# Measure Theory, Integration Theory, and $L^p$ Spaces: Reading Group Notes

K. R. Chickering
*krchicke@ucdavis.edu*

April 29, 2019

## Contents

# 1 Introduction

In the late 19th and early 20th century, analysts began to realize that the Riemann integral was not adequate for many applications. Of main concern was the fact that the space $\mathcal{R}$ of Riemann integrable functions is not closed under taking point-wise limits of sequences in $\mathcal{R}$. That is, it is not true in general that:

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

Consider the following, defined for all $n \geqslant 2$ on $[0,1]$ (Bruckner, 1997):

$$f_n(0) = f_n(2/n) = 0 \qquad f_n(1/n) = n$$

This function defines a triangle of height $n$ and base $1/n$. We take the function $f_n$ to be zero on $[2/n, 1]$. Each of these triangles is Riemann integrable (with integral 1 on $[0,1]$) and we find that:

$$\lim_{n\to\infty}\int_0^1 f_n(x)\ dx = 1 > 0 = \int_0^1 \lim_{n\to\infty} f_n(x)\ dx$$

The inability to take limits of functions in $\mathcal{R}$ will be the main focus here, but there are several other problems with the Riemann integral. For example, the fundamental theorem of calculus for the Riemann integral requires that the the function $f$ has an integrable derivative on the interval $(a,b)$ so that we have:

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

This is, in practical cases, irrelevant, but continuous functions do exist which poses non-integrable derivatives. The existence of such functions motivates a better theory, even if they are extreme constructions. The construction and investigation of the properties for such functions proves to be quite subtle, and for further information on this point the reader is encouraged to consult the references below.

Another problem is that there are functions which our intuition says should be integrable, but which are not Riemann integrable. For example, the Dirchlet function:

$$D(x) = \begin{cases} 0\ ,\ x \in \boldsymbol{Q} \\ 1\ ,\ x \notin \boldsymbol{Q} \end{cases}$$

is not Riemann integrable (if the reader has not seen this fact before they should prove it by showing that the upper and lower sums cannot be made to be arbitrarily close). If we consider that the irrational numbers vastly outnumber the rational numbers, we may expect that we could simply ignore the contributions to the integral given by the $D(x)$ for $x \in \boldsymbol{Q}$. This informal argument leads us to suspect that this integral *should* have a definite value, namely:

$$\int_a^b D(x)\ dx = b - a \tag{1.1}$$

Going back to the previous point, we can construct functions for which the integral of the limit is not even Riemann integrable. Consider an enumeration $\{q_n\}_{n=1}^{\infty}$, $q_n \in \boldsymbol{Q}$, such that $0 < q_n < 1$ for all $n$. Then define the following function:

$$D_n(x) = \begin{cases} 0 \ , \ x = q_m \ , \ m \leqslant n \\ 1 \ , \ \ \text{otherwise} \end{cases}$$

Clearly $D_n(x) \to D(x)$ point-wise, and each $D_n(x)$ is Riemann integrable on $[0,1]$ since there are only finitely many discontinuities. However the Dirchlet function is not integrable, so we have explicitly constructed a sequence of Riemann integrable functions whose limit isn't even Riemann integrable.

We will see that with the notions of measurable functions and Lebesgue integration we will be able to construct complete spaces of integrable functions and integrate strange functions like the Dirchlet function. In Section 3 we outline the conditions under which the limit of a sequence of integrable functions is integrable, and in Section 4 we look at spaces of $p$ power summable functions and show that these spaces are complete.

We will proceed in two stages. First, we introduce the minimum amount of measure theory to understand measurable functions and give a foundation for integration theory. Next, we develop the Lebesgue integral with the goal of understanding the dominated convergence theorem, which can be thought of as **the** result of introductory Lebesgue integration. Finally, we will explore one of the most important applications of Lebesgue integration theory, which is the construction of complete functions spaces of integrable functions.

Unfortunately, we must omit many important and instructive results in our current discussion, and we will necessarily only cover the bare minimum to glean the essence of these theories. The reader is encouraged to look through the references and study further the notions of measure theory, integration, and $L^p$ spaces.

