UiO : Department of Mathematics
University of Oslo

A measure of Lebesgue measure

MAT3400/MAT4400: Linear analysis with applications

Ole Fredrik Brevig

Version 1.02

Abstract

The main texts of MAT3400/MAT4400 are

[4] Tom L. Lindstrøm, *Spaces—an introduction to real analysis.*

[1] Erik Bédos, *Notes on Elementary Linear Analysis.*

The purpose of the present note is to supplement the main texts, and it is intended that the note be read after [4] and before [1]. Specifically, it is expected that the reader is familiar with Sections 7.1–7.6 and 8.1–8.4 in [4]. We will replace and expand on the material found in the second part of Section 7.5 and in Section 8.5 of [4]. Some material from Section 7.8 is also included. Although the main focus of this note is on the Lebesgue measure, we will at times state and prove results for general measure spaces when no additional work is required[1].

[1]This pertains specifically to Lemma 2.2.2, Lemma 2.2.4 and Theorem 2.3.1.

Contents

CHAPTER 1

Riemann and Lebesgue

> Fredholm and Schmidt had nothing else at their disposal beyond the horrible and useless "Riemann integral", and it is likely that progress in Functional Analysis might have been appreciably slowed down if the invention of the Lebesgue integral had not appeared, by a happy coincidence, exactly at the beginning of Hilbert's work.
>
> Jean Dieudonné [2, pp. 119–120].

The main purpose of this chapter is to compare the Riemann integral and the Lebesgue integral. Recall from calculus that the Riemann integral is defined only for certain bounded functions on a finite interval $I = [a, b]$. The Lebesgue integral was constructed in [4, Section 8.4] for the entire real line, but we can easily restrict it to I as follows.

Let $\mathcal{A}$ denote the σ-algebra of Lebesgue measurable sets on $\mathbb{R}$ and let μ be the Lebesgue measure. The σ-algebra of Lebesgue measurable sets on I is $\mathcal{A}_I = \{X \cap I : X \in \mathcal{A}\}$ and μ_I denotes the Lebesgue measure restricted to sets from $\mathcal{A}_I$. Suppressing the subscript I on $\mathcal{A}$ and μ, we consider the measure space $([a, b], \mathcal{A}, \mu)$. If a $\mathcal{A}$-measurable function $f : [a, b] \to \mathbb{R}$ is μ-integrable, then its Lebesgue integral is denoted

$$\int_{[a,b]} f \, d\mu.$$

In the construction of the Riemann integral we only measure the area of rectangles, so the only thing we need to *know* is what the length of an interval is. The construction of the Lebesgue integral relies on the Lebesgue measure, so in this case we can measure many other sets. Hence we expect that the Lebesgue integral can integrate more functions than the Riemann integral.

1.1 Riemann integrable implies Lebesgue integrable

To compare the Riemann integral and the Lebesgue integral, we require the following definition. In the present section (as in [4, Section 8.4]), we will work with half-open intervals.

Definition. A function $g\colon \mathbb{R} \to \mathbb{R}$ is called a *step function* if

$$g = \sum_{j=1}^{J} c_j \mathbf{1}_{I_j}$$

where $c_j \neq 0$ and $I_j = (a_j, b_j]$ for $j = 1, 2, \ldots, J$.

We will also consider step functions defined on an interval I, in which case we require that $I_j \subseteq I$ for $j = 1, 2, \ldots, J$. In the context of the Lebesgue measure, a step function is just a particular example of a simple function since intervals are Lebesgue measurable.

If $g\colon [a, b] \to \mathbb{R}$ is a step function and $I_j = (a_j, b_j]$ for $j = 1, 2, \ldots, J$, then

$$\int_{[a,b]} g \, d\mu = \sum_{j=1}^{J} c_j \mu(I_j) = \sum_{j=1}^{J} c_j (b_j - a_j) = \int_a^b g(x) \, dx, \tag{1.1.1}$$

where the integral on the left hand side is the Lebesgue integral and the integral on the right hand side is the Riemann integral. Having now established the completely expected fact that the Riemann and Lebesgue integrals coincide for step functions, our next goal is to consider more complicated functions.

We begin by discussing a slightly different (but equivalent[1]) construction of the Riemann integral compared to the standard one typically presented in calculus textbooks using partitions. If $f\colon [a, b] \to \mathbb{R}$ is a bounded function, set

$$\mathcal{L}(f) = \sup \left\{ \int_a^b g(x) \, dx \;:\; g \text{ is a step function and } g \leq f \right\},$$
$$\mathcal{U}(f) = \inf \left\{ \int_a^b g(x) \, dx \;:\; g \text{ is a step function and } g \geq f \right\}.$$

The function f is Riemann integrable if and only if $\mathcal{L}(f) = \mathcal{U}(f)$, in which case

$$\int_a^b f(x) \, dx = \mathcal{L}(f) = \mathcal{U}(f).$$

Note that for a function to be Riemann integrable, we have to be able to approximate it well from both above and below using *step* functions. Recall also that for a function to be Lebesgue integrable, we only need to be able to approximate it well from below using the more general *simple* functions.

