|g\|_q = 1$. From Young's inequality, we have that

$$|f(x)g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}$$

for all $x \in X$. We can integrate both sides of this inequality with respect to μ to demonstrate that $\|fg\|_1 \leq 1 = \|f\|_p \|g\|_q$, which completes the proof in this special case.

If $\|f\|_p = 0$ or $\|g\|_q = 0$, then $\|fg\|_1 = 0$ and the desired inequality holds. We can also see that the inequality will hold if $\|f\|_p = \infty$ or $\|g\|_q = \infty$. Thus, we assume that $0 < \|f\|_p < \infty$ and $0 < \|g\|_q < \infty$.

We now define $\mathcal{S}$-measurable functions $f_1, g_1 : X \to \mathbb{F}$ by

$$f_1 = \frac{f}{\|f\|_p} \text{ and } g_1 = \frac{g}{\|g\|_q}.$$

Then $\|f_1\|_p = 1$ and $\|g_1\|_q = 1$. From the result of our special case, we then have $\|f_1 g_1\|_1 \leq 1$ which implies that $\|fg\| \leq \|f\|_p \|g\|_q$ □

Theorem 6.4. *(Minkowski's inequality). Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $1 \leq p \leq \infty$, and $f, g \in L^p$. Then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof. It follows from Theorem 4.27 that our statement is satisfied when $p = 1$. Therefore, suppose that $p > 1$. Then the sum $f + g$ must be measurable. We also know that

$$|f + g|^p \leq (2 \sup\{|f|, |g|\})^p \leq 2^p \{|f|^p + |g|^p\},$$

so it follows from Theorems 4.26 and (4.27) that $f + g \in L_p$. Furthermore, we have that

$$|f + g|^p = |f + g||f + g|^{p-1} \leq |f||f + g|^{p-1} + |g||f + g|^{p-1}. \tag{6.5}$$

Because $f + g \in L_p$, then $|f + g|^p \in L_1$. Also, since $p = (p - 1)q$, it follows that $|f + g|^{p-1} \in L_q$. Therefore, we can apply Hölder's inequality to see that

$$\begin{aligned} \int |f||f + g|^{p-1} \, d\mu &\leq \|f\|_p \Big\{ \int |f + h|^{(p-1)q} \, d\mu \Big\} \\ &= \|f\| |f + g|. \end{aligned}$$

□

With the proof of Minkowski's inequality, we now can say that L_p spaces are closed under addition; in fact, the triangle inequality holds for L_p spaces and thus they are normed vector spaces.

7. Completeness of L_p

We've already shown that L_p is a normed vector space, but as we've also pointed out, normed vector spaces can be metrically incomplete; this is potentially problematic, as we prefer our normed vectors spaces to be complete so that we can apply to them more interesting and generalized results.

We've discussed pointwise and uniform convergence, but to show the completeness of L_p spaces, we will need to introduce a new type of convergence. We begin by tailoring the definition of a Cauchy sequence fit our needs.

Definition 7.1. A sequence (f_n) in L_p is a *Cauchy sequence in* L_p if for every $\epsilon > 0$ there exists an $M(\epsilon)$ such that if $m, n \geq M(\epsilon)$, then $\|f_m - f_n||_p < \epsilon$.

We now rigorously explore what it means for a sequence of functions, each of which has a measurable value, to converge to a given size.

Definition 7.2. A sequence (f_n) in L_p is *convergent to* f *in* L_p if for every $\epsilon > 0$ there exists an $N(\epsilon)$ such that if $n \geq N(\epsilon)$, then $\|f - f_n||_p < \epsilon$.

Lemma 7.3. *If the sequence (f_n) converges to f in L_p, then it is a Cauchy sequence.*

Proof. If $m, n \geq N(\frac{\epsilon}{2})$, then

$$\|f - f_m||_p < \frac{\epsilon}{2}, \text{ and } \|f - f_n||_p < \frac{\epsilon}{2}.$$

Therefore, we have that

$$\|f_m - f_n||_p \leq \|f_m - f||_p + \|f - f_n||_p < \epsilon.$$

□

Now recall that, as we introduced earlier, a normed linear space such that every Cauchy sequence in the space converges to some element in that space is called a Banach space. We conclude by giving a proof that a particular set is a Banach space.