## 1.1 Notation and Definitions

Before we begin, we consider some notation and definitions which will be used throughout the remainder of our discussion. We denote the reals, complex numbers, integers, natural numbers, and rationals by $\boldsymbol{R}, \boldsymbol{C}, \boldsymbol{Z}, \boldsymbol{N}$, and $\boldsymbol{Q}$ respectively. The natural numbers start at 1, and we denote by $\boldsymbol{R}^+$ the set of non-negative real numbers, that is $\boldsymbol{R}^+ := \{ x \in \boldsymbol{R} \,|\, 0 \leqslant x \}$. We shall denote by $\Omega$ an open subset of $\boldsymbol{R}^d$.

**Definition 1.1:** *We say that a sequence $f_n : \Omega \to \boldsymbol{R}$ converges point-wise (or converges) to a function $f : \Omega \to \boldsymbol{R}$, denoted $f_n \to f$, if for every $x \in \Omega$:*

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

**Definition 1.2:** *We say that a sequence $f_n : \Omega \to \boldsymbol{R}$ converges uniformly to a function $f : \Omega \to \boldsymbol{R}$, denoted $f_n \rightrightarrows f$, if given any $\varepsilon > 0$, there exists an $N \in \boldsymbol{N}$ such that whenever $n \geqslant N$ we have:*

$$|f_n(x) - f(x)| < \varepsilon$$

*for all $x \in \Omega$.*

Recall that uniform convergence is stronger than point-wise convergence in two senses. First, any sequence which converges uniformly necessarily also converges point-wise. Second, if the sequence of functions $f_n$ are each continuous on $\Omega$, then $f$ will be continuous on $\Omega$. If each $f_n$ is differentiable on $\Omega$, then $f$ will be differentiable on $\Omega$.

## 2 Measure Theory

Measure theory is central to our new formulation of the integral which will be presented in the next section. The central idea in measure theory is to generalize our intuitive notions of volume to include a much broader class of sets. In one dimension, we think of volume as length, in two dimensions area, and so on and so forth.

Fundamental to the notion of measure is specifying a class of sets which will be 'measurable'. A measure space will then be the measurable sets along with a function $\mu : \Omega \to \mathbb{R}^+$ which assigns to each measurable set in $\Omega$ a non-negative 'measure'.

Clearly we would like the class of sets which are measurable to behave 'well', and the following discussion will seem, at first, somewhat unmotivated. This is an unfortunate quirk of the present exposition, but when we define measure, the following will make more sense.

We mention at the start that there are sets, most notably subsets of $\boldsymbol{R}^d$, which are not measurable. These sets are 'pathological' by any stretch of the imagination, and we defer the construction of such sets to the references. It is, of course, instructive to realize such sets do exist.

**Definition 2.1:** *A collection $\Sigma$ of subsets $\sigma \subset \Omega$ is called a sigma-algebra if it satisfies:*

1. *If $\sigma \in \Sigma$, then $\Omega \setminus \sigma \in \Sigma$.*
2. *If $\sigma_1, \sigma_2, \cdots$ is a countable family of sets in $\Sigma$, then $\bigcup_{j=1}^{\infty} \sigma_j \in \Sigma$.*
3. *$\Omega \in \Sigma$.*

In addition to being closed under taking unions, sigma-algebras are also closed under countable intersections and pairwise complements. We denote $\Omega \setminus \sigma$ by $\sigma^C$, where it is understood that the complement is taken in $\Omega$.

**Theorem 2.2:** *Let $\Sigma$ be a sigma-algebra of $\Omega$, then:*

1. *$\varnothing$ in $\Sigma$.*
2. *If $\sigma_1, \sigma_2, \cdots$ is a countable family of sets in $\Sigma$, then $\bigcap_{i=1}^{\infty} \sigma_i \in \Sigma$.*
3. *If $\sigma_1, \sigma_2 \in \Sigma$, then $\sigma_1 \setminus \sigma_2 \in \Sigma$.*

**Proof:** The first statement is obvious since $\Omega \setminus \Omega = \varnothing \in \Sigma$ by property (1) in Definition 2.

The second statement is a consequence of Demorgan's law:

$$\bigcap_{i=1}^{\infty} \sigma_i = \bigcup_{i=1}^{\infty} {\sigma_i}^C \in \Sigma$$

since the countable union of elements in $\Sigma$ is in $\Sigma$ and ${\sigma_i}^C \in \Sigma$.