[1] We leave it to the interested reader to prove that our construction is equivalent to whatever is presented their favorite calculus textbook.

Theorem 1.1.1. *Suppose that a bounded function $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable. Then f is Lebesgue integrable and*

$$\int_{[a,b]} f\, d\mu = \int_a^b f(x)\, dx.$$

Proof. Let M denote the maximum of $|f|$ on $[a,b]$. Since f is Riemann integrable, we can find a sequence of step functions $\{g_j\}_{j\geq 1}$ and a sequence of step functions $\{h_j\}_{j\geq 1}$ such that

$$-M \leq g_j(x) \leq f(x) \leq h_j(x) \leq M \tag{1.1.2}$$

for every $x \in [a,b]$ and such that

$$\lim_{j\to\infty} \int_a^b g_j(x)\, dx = \lim_{j\to\infty} \int_a^b h_j(x)\, dx = \int_a^b f(x)\, dx.$$

By iteratively replacing g_j by $\max(g_j, g_{j-1})$ for $j \geq 2$ and h_j by $\min(h_j, h_{j-1})$ for $j \geq 2$, we may assume without loss of generality that $\{g_j\}_{j\geq 1}$ is increasing and that $\{h_j\}_{j\geq 1}$ is decreasing. Hence they will converge pointwise to functions g and h, respectively, which in view of (1.1.2) satisfy

$$-M \leq g(x) \leq f(x) \leq h(x) \leq M$$

for every $x \in [a,b]$. Note that g and h are measurable by [4, Proposition 7.3.9], since the step functions $\{g_j\}_{j\geq 1}$ and $\{h_j\}_{j\geq 1}$ are measurable. Since

$$0 \leq h_j(x) - g_j(x) \leq 2M$$

we can use the Dominated Convergence Theorem and (1.1.1) to conclude that

$$0 \leq \int_{[a,b]} (h-g)\, d\mu = \lim_{j\to\infty} \int_{[a,b]} (h_j - g_j)\, d\mu = \lim_{j\to\infty} \int_a^b (h_j(x) - g_j(x))\, dx = 0.$$

Since $g \leq h$, we conclude from this that $g = h$ almost everywhere. Since $g \leq f \leq h$, we find that $f = g = h$ almost everywhere. Since the Lebesgue measure is complete, this shows that f is measurable. Since f is bounded by assumption, it is therefore Lebesgue integrable. Moreover, the Lebesgue integral of f is equal to the Lebesgue integral of g. Using that $|g_j| \leq M$ for every $j \geq 1$ from (1.1.2), we can appeal to the Dominated Convergence Theorem and (1.1.1) to compute

$$\int_{[a,b]} f\, d\mu = \int_{[a,b]} g\, d\mu = \lim_{j\to\infty} \int_{[a,b]} g_j\, d\mu = \lim_{j\to\infty} \int_a^b g_j(x)\, dx = \int_a^b f(x)\, dx. \quad \square$$

Theorem 1.1.1 can be used to compute Lebesgue integrals of continuous functions over finite intervals using Riemann integrals, where we can use the techniques of calculus. We conclude this section with two examples intending to clarify how we can handle Lebesgue integrals over $\mathbb{R}$.

Example 1.1.2. Let $f(x) = (1+x^2)^{-1}$ and suppose that we want to compute the Lebesgue integral $\int_{\mathbb{R}} f\,d\mu$. By the Monotone Convergence Theorem

$$\int_{\mathbb{R}} f\,d\mu = \lim_{n\to\infty} \int_{\mathbb{R}} \mathbf{1}_{[-n,n]} f\,d\mu.$$

The integrals on the right hand side can be computed using Theorem 1.1.1 since f is continuous on $[-n,n]$. We obtain

$$\int_{\mathbb{R}} \mathbf{1}_{[-n,n]} f\,d\mu = \int_{[-n,n]} f\,d\mu = \int_{-n}^{n} \frac{1}{1+x^2}\,dx = 2\arctan(n).$$

Computing the limit, we find that $\int_{\mathbb{R}} f\,d\mu = \pi$. ♣

Example 1.1.3 (The Dirichlet integral)**.** Consider the function

$$f(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x} & x \neq 0, \\ 1 & x = 0. \end{cases}$$

A fun calculus exercise is to compute the *improper* integral

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{n\to\infty} \int_{-n}^{n} f(x)\,dx = 1.$$

However, f is *not* Lebesgue integrable on $\mathbb{R}$. We recall from the definition of integrable functions in [4, Section 7.6] that f is Lebesgue integrable if and only if $|f|$ is Lebesgue integrable. We can show that $\int_{\mathbb{R}} |f|\,d\mu = \infty$ (see Exercise 1.2), so f cannot be Lebesgue integrable. ♣

1.2 Lebesgue's criterion

In this section we shall present a criterion which describes precisely when a bounded function $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable. The criterion is, amusingly, due to Lebesgue. To state it, we require the following definition.