Theorem 7.4. *(Riesz-Fischer Theorem). The set $(L_p, \|\cdot||_p)$ with $1 \leq p < \infty$ is a Banach space.*

Proof. From [Bartle1966]. We know that L_p is a normed linear space. Then to establish the completeness of L_p, let (f_n) be a Cauchy sequence relative to the norm $\|\cdot||_p$. Then, if $\epsilon > 0$, there exists an $M(\epsilon)$ such that if $m, n \geq M(\epsilon)$, then

$$\int |f_m - f_n|^p \, d\mu = \|f_m - f_n||_p^p < \epsilon^p. \tag{7.5}$$

There exists a subsequence (g_k) of (f_n) such that $\|g_{k+1} - g_k||_p < 2^{-k}$ for $k \in \mathbb{N}$. Define g by

$$g(x) = |g_1(x)| + \sum_{k=1}^{\infty} |g_{k+1}(x) - g_k(x)|, \tag{7.6}$$

so that g is in $(X, \mathcal{S}, \mu)^+$. By Fatou's Lemma, we have that

$$\int |g|^p \; d\mu \leq \liminf_{n\to\infty} \int \Big\{ |g_1| = \sum_{k=1}^{n} |g_{k+1} - g_k| \Big\}^p \; d\mu.$$

We then take the pth root of both sides before applying Minkowski's inequality to derive that

$$\Big\{ \int |g|^p \; d\mu \Big\}^{\frac{1}{p}} \leq \liminf_{n\to\infty} \Big\{ \|g_1\|_p + \sum_{k=1}^{n} \|g_{k+1} - g_k\|_p \Big\}.$$

$$\leq \|g_1\|_p + 1.$$

Thus, if $E = x \in X : g(x) < \infty$, then $E \in \mathcal{S}$ and $\mu(X \backslash E) = 0$. As a result, we see that (7.6) converges almost everywhere and $g\chi_E$ belongs to L_p. Now, we define f on X by

$$f(x) = g_1(x) + \sum_{k=1}^{\infty} g_{k+1}(x) - g_k(x), \text{ if } x \in E,$$

$$= 0, \text{ if } x \notin E.$$

Because $|g_k| \leq |g_1| + \sum_{j=1}^{k-1} |g_{j+1} - g_j| \leq g$ and since (g_k) converges almost everywhere to f, we can deduce from the Lebesgue Dominated Convergence Theorem that $f \in L_p$. Since $|f - g_k|^p \leq 2^p g^p$, we see again from the Dominated Convergence Theorem that $0 = \lim \|f - g_k\|_p$, so that (g_k) converges in L_p to f. Going back to (7.5), if $m \geq M(\epsilon)$ and k is sufficiently large, then

$$\int |f_m - g_k|^p \; d\mu < \epsilon^p.$$

Then we apply Fatou's Lemma to find that

$$\int |f_m - f|^p \; d\mu \leq \liminf_{k\to\infty} \int |f_m - g_k|^p \; d\mu \leq \epsilon^p,$$

whenever $m \geq M(\epsilon)$. Thus, we've proven that the sequence (f_n) converges to f in the norm of L_p, so we conclude that L_p is a Banach space. $\square$

Acknowledgments

I would like to offer a special thanks my mentor, Chris Wilson, for serving as an invaluable source of guidance, knowledge, and support over this past summer. Thank you as well to the other members of the analysis research group, Danny Espejo and Robert Hale, for their encouragement and for sharing their research with me. I would also like to express my gratitude to Professor Justin Lanier for introducing me to Lebesgue spaces and furthering my interest in functional analysis. Lastly, I would like to thank Peter May for organizing the REU program and all of the lecturers who presented their work. I am grateful to have had the opportunity to develop as a mathematician alongside so many other talented individuals.

References

[Bartle1966] Robert Bartle. The Elements of Integration and Lebesgue Measure. John Wiley & Sons, Inc, 1966.

[Pugh2015] Charles Pugh. Real Mathematical Analysis 2e. Springer Undergraduate Mathematics Series, 2015.

[Dahya2019] Raj Dahya. A Globabl Analysis of Spaces in Operator Theory, 2019.

[Magnus2020] Robert Magnus. Function Sequences and Function Series. In: Fundamental Mathematical Analysis. Springer Undergraduate Mathematics Series. Springer, Cham, 2020.

[Rudin1976] Walter Rudin. Principles of Mathematical Analysis 3e. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc, 1976.

[Royden2018] Halsey Royden and Patrick Fitzpatrick. Real Analysis 4e. Pearson, 2018.

[NLBE 2011] Narici, Lawrence; Beckenstein, Edward (2011). Topological Vector Spaces. Pure and applied mathematics (Second ed.). Boca Raton, FL: CRC Press. ISBN 978-1584888666. OCLC 144216834

[Axler2020] Sheldon Axler. Measure, Integration, & Real Analysis. Springer, 2020.

[Tao2011] Terence Tao. An Introduction to Measure Theory. American Mathematical Society, 2011.

[Solovay1970] Solovay, Robert M. (1970), "A model of set-theory in which every set of reals is Lebesgue measurable", Annals of Mathematics, Second Series, 92 (1): 1–56, doi:10.2307/1970696, ISSN 0003-486X, JSTOR 1970696, MR 0265151