The third property follows from:

$$\sigma_1 \setminus \sigma_2 = (\Omega \setminus \sigma_2) \cap \sigma_1$$

where $\Omega \setminus \sigma_2 \in \Sigma$ by property (1) in Definition 2 and the intersection of two elements of $\Sigma$ is in $\Sigma$ by property (2) which we just proved.

■

This sigma-algebras prove to be the proper setting to discuss measure because they will satisfy the countable additivity (defined below) of the measure in a very nice way.

## 2.1 Measure Spaces

We start our discussion of measure spaces by formally defining a measure:

**Definition 2.3:** *Given a sigma algebra $\Sigma$, a function $\mu : \Sigma \to \boldsymbol{R}^+$ is called a **measure** if:*

1. *$\mu(\varnothing) = 0$ and*
2. *(Countable Additivity) given a sequence of disjoint sets $A_1, A_2, \cdots$ with $A_j \in \Sigma$ we have:*

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j)$$

Students familiar with probability will immediately see that probability spaces are measure spaces, where $\Sigma$ is the state space and $\mu(\sigma)$, $\sigma \in \Sigma$ is the probability of the event $\sigma$ occurring.

**Definition 2.4:** *A **measure space** is a three-tuple consisting of a measure $\mu$, set $\Omega$, and sigma algebra $\Sigma$ of $\Omega$, denoted $(\Omega, \Sigma, \mu)$.*

A measure space is useful because it specifies the space, the sigma-algebra on that space, and how we are measuring the sets in the sigma-algebra. Note that there can be many different sigma-algebras for a space $\Omega$, much like there can be many different topologies for a space of points.

The notion of measure allows us to make a convenient statement about functions. If two functions $f$ and $g$ are such that $f(x) = g(x)$ for all $x \in \Omega$ except for on some subset $E \subset \Omega$ such that $\mu(E) = 0$, we say that $f$ and $g$ agree **$\mu$-almost everywhere**, or $\mu$-a.e. Our intuition about the integral leads us to suspect the following. Let $f, g : \Omega \to \boldsymbol{R}$ be two functions which agree $\mu$-a.e., then a reasonable definition of integration should have:

$$\int_{\Omega} (f - g)\ dx = 0$$

and indeed, as we shall see below, this is the case.

## 2.2 The Lebesgue Measure on $\boldsymbol{R}^d$

We leave construction of the Lebesgue measure to the references, however we would like to point out that the Lebesgue measure is indeed our natural idea of a measure on $\boldsymbol{R}^d$; that is, it assigns to $d$-dimensional volumes our intuitive notion of volume. For example, the familiar equation of the volume of the $d$-dimensional sphere is given by:

$$\mathcal{L}^d(B_r(x)) = \frac{2\pi^{d/2} r^d}{d\Gamma(d/2)}$$

With regards to measures on $\boldsymbol{R}^d$, the Borel sigma-algebra $\mathcal{B}$ is the natural sigma-algebra to use in the Lebesgue measure space of $\boldsymbol{R}^d$. $\mathcal{B}$ is the sigma-algebra generated by the open balls of $\boldsymbol{R}^d$, and contains all open and closed sets of $\boldsymbol{R}^d$. The Borel sigma-algebra will come up again and again in further study of measure theory.

There are some things to keep in mind about the Lebesgue measure on $\boldsymbol{R}^d$. For example, the Lebesgue measure is translationally invariant, and moreover it is the *only* measure on $\boldsymbol{R}^d$ which is translationally invariant.

**Proposition 2.5:** *Let $(\boldsymbol{R}, \mathcal{B}, \mu)$ be the Lebesgue measure space of $\boldsymbol{R}$. Then $\mu(\boldsymbol{Q} \cap [a,b]) = 0$.*

**Proof:** The Lebesgue measure of the set containing a single point is zero. That is, given $q \in \boldsymbol{R}$, $\mu(\{q\}) = 0$. Note that $\{q\} \in \mathcal{B}$ and that we can construct $\boldsymbol{Q} \cap [a,b]$ as a countable union of disjoint points. Let $\{q_j\}_{j=1}^{\infty}$ be an enumeration of the rational numbers, then:

$$\boldsymbol{Q} \cap [a,b] = \bigcup_j \{q_j\} \quad \Rightarrow \quad \mu(\boldsymbol{Q} \cap [a,b]) = \sum_j \mu(q_j) = 0$$

by countable additivity as desired.