Definition. Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function. The *set of discontinuities* of f is $\mathcal{D}_f = \{x \in [a,b] : f \text{ is discontinuous at } x\}$.

A well-known result from calculus states that if $\mathcal{D}_f$ is *finite*, then f is Riemann integrable. It turns out that the Riemann integral can handle an infinite number of discontinuities, provided their Lebesgue measure is zero.

Theorem 1.2.1 (Lebesgue's criterion)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable if and only if $\mu(\mathcal{D}_f) = 0$.*

We will not present a proof of Theorem 1.2.1, but refer instead the interested reader to [3, Theorem 2.28]. While Theorem 1.2.1 seems to be an impressive result, it is often not very useful for practical computations. Example 1.2.2 illustrates one case where it happens to be useful.

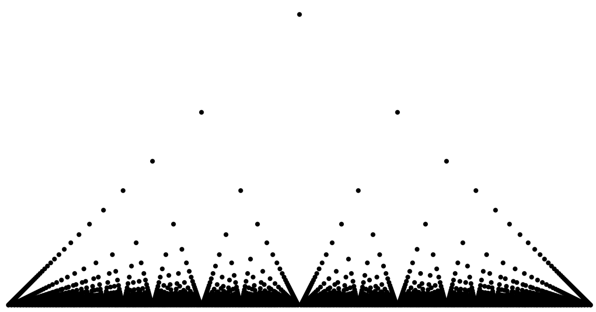

Figure 1.1: Plot of the function from Example 1.2.2 for $q \leq 200$.

Example 1.2.2 (The popcorn function)**.** Consider the function $f\colon [0,1] \to \mathbb{R}$ defined as follows:

- $f(0) = f(1) = 0$,
- if $0 < x < 1$ and x is irrational, then $f(x) = 0$,
- if $0 < x < 1$ and $x = p/q$ for positive integers p and q with no common factor, then $f(x) = 1/q$.

See Figure 1.1 for a plot of an approximation to f. Since $\mu(\mathbb{Q}) = 0$, it follows that f is Lebesgue integrable and that the Lebesgue integral is

$$\int_{[0,1]} f\,d\mu = 0.$$

Is f Riemann integrable? It is not too difficult to check that f is continuous at each irrational x (since any sufficiently close rational number must have a large denominator) and discontinuous at each rational $0 < x < 1$. It therefore follows from Theorem 1.2.1 that f is Riemann integrable and that

$$\int_0^1 f(x)\,dx = 0$$

by Theorem 1.1.1. ♣

1.3 Cantor sets

In this section we shall study some interesting subsets of the unit interval $[0,1]$. We are used to expressing real numbers on this interval in *decimal* expansion

$$x = \sum_{j=1}^{\infty} \frac{a_j}{10^j}, \qquad \text{where} \qquad a_j \in \{0,1,2,3,4,5,6,7,8,9\}.$$

It is also possible to express numbers with respect to other bases, for instance in *binary* expansion

$$x = \sum_{j=1}^{\infty} \frac{b_j}{2^j}, \qquad \text{where} \qquad b_j \in \{0,1\}$$

or in *ternary* expansion

$$x = \sum_{j=1}^{\infty} \frac{c_j}{3^j}, \qquad \text{where} \qquad c_j \in \{0,1,2\}.$$

Just as some rational numbers have two distinct decimal expansions

$$\frac{2}{5} = 0.2000\ldots = 0.1999\ldots,$$

some rational numbers have two distinct binary and ternary expansions. For example, $1/3$ have the ternary expansions

$$\frac{1}{3} = 0.1000\ldots = 0.0222\ldots.$$

Using ternary expansion we can introduce the protagonist of the present section, which is the set of real numbers on $[0,1]$ that have a ternary expansion which does not contain any ones. In particular, $1/3 = 0.0222\ldots$ is in the set.

Definition. The *Cantor set* is

$$\mathcal{C} = \left\{ x = \sum_{j=1}^{\infty} \frac{c_j}{3^j} \ :\ c_j \in \{0,2\} \right\}.$$

We will always choose a ternary expansion for $x \in \mathcal{C}$ without any ones.

The Cantor set is the canonical example of *fractal* subset of the real line. Before discussing this further, let us consider the following question: Is the Cantor set big or small? The following result provides one possible answer.

Theorem 1.3.1. *The Cantor set is uncountable.*

Proof. The function $T\colon \mathcal{C} \to [0,1]$ defined by

$$T(x) = T\left(\sum_{j=1}^{\infty} \frac{c_j}{3^j}\right) = \sum_{j=1}^{\infty} \frac{c_j/2}{2^j}$$

is surjective (see Exercise 1.3). Since $[0,1]$ is uncountable, so is $\mathcal{C}$. □

Since $\mathcal{C}$ is a subset of $[0,1]$, the proof of Theorem 1.3.1 also shows that $\mathcal{C}$ (in some sense) is as big as the unit interval $[0,1]$, so we might answer that the Cantor set is big.

The following result hints at the fractal nature of $\mathcal{C}$, since it states that $\mathcal{C}$ is equal to *two* copies of itself if each copy is scaled down by a *third*.

Figure 1.2: The sets $\mathcal{C}_n$ in the proof of Theorem 1.3.3 for $n = 0, 1, 2, \ldots, 6$.

Theorem 1.3.2. *Let $\mathcal{C}$ denote the Cantor set. Then*

$$\mathcal{C} = \frac{\mathcal{C}}{3} \cup \left(\frac{2}{3} + \frac{\mathcal{C}}{3}\right)$$

and this is a union of two disjoint sets.

Proof. If $x = \sum_{j \geq 1} c_j 3^{-j}$ let us write $[x]_3 = (c_1, c_2, \ldots)$. Then

$$\left[\frac{x}{3}\right]_3 = (0, c_1, c_2, \ldots) \qquad \text{and} \qquad \left[\frac{2}{3} + \frac{x}{3}\right]_3 = (2, c_1, c_2, \ldots).$$

The result now follows after recalling that the Cantor set is defined by the requirement that $c_j \in \{0, 2\}$ for each $j \geq 1$. □

Let us *assume* that $\mathcal{C}$ is Lebesgue measurable. Combining Theorem 1.3.2 with [4, Proposition 8.4.6] and [4, Exercise 8.4.5], we find that

$$\mu(\mathcal{C}) = \frac{1}{3}\mu(\mathcal{C}) + \frac{1}{3}\mu(\mathcal{C}) = \frac{2}{3}\mu(\mathcal{C}),$$

which ensures that $\mu(\mathcal{C}) = 0$ since $\mu(\mathcal{C}) \leq \mu([0,1]) = 1 < \infty$. Hence we might answer that the Cantor set is small. The next result will be proved by establishing that $\mathcal{C}$ is Lebesgue measurable. As a byproduct, we shall obtain a geometric definition of the Cantor set which is probably more useful and more natural than the one we gave in the definition above.

Theorem 1.3.3. *The Lebesgue measure of the Cantor set is* 0.

Proof. We begin with $\mathcal{C}_0 = [0, 1]$. If $1/3 < x < 2/3$, then the first digit in any ternary expansion of x must be a 1. Hence we can safely delete the middle interval $(1/3, 2/3)$ from $\mathcal{C}_0$ without loosing any real number in the Cantor set. We therefore define

$$\mathcal{C}_1 = \mathcal{C}_0 \setminus (1/3, 2/3).$$

A moments thought now reveals to us that if $1/9 < x < 2/9$ or $7/9 < x < 8/9$ then the second digit in any ternary expansion of x must be a 1. Note also that the intervals $(1/9, 2/9)$ and $(7/9, 8/9)$ are the middle third of the intervals $[0, 1/3]$ and $[2/3, 1]$, which together constitute $\mathcal{C}_1$. Since no real number from the former two intervals can be in the Cantor set, we may safely delete them. Hence we set

$$\mathcal{C}_2 = \mathcal{C}_1 \setminus \big((1/9, 2/9) \cup (7/9, 8/9)\big).$$

Note that $\mathcal{C}_2$ consists of four intervals. A similar argument shows that we can safely delete the middle third of any of them. By continuing iteratively in this way, we obtain a sequence of sets $\mathcal{C}_n$. The first seven sets are displayed in Figure 1.2. When going from $\mathcal{C}_{n-1}$ to $\mathcal{C}_n$ we remove precisely the real numbers $0 < x < 1$ whose ternary expansions must have a 1 at the nth digit (and which have not previously been removed). This shows that

$$\mathcal{C} = \bigcap_{n=0}^{\infty} \mathcal{C}_n$$

and hence $\mathcal{C}$ is measurable. By the discussion above, it follows that $\mu(\mathcal{C}) = 0$. It is also possible to give a direct proof of this claim based on the geometric construction as follows: Note that $\mu(\mathcal{C}_n) = \frac{2}{3}\mu(\mathcal{C}_{n-1})$ since we always delete the middle third of each interval. By induction we obtain

$$\mu(\mathcal{C}_n) = \left(\frac{2}{3}\right)^n \mu(\mathcal{C}_0) = \left(\frac{2}{3}\right)^n .$$

Since evidently $\mathcal{C}_n \subseteq \mathcal{C}_{n-1}$, it follows from continuity of measure for decreasing sequences (see [4, Proposition 7.1.5]) that

$$\mu(\mathcal{C}) = \lim_{n\to\infty} \mu(\mathcal{C}_n) = 0. \qquad \square$$

It is possible to exploit the construction of the Cantor set from the proof of Theorem 1.3.3 to construct sets are similar to the Cantor set, but which have positive measure.