$\blacksquare$

Recall that in Section 1 we discussed the Dirchlet function and our intuition told us that we could ignore the contributions to the integral from $x \in \boldsymbol{Q}$. The Proposition 2.2 is the first step towards making this intuition precise.

## 2.3 Measurable Functions

Consider some function $f : \Omega \to \boldsymbol{R}$, the level sets of this function are given by:

$$S_f(t) = \left\{ x \in \Omega \,\middle|\, f(x) < t \right\} \tag{2.1}$$

We say that the function $f$ is measurable if for every $t \in \mathbb{R}$ the level set $S_f(t)$ is measurable, that is $S_f(t) \in \Sigma$. Note that we do not need to specify a particular measure, only that the level sets themselves are measurable. That is, given a measure space, the measurability of a function depends on the sigma-algebra we have chosen *not* the measure!

**Definition 2.6:** *Let $\Sigma$ be a sigma-algebra of $\Omega$. A non-negative function $f : \Omega \to \boldsymbol{R}^+$ is said to be measurable if for all $t \in \boldsymbol{R}^+$, the level set $S_f(t) \in \Sigma$.*

We address now an important result which is used (not by us here) extensively in the further study of integration.

**Proposition 2.7:** *Let $\{f_j\}_{j=1}^{\infty}$ be a sequence of measurable functions $f_j : \Omega \to \boldsymbol{R}$. Then:*

$$\inf_j f_j(x) \quad \textit{and} \quad \sup_j f_j(x) \quad \textit{and} \quad \liminf_{j\to\infty} f_j(x) \quad \textit{and} \quad \limsup_{j\to\infty} f_j(x)$$

*exist and are measurable (where existence is into the extended real numbers).*

**Proof:** Note that:

$$\left\{ x \in \Omega \,\middle|\, \sup_j f_j(x) > t \right\} = \bigcup_j \{ x \in \Omega \mid f_j(x) > t \}$$

and by property (2) in Definition 2, $\sup_j f_j$ is measurable. $\inf_j f_j$ is also measurable since:

$$\inf_j f_j(x) = -\sup_j(-f_j(x))$$

Finally, $\liminf_{j\to\infty} f_j$ and $\limsup_{j\to\infty} f_j$ are measurable since:

$$\limsup_{j\to\infty} f_j(x) := \inf_k \{\sup_{j\geqslant k} f_j\} \quad \text{and} \quad \liminf_{j\to\infty} f_j(x) := \sup_k \{\inf_{j\geqslant k} f_j\}$$

■

And finally, we show that the limit of a sequence of measurable functions is itself measurable.

**Corollary 2.8:** *Let $\{f_j\}_{j=1}^{\infty}$ be a sequence of measurable functions $f_j : \Omega \to \boldsymbol{R}$. Then if we have:*

$$\lim_{j\to\infty} f_j(x) = f(x)$$

*then $f$ is measurable.*

**Proof:** Since $f(x) = \limsup_{j\to\infty} f_j(x) = \liminf_{j\to\infty} f_j(x)$, the result follows from Proposition 2.3 above.

■

This makes handling functions 'nice', since we can pass the limits, and even limit superiors and inferiors around without worrying about the measurability of the resulting function.

# 3 Integration

We are now equipped to define an integral based on measurable functions. Let $f : \Omega \to \boldsymbol{R}$ be a function and $\mu : \Omega \to \boldsymbol{R}^+$, that is $f$ takes on only positive values and zero. Let

$$F_f(t) := \mu(S_f(t))$$

We now define the Lebesgue integral of the function $f$ on $\Omega$ as:

$$\int_\Omega f \ \mu(dx) := \int_0^\infty F_f(t) \ dt \tag{3.1}$$

where the integral on the RHS is a Riemann integral and we interpret divergence of the Riemann integral on the RHS as the definition of divergence of the integral on the LHS.

It is important to point out *why* this definition works, that is the Riemann integral on the RHS converges in an appropriate manner. The function $F_f(t)$ is a monotonically decreasing function since for all $t_1 < t_2$, $S_f(t_2) \subset S_f(t_1)$ and countable additivity then implies that $F_f(t_2) \leqslant F_f(t_1)$. We know from undergraduate analysis that such a function will be Riemann integrable.