Example 1.3.4 (Fat Cantor sets)**.** Suppose that $0 < r < 1/3$. We set $\mathcal{F}_0 = [0,1]$ and iteratively define $\mathcal{F}_n$ by deleting a segment of length r^n from the middle of each interval in $\mathcal{F}_{n-1}$. The next two sets in the sequence are

$$\mathcal{F}_1 = \left[0, \frac{1-r}{2}\right] \cup \left[\frac{1+r}{2}, 1\right]$$

and

$$\begin{aligned}\mathcal{F}_2 = \left[0, \frac{1-r}{4} - \frac{r^2}{2}\right] &\cup \left[\frac{1-r}{4} + \frac{r^2}{2}, \frac{1-r}{2}\right] \\ &\cup \left[\frac{1+r}{2}, \frac{3+r}{4} - \frac{r^2}{2}\right] \cup \left[\frac{3+r}{4} + \frac{r^2}{2}, 1\right].\end{aligned}$$

Since we double the number of intervals in each step, it is clear that $\mathcal{F}_{n-1}$ is comprised of 2^{n-1} intervals. When we proceed to $\mathcal{F}_n$, we delete a segment of length r^n from each of these intervals. By induction, this shows that

$$\mu(\mathcal{F}_n) = \mu(\mathcal{F}_{n-1}) - 2^{n-1}r^n = 1 - \sum_{k=1}^{n} 2^{k-1}r^k.$$

Figure 1.3: The sets $\mathcal{F}_n$ from Example 1.3.4 with $r = 1/4$ for $n = 0, 1, 2, \ldots, 6$.

We now define $\mathcal{F} = \cap_{n \geq 0} \mathcal{F}_n$. As in the proof of Theorem 1.3.3, we find that

$$\mu(\mathcal{F}) = 1 - \sum_{k=1}^{\infty} 2^{k-1} r^k = 1 - \frac{r}{1-2r} = \frac{1-3r}{1-2r}.$$

It is interesting to note that even though $\mathcal{F}$ has positive measure, there is no interval I such that $I \subseteq \mathcal{F}$ (see Exercise 1.6). ♣

Remark. Let us close this section by briefly discussing the fractal nature of $\mathcal{C}$ again, which from our point of view is encoded in Theorem 1.3.2. Every subset $X \subseteq \mathbb{R}$ can be assigned a real number $0 \leq H(X) \leq 1$ which is called its *fractal dimension* (or *Hausdorff dimension*). Every countable set has fractal dimension 0 and every set which contains an interval has fractal dimension 1. The Cantor set has fractal dimension

$$H(\mathcal{C}) = \frac{\log 2}{\log 3} = 0.6309\ldots$$

which gives yet another answer to the question: Is the Cantor set big or small? The interested reader is referred to [3, Chapter 11].

1.4 Exercises

Exercise 1.1. Explain why the Monotone Convergence Theorem applies in Example 1.1.2.

Exercise 1.2. Let f be the function defined in Example 1.1.3.

(a) Prove that

$$\int_{\mathbb{R}} |f| \, d\mu = \infty.$$

(b) Explain why the Monotone Convergence Theorem cannot be used to compute

$$\int_{\mathbb{R}} f \, d\mu$$

in Example 1.1.3. What about the Dominated Convergence Theorem?

Exercise 1.3. Prove that the function T defined in the proof of Theorem 1.3.1 is a surjection.

Exercise 1.4. Let $\mathcal{C}_n$ denote the sets constructed in the proof of Theorem 1.3.3. Prove that

$$\mathcal{C}_n = \frac{\mathcal{C}_{n-1}}{3} \cup \left(\frac{2}{3} + \frac{\mathcal{C}_{n-1}}{3}\right)$$

and that the union is disjoint for $n \geq 1$.

Exercise 1.5. Let $f\colon [0,1] \to \mathbb{R}$ be the function $f = \mathbf{1}_{\mathcal{C}}$ where $\mathcal{C}$ is the Cantor set. Prove that f is Riemann integrable.

Hint. Do *not* use Theorem 1.2.1, but take a look at the proof of Theorem 1.3.3.

Exercise 1.6. Fix $0 < r < 1/3$ and consider the set $\mathcal{F}$ from Example 1.3.4.

(a) Prove that there is no interval I such that $I \subseteq \mathcal{F}$.

(b) Let $f\colon [0,1] \to \mathbb{R}$ be the function $f = \mathbf{1}_{\mathcal{F}}$ where $\mathcal{F}$ a fat Cantor set. Prove that f is not Riemann integrable.

Hint. Use (a) and Theorem 1.2.1.

CHAPTER 2

Littlewood's three principles

> If one of the principles would be the obvious means to settle a problem if it were "quite" true, it is natural to ask if the "nearly" is near enough, and for a problem that is actually soluble it generally is.
>
> J. E. Littlewood [5, p. 27].