We will often write $\int_\Omega f$, $\int_\Omega f \ d\mu$ or $\int_\Omega f \ dx$ depending on the situation and which is most convenient. Many authors use differing conventions. Unless otherwise stated, all integrals are to be taken as Lebesgue integrals.

We can easily extend this definition to all real-valued functions in the following way. Let $f : \Omega \to \boldsymbol{R}$ be given and define:

$$f_+(x) := \begin{cases} f(x) & , \ f(x) \geqslant 0 \\ 0 & , \ \text{otherwise} \end{cases} \qquad f_-(x) := \begin{cases} -f(x) & , \ f(x) \leqslant 0 \\ 0 & , \ \text{otherwise} \end{cases}$$

We now have that $f = f_+ - f_-$. Since both $f_+, f_- : \Omega \to \boldsymbol{R}^+$, we can define a Lebesgue integral for each and we get the more general definition:

$$\int_\Omega f = \int_\Omega f_+ - \int_\Omega f_- \tag{3.2}$$

where the integral on the left is the Lebesgue integral and the integrals on the right are as defined as in (3.1). In a similar way the definition can be extended to any complex-valued $f = u + iv$, $u, v : \Omega \to \boldsymbol{R}$ by:

$$\int_\Omega f = \int_\Omega u_+ - \int_\Omega u_- + i \left( \int_\Omega v_+ - \int_\Omega v_- \right)$$

For the rest of our discussion we will use the integral defined in (3.2).

It is pedagogical to think geometrically about the integral we have just defined and to compare it to the Riemann integral. In addition to gaining deeper understanding, the reason for agreement between the Lebesgue and Riemann formulations will become more clear. In (3.2), we are effectively integrating 'bottom up', rather than 'left to right'.

Finally, we leave off with a theorem linking the Riemann integral to the Lebesgue integral.

**Theorem 3.1:** *Let $f : \boldsymbol{R}^d \to \boldsymbol{R}$ be a Riemann integrable function on a compact set $\Gamma \subset \boldsymbol{R}^d$, then the Riemann integral agrees with the Lebesgue integral* (3.2).

There is a caveat to Theorem 3, a function $f$ may be indefinitely Riemann integrable on a set but not Lebesgue integrable at all.

Before moving on, we mention an identity which can be useful in general. Let $\Theta(s)$ be the Heaviside step function:

$$\Theta(s) := \begin{cases} 1 & \text{if} \quad s \geqslant 0 \\ 0 & \text{if} \quad s < 0 \end{cases}$$

then we have the following for the Lebesgue integral:

$$\begin{aligned} \int_0^\infty F_f(t)\; dt &= \int_0^\infty \left( \int_\Omega \Theta(F_f(x) - t)\; \mu(dx) \right)\; dt \\ &= \int_\Omega \left( \int_0^{f(x)} dt \right)\; \mu(dx) = \int_\Omega f(x)\; \mu(dx) \end{aligned}$$

which agrees with our definition (3.1).

## 3.1 Integration Theorems

We now move on to theorems regarding our new integral, culminating in the dominated convergence theorem. We begin by proving that our integral is linear, as is to be expected from a reasonable definition of integration. The proof of this is trivial for Riemann integrals, but we find that it is not trivial in the case of the Lebesgue integral.

**Theorem 3.2:** *Let $f, g : \Omega \to \boldsymbol{R}$ be summable functions and let $\lambda, \gamma \in \boldsymbol{R}$, then:*

$$\int_\Omega (\lambda f + \gamma g)\; d\mu = \lambda \int_\Omega f\; d\mu + \gamma \int_\Omega g\; d\mu$$

*In other words, the integral* (3.2) *is linear.*

Proving that the Lebesgue integral is surprisingly tricky, unlike the case for the Riemann integral. For a proof of this result, see (Stein, 2005).

Our next result begins our discussion of convergence.

**Theorem 3.3:** ***(Monotone Convergence Theorem)*** *Let $f_j \to f$ be an increasing sequence of summable functions on $(\Omega, \Sigma, \mu)$. Then $f$ is measurable and:*

$$\lim_{j\to\infty} \int_\Omega f_j\; \mu(dx) = \int_\Omega f\; \mu(dx)$$

*where the RHS is finite and $f$ is summable if and only if the LHS is finite.*

Note that for a sequence to be increasing on $\Omega$ we mean that for all $j$, $f_{j+1}(x) \geqslant f_j(x)$ for a.e. $x \in \Omega$.