Littlewood's three principles are heuristic devices which enable us to better understand Lebesgue measurable sets and functions on the real line, which we have seen in Section 1.3 can be quite erratic. The principles can informally be stated as follows.

1. Every measurable set is nearly a finite union of intervals.

2. Every integrable function is nearly continuous.

3. Every convergent sequence of measurable functions is nearly uniformly convergent.

Note that we have not assigned the word *nearly* a mathematical meaning. In fact, we shall see that this word takes on different meanings in the different principles. The purpose of this chapter is to present three theorems which capture the essence of the three principles.

2.1 The first principle

To state our version of Littlewood's first principle, we shall make use of the following definition.

Definition. Let A and B be subsets of some set X. The *symmetric difference* of A and B is

$$A \triangle B = (A \setminus B) \cup (B \setminus A).$$

The symmetric difference of two sets is the set comprised of the elements which lie in one of the two sets, but not in both. Note that $A = B$ if and only if $A \triangle B = \emptyset$. The symmetric difference is therefore natural to use if we want to measure if two sets are *nearly* equal.

Theorem 2.1.1 (Littlewood's first principle)**.** *Let E be a Lebesgue measurable subset of $\mathbb{R}$ with finite measure. For every $\varepsilon > 0$, there is a finite union of finite intervals F such that*

$$\mu(E \triangle F) \leq \varepsilon.$$

Remark. Theorem 2.1.1 assigns the word *nearly* in Littlewood's first principle the meaning that the measure of the symmetric difference of E and a finite union of intervals can be made as small as we wish.

Note that Theorem 2.1.1 is stated for general finite intervals. Since the Lebesgue measure of a finite union of endpoints is 0, we can assume that the intervals in question are open, closed or half-open without loss of generality.

Proof. Fix $\varepsilon > 0$. By the assumption that $\mu(E) < \infty$, it follows from the outer measure construction (see [4, Chapter 8]) that there is a sequence of finite half-open intervals $\{I_j\}_{j \geq 1}$ such that

$$E \subseteq \bigcup_{j=1}^{\infty} I_j \tag{2.1.1}$$

and such that

$$\sum_{j=1}^{\infty} \mu(I_j) \leq \mu(E) + \frac{\varepsilon}{2}. \tag{2.1.2}$$

Since $\mu(E) < \infty$, the left hand side of (2.1.2) is a convergent sum of nonnegative real numbers. This means that there is an integer J such that $\sum_{j>J} \mu(I_j) \leq \varepsilon/2$. We now choose our finite union of intervals as

$$F = \bigcup_{j=1}^{J} I_j. \tag{2.1.3}$$

To estimate the Lebesgue measure of $E \triangle F$, it is sufficient to estimate the Lebesgue measure of $E \setminus F$ and $F \setminus E$. The latter two sets are disjoint, so

$$\mu(E \triangle F) = \mu(E \setminus F) + \mu(F \setminus E). \tag{2.1.4}$$

To estimate $\mu(E \setminus F)$, we first use (2.1.1) and (2.1.3) to conclude that

$$E \setminus F \subseteq \left(\bigcup_{j=1}^{\infty} I_j \right) \setminus \left(\bigcup_{j=1}^{J} I_j \right) \subseteq \bigcup_{j=J+1}^{\infty} I_j.$$

This implies that $\mu(E \setminus F) \leq \varepsilon/2$ by countable subadditivity and our choice of J. To estimate $\mu(F \setminus E)$, we first use (2.1.3) to see that

$$F \setminus E = \left(\bigcup_{j=1}^{J} I_j \right) \setminus E \subseteq \left(\bigcup_{j=1}^{\infty} I_j \right) \setminus E.$$

Since $\mu(A \setminus B) = \mu(A) - \mu(B)$ for measurable sets $B \subseteq A$ when $\mu(A) < \infty$, we conclude from this, (2.1.1), countable subadditivity and (2.1.2) that

$$\mu(F \setminus E) \leq \mu\left(\bigcup_{j=1}^{\infty} I_j \right) - \mu(E) \leq \left(\sum_{j=1}^{\infty} \mu(I_j) \right) - \mu(E) \leq \frac{\varepsilon}{2}.$$

Inserting our estimates for $\mu(E \setminus F)$ and $\mu(F \setminus E)$ into (2.1.4) yields the desired estimate $\mu(E \triangle F) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$. □

The assumption that E has finite measure in Theorem 2.1.1 cannot be dropped (see Exercise 2.1).