**Theorem 3.4:** ***(Dominated Convergence Theorem)*** *Let $\{f_n\}_{n=1}^\infty$ be a sequence of functions $f_n : \Omega \to \mathbb{R}$ converging point-wise to $f$. Suppose that there is a function $G$ such that $|f_n(x)| \leqslant G(x)$ for all $x \in \Omega$ and $n \in \mathbb{N}$. Then:*

$$\lim_{n\to\infty} \int_\Omega f_n\; \mu(dx) = \int_\Omega f\; \mu(dx) \tag{3.3}$$

The importance of this theorem cannot be overstated. In particular, this gives us the ability to take the point-wise limits of functions and exchange taking limits of functions with integration. The same does not hold for Riemannian integration, and in particular, bounded functions and most 'nice' functions accept this interchange of limit and integral.

The next result is a strengthening of the dominated convergence theorem which can be useful in practice.

**Theorem 3.5:** *Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of functions $f_n : \Omega \to \mathbb{R}$ converging pointwise to $f$. Let $\{G_n\}_{n=1}^{\infty}$ be another sequence of functions $G_n : \Omega \to \mathbb{R}^+$ converging to $G$. If $|f_n(x)| \leqslant G_n(x)$ for all $x \in \Omega$, $n \in \mathbb{N}$ and:*

$$\lim_{n\to\infty} \int_{\Omega} |G(x) - G_n(x)| \; \mu(dx) = 0$$

*Then:*

$$\lim_{n\to\infty} \int_{\Omega} f_n \; d\mu = \int_{\Omega} f \; d\mu$$

## 3.2 The Dirchlet Function

We return to a discussion of the Dirchlet function:

$$D(x) = \begin{cases} 0 \;, \; x \in \boldsymbol{Q} \\ 1 \;, \; x \notin \boldsymbol{Q} \end{cases} \tag{3.4}$$

which we discussed in Section 1. We argued informally that $\mu(\boldsymbol{Q}) = 0$, and we saw in Proposition (**??**) that this is indeed the case. It should be immediately obvious, given our definition of the integral, that the Dirchlet function (3.4) is integrable on any bounded domain.

**Corollary 3.6:** *Let $[a, b]$ be an interval in $\boldsymbol{R}$. Then the integral of the Dirchlet function* (3.4) *is:*

$$\int_a^b D(x) \; dx = b - a$$

**Proof:** This follows from Corollary **??** and the definition (3.1) of the Lebesgue integral.

∎

This result in effect tells us that the 'shadow' of the Dirchlet function onto the real line is measurable, or has 'volume'. This agrees with our intuition.

# 4 Introductory $L^p$ Space Theory

The completion of spaces of integrable functions under point-wise limits give us the opportunity to work with several 'nice' function spaces. Nice in the sense that they are vector spaces with some notion of distance, and are complete, that is any Cauchy sequence converges.

Many of the proofs will be omitted in this section, but can be found in the references below.

**Definition 4.1:** *The space $L^p(\Omega, \Sigma, \mu)$, often abbreviated to $L^p(\Omega)$, is the space of p-power summable functions $f$ on $\Omega$:*

$$\int_\Omega |f|^p \, d\mu < \infty \tag{4.1}$$

*along with the $L^p$ norm:*

$$||f||_p = ||f||_{L^p(\Omega)} := \left( \int_\Omega |f|^p d\mu \right)^{1/p} \tag{4.2}$$

*defined whenever $1 \leqslant p < \infty$.*

We make two important notes, the first is that the reason for excluding $p < 1$ is that in this case, the $L^p$ norm fails to be a norm as can be easily checked (specifically it does not satisfy the triangle inequality). Second, we will extend the definition of $L^p$-spaces to include $L^\infty(\Omega)$, but we defer the discussion to Section 4.3 below.

## 4.1 Hilbert and Banach Spaces

The relegation of Hilbert and Banach spaces to a single subsection should not fool the reader into thinking that these notions are in any way unimportant. Quite to the contrary, these two classes of spaces need to be studied extensively in analysis. We will only touch the surface of these spaces, and learn enough to apply our knowledge to the topic of $L^p$ spaces.