Example 2.1.2. Fix $0 < r < 1/3$ and consider the fat Cantor set $\mathcal{F}$ from Example 1.3.4. Recall that the set $\mathcal{F}_n$ used in that construction is a union of 2^n intervals. Since $\mathcal{F} \subseteq \mathcal{F}_n$ by construction and since $\mu(\mathcal{F}_n) \leq 1$, we find that

$$\mu(\mathcal{F} \triangle \mathcal{F}_n) = \mu(\mathcal{F}_n \setminus \mathcal{F}) = \mu(\mathcal{F}_n) - \mu(\mathcal{F}).$$

Combining this with the computations in Example 1.3.4, we find that

$$\mu(\mathcal{F} \triangle \mathcal{F}_n) = \left(1 - \sum_{k=1}^{n} 2^{k-1} r^k \right) - \left(1 - \sum_{k=1}^{\infty} 2^{k-1} r^k \right) = \sum_{k=n+1}^{\infty} 2^{k-1} r^k = \frac{2^n r^{n+1}}{1 - 2r}.$$

Since $0 < r < 1/3$, we can make this as small as we like by picking n sufficiently large. Of course, we knew this was possible already by Theorem 2.1.1! ♣

We will have use of the following corollary to Theorem 2.1.1 in the next section. Recall from [4, Chapter 7] that a every simple function f can be written in standard form $f = \sum_{j=1}^{J} a_j \mathbf{1}_{A_j}$, where $a_j \neq a_k$ and $A_j \cap A_k = \emptyset$ for $j \neq k$. Recall also that f is integrable if and only if $\mu(A_j) < \infty$ for every j such that $a_j \neq 0$.

From the definition of step functions in Section 1.1, we see that step functions are Lebesgue integrable simple functions.

Corollary 2.1.3. *Suppose that $f\colon \mathbb{R} \to \mathbb{R}$ is a Lebesgue integrable simple function. For every $\varepsilon > 0$, there is a step function $g\colon \mathbb{R} \to \mathbb{R}$ such that*

$$\int_{\mathbb{R}} |f - g| \, d\mu \leq \varepsilon.$$

Proof. Left to the reader (see Exercise 2.2). □

2.2 The second principle

The are (at least) two different theorems which can be chosen to exemplify Littlewood's second principle. It is probably most common to choose Lusin's theorem (see e.g. [3, Exercise 2.4.44]), the proof of which actually relies on a version of Littlewood's third principle! We will instead choose to present a theorem which is easier to prove and more useful in applications. Before stating the result, we recall the following.

Definition. A function $f\colon \mathbb{R} \to \mathbb{R}$ is said to be *compactly supported* if there is a real number $M \geq 0$ such that $f(x) = 0$ whenever $|x| \geq M$.

Theorem 2.2.1 (Littlewood's second principle)**.** *Let $f\colon \mathbb{R} \to \mathbb{R}$ be a Lebesgue integrable function. For every $\varepsilon > 0$ there is a compactly supported continuous function $g\colon \mathbb{R} \to \mathbb{R}$ such that*

$$\int_{\mathbb{R}} |f - g| \, d\mu \leq \varepsilon.$$

Remark. Theorem 2.2.1 assigns the word *nearly* in Littlewood's second principle the meaning that we can choose a continuous function g such that the Lebesgue integral of $|f - g|$ can be made as small as we wish.

Our strategy for the proof of Theorem 2.2.1 is to perform a sequence of approximations. In each step, we will replace our integrable function f by a more manageable function:

$$\text{Integrable} \quad \rightsquigarrow \quad \text{Simple} \quad \rightsquigarrow \quad \text{Step} \quad \rightsquigarrow \quad \text{Continuous}$$

The middle step is a consequence of Littlewood's first principle and can be found in in Corollary 2.1.3 above. Our job is therefore to obtain the first and last approximations. The following result, which is stated and proved for a general measure space, supplies the first.

Lemma 2.2.2. *Suppose that $(X, \mathcal{A}, \mu)$ is a measure space and let $f\colon X \to \mathbb{R}$ be an integrable function. For every $\varepsilon > 0$, there is a simple integrable function g such that*

$$\int_X |f - g| \, d\mu \leq \varepsilon.$$

Proof. Since f is measurable, we can decompose $f = f^+ - f^-$, where f^+ and f^- are nonnegative measurable functions. By [4, Lemma 7.5.3], there are increasing sequences of simple nonnegative functions $\{h_n^+\}_{n\geq 1}$ and $\{h_n^-\}_{n\geq 1}$ such that $\{h_n^\pm\}_{n\geq 1}$ converges pointwise $f^\pm$. Set

$$g_n = h_n^+ - h_n^-.$$

Then $\{g_n\}_{n\geq 1}$ is a sequence of simple functions which converges pointwise to f and which enjoy the estimates

$$|g_n(x)| = |h_n^+(x) - h_n^-(x)| = |h_n^+(x)| + |h_n^-(x)| \leq |f^+(x)| + |f^-(x)| = |f(x)|$$

for every $n \geq 1$ and every $x \in X$. In particular, we conclude from this that

$$|f(x) - g_n(x)| \leq |f(x)| + |g_n(x)| \leq 2|f(x)|$$

for every $n \geq 1$ and every $x \in X$. Since f is integrable, we can therefore appeal to the Dominated Convergence Theorem to establish that

$$\lim_{n\to\infty} \int_X |f - g_n|\, d\mu = \int_X \lim_{n\to\infty} |f - g_n|\, d\mu = 0.$$

In the final equality we used that $\{g_n\}_{n\geq 1}$ converges pointwise to f. Since this limit is equal to 0, there must be some positive integer N such that

$$\int_X |f - g_N|\, d\mu \leq \varepsilon,$$

which is the desired estimate with the simple function $g = g_N$. $\square$

The last approximation in the proof of Theorem 2.2.1 is the easiest one.