**Definition 4.2:** *A* ***Banach space*** *is a complete, normed, vector space.*

It should be clear why a Banach space is nice to work with. We can freely take (convergent) limits, we have a geometric notion of distance provided by the norm, and we inherit all of the nice properties of a vector space (we normally take the vector space over $\boldsymbol{C}$).

The importance of this notion is that $L^p$-spaces are conveniently Banach spaces.

**Theorem 4.3:** *The space $L^p(\Omega, \Sigma, \mu)$, for some $1 \leqslant p < \infty$ as defined in (4) is Banach space in the norm defined by (4.2).*

This theorem establishes that the $L^p$ spaces are nice. We can immediately see as a consequence of $L^p(\Omega)$ being a Banach space, any sequence $f_k \to f$ with $f_k \in L^p(\Omega)$ necessarily implies that $f$ is also in $L^p(\Omega)$.

Recall the class of functions $\mathcal{R}$ of Riemann integrable functions discussed in the introduction. It is not entirely obvious that we could not formulate $L^p$ spaces using Riemann integrable functions, rather than Lebesgue integrable ones. However, recall that $\mathcal{R}$ is not closed under taking point-wise limits under integrals. That is, $\mathcal{R}$ is not complete in the $L^p$ norm! This lack of completeness is what makes $L^p$ a much better space to work with then $\mathcal{R}$.

**Definition 4.4:** *A* ***Hilbert space*** $\mathscr{H}$ *is a Banach space equipped with an inner product* $(\cdot,\cdot) : \mathscr{H} \times \mathscr{H} \to \boldsymbol{R}^+$ *such that* $||\cdot|| = (\cdot,\cdot)$*. That is, a Hilbert space is a Banach space equipped with an inner product generated by its norm.*

Hilbert spaces arise as infinite-dimensional generalizations of Euclidean spaces, and because they are equipped with an inner product, carry a notion of orthogonality. We are concerned with the space $L^2$, which shows up in many branches of analysis, most notably the study of the Fourier transform.

**Theorem 4.5:** *The space* $L^2(\Omega, \Sigma, \mu)$ *is a Hilbert space with the inner product:*

$$||f||_{L^2(\Omega)} = (f,f) = \left(\int_\Omega f^2 \ \mu(dx)\right)^{1/2} \tag{4.3}$$

*generating the* $L^2$ *norm* $||f||_{L^2(\Omega)}$ (4.2).

One last point to make about $L^p$-spaces is that their elements are not *functions*! To see why this is the case, consider two functions $f$ and $g$ on $\Omega$ which agree $\mu$-a.e.. Then certainly their integrals will agree, and:

$$||f-g||_{L^p} = \left(\int_\Omega |f-g|^2 \ \mu(dx)\right)^{1/p} = 0$$

which, according to the vector space axioms, means that $f = g$. Since the functions *don't* equal each other, the elements of the $L^p$ space must be equivalence classes of functions which agree $\mu$-a.e. and we say that $f \equiv g$. Thinking of the elements of an $L^p$ space is rarely dangerous, but it is an important thing to keep in the back of your head.

## 4.2 $L^p$ Inequalities

$L^p$ theory has several fundamental inequalities which color the study of these spaces.

**Theorem 4.6:** ***(Hölder's inequality)*** *Let* $p,q$ *be dual indices (meaning that* $1/p + 1/q = 1$*) with* $1 \leqslant p \leqslant \infty$*. Let* $f \in L^p$ *and* $g \in L^q$*. Then the product* $f(x)g(x)$ *(taken point-wise) is in* $L^1$ *and:*

$$||fg||_{L^1} \leqslant ||f||_{L^p} \, ||g||_{L^q} \tag{4.4}$$

We will not prove this inequality here, see (Lieb 2001) for a complete proof and discussion. The proof requires another inequality known as Jensen's inequality which we do not state. We remark that the case that $p = q = 2$, Hölder's inequality becomes the Schwartz inequality:

$$\left|\int_\Omega fg\right|^2 \leqslant \int_\Omega |f|^2 \int_\Omega |g|^2$$

The following is another classic inequality in the theory of $L^p$ spaces.

**Theorem 4.7:** ***(Minkowski's inequality)*** *Let $1 \leqslant p < \infty$ and $f, g \in L^p$, then $f + g \in L^p$ and:*

$$||f+g||_{L^p} \leqslant ||f||_{L^p} + ||g||_{L^p} \tag{4.5}$$

This is a generalization of the triangle inequality, which is actually used in proving that $L^p$ is a Banach space, but we have gone slightly out of order here since we are not providing proofs to these theorems.

**Theorem 4.8:** ***(Separability of $L^p$)*** *The space $L^p(\Omega)$ is separable. That is, there exists a countable collection $\{f_k\}$ with $f_k \in L^p(\Omega)$ for each $k$ such that linear combinations of $f_k$s are dense in $L^p(\Omega$.*

Separability is useful in general because linear functionals on $L^p$ spaces are characterized by their behavior on dense subsets of $L^p$. This can be advantageous in some situations.

## 4.3 The Space $L^\infty$

The next logical question is what generalizations can we make to $L^p$ spaces. The immediate generalization is the dual space of $L^1$, which fundamentally means we need to define $L^\infty$ in a way that is reasonable.

**Definition 4.9:** *The $L^\infty$ norm is defined by:*

$$||f||_\infty = ||f||_{L^\infty(\Omega)} := \inf \left\{ C \geqslant 0 \,\middle|\, |f(x)| \leqslant C \ \ a.e. \ \forall \ x \in \Omega \right\} \tag{4.6}$$

*and the space $L^\infty(\Omega)$ is the space of functions:*

$$L^\infty(\Omega) := \left\{ f \,\middle|\, ||f||_\infty < \infty \right\} \tag{4.7}$$

*That is, $L^\infty(\Omega)$ is the space of bounded (more precisely essentially bounded) functions on $\Omega$.*

We want to be sure that this definition agrees with the definition that we had before, in other words, we want to show that:

$$\lim_{p\to\infty} ||f||_p = \lim_{p\to\infty} \left( \int_\Omega |f|^p \ \mu(dx) \right)^{1/p} = ||f||_\infty$$

The following proof is provided by (Stein 2011).

**Theorem 4.10:** *Suppose $f \in L^\infty$ is supported on a set $E \subset \Omega$ of finite measure. Then $f \in L^p$ for all $1 \leqslant p < \infty$ and:*

$$\lim_{p\to\infty} ||f||_p = ||f||_\infty$$

**Proof:** If $\mu(E) = 0$ we are done, since $f \equiv 0$ and $||f||_p = ||f||_\infty = 0$ for all $p$. Therefore, assume $\mu(E) > 0$. Then:

$$||f||_p = \left( \int_\Omega |f|^p d\mu \right)^{1/p} \leqslant \left( \int_E ||f||_\infty^p \, d\mu \right)^{1/p} \leqslant ||f||_\infty \, \mu(E)^{1/p}$$

where $\mu(E)^{1/p} \to 1$ as $p \to \infty$ so $\limsup_{p\to\infty} ||f||_p \leqslant ||f||_\infty$.

Conversely, given any $\varepsilon > 0$, we have:

$$\mu(\{\, x \,|\, |f(x)| \geqslant ||f||_\infty - \varepsilon \,\}) \geqslant \delta$$

for some $\delta > 0$. Then we have:

$$\int_\Omega |f|^p \; d\mu \geqslant \delta(||f||_\infty - \varepsilon)^p$$

Therefore, $\liminf_{p\to\infty} ||f||_p \geqslant ||f||_\infty - \varepsilon$. Since $\varepsilon > 0$ was arbitrary, we have that:

$$\limsup_{p\to\infty} ||f||_p \leqslant ||f||_\infty \leqslant \liminf_{p\to\infty} ||f||_p$$

and the result follows.

$\blacksquare$

# References

[1] A. Bruckner, J. Bruckner, B. Thompson, **Real Analysis**, Prentice Hall, 1997.

[2] E. Lieb, M. Loss, **Analysis**, Graduate Studies in Mathematics, 2001.

[3] E. Stein, R. Shakarchi, **Real Analysis**, Princeton University Press, 2005.

[4] E. Stein, R. Shakarchi, **Functional Analysis**, Princeton Univerisy Press, 2011

[5] B. Thompson, J. Bruckner, A. Bruckner, **Elementary Real Analysis**, Prentice Hall, 2008