Lemma 2.2.3. *Suppose that $f\colon \mathbb{R} \to \mathbb{R}$ is a step function. For every $\varepsilon > 0$, there is a compactly supported continuous function g such that*

$$\int_{\mathbb{R}} |f - g|\, d\mu \leq \varepsilon.$$

Proof. Consider a step function $f = \sum_{j=1}^{J} c_j \mathbf{1}_{I_j}$. Recall from the definition in Section 1.1 that the real numbers c_j are nonzero and that the intervals $I_j = (a_j, b_j]$ are finite. For each $j = 1, 2, \ldots, J$, we let g_j denote the function defined in Figure 2.1 with $\delta_j = \varepsilon/(J|c_j|)$. Since each g_j is continuous and compactly supported, so is the function $g = g_1 + g_2 + \cdots + g_J$. By the triangle inequality, we find that

$$\int_{\mathbb{R}} |f - g|\, d\mu \leq \sum_{j=1}^{J} \int_{\mathbb{R}} |c_j \mathbf{1}_{I_j} - g_j|\, d\mu = \sum_{j=1}^{J} \frac{\varepsilon}{J} = \varepsilon. \qquad \square$$

We are finally ready to wrap up the proof of our version of Littlewood's second principle. Note that the heavy lifting is done by Theorem 2.1.1 through Corollary 2.1.3.

Proof of Theorem 2.2.1. Use Lemma 2.2.2, Corollary 2.1.3 and Lemma 2.2.3. Note that the function produced by Corollary 2.1.3 will be compactly supported. The details are left to the reader (see Exercise 2.3). $\square$

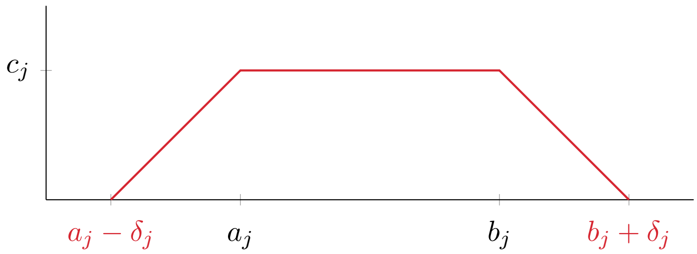

Figure 2.1: The functions g_j from the proof of Lemma 2.2.3.

We will conclude this section by applying the techniques we have developed to Fourier analysis. We shall use the triangle inequality for integrals of complex-valued functions, which has a rather cute proof.

Lemma 2.2.4. *Let $(X, \mathcal{A}, \mu)$ be a measure space and suppose that $f : X \to \mathbb{C}$ is integrable. Then*

$$\left|\int_X f\,d\mu\right| \leq \int_X |f|\,d\mu.$$

Proof. There is some $\theta \in \mathbb{R}$ such that

$$\left|\int_X f\,d\mu\right| = e^{i\theta}\int_X f\,d\mu = \int_X e^{i\theta} f\,d\mu.$$

Since the left hand side is real, so is the right hand side. Hence

$$\left|\int_X f\,d\mu\right| = \mathrm{Re}\int_X e^{i\theta} f\,d\mu = \int_X \mathrm{Re}(e^{i\theta} f)\,d\mu \leq \int_X |f|\,d\mu,$$

which completes the proof. □

If $f : \mathbb{R} \to \mathbb{R}$ is an integrable function, then its *Fourier transformation* is

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-i\xi x}\,d\mu(x).$$

The complex-valued function $x \mapsto f(x)e^{-i\xi x}$ is integrable for each $\xi \in \mathbb{R}$, since $|e^{-i\xi x}| = 1$. It follows at once from Lemma 2.2.4 that

$$|\widehat{f}(\xi)| \leq \int_X |f|\,d\mu$$

so the function $\widehat{f} : \mathbb{R} \to \mathbb{C}$ is bounded when the function f is integrable. It is possible to use Lemma 2.2.4 and the Dominated Convergence Theorem to prove that $\widehat{f}$ is continuous (see Exercise 2.5). We do not need this fact to establish the following result.

Theorem 2.2.5 (The Riemann–Lebesgue lemma)**.** *Suppose that* $f\colon \mathbb{R} \to \mathbb{R}$ *is Lebesgue integrable. Then*

$$\lim_{\xi\to\pm\infty} \widehat{f}(\xi) = 0.$$

Proof. Suppose first that $f = c\mathbf{1}_I$ for a constant c and a finite interval $I = [a, b]$. By direct computation using the Riemann integral (justified by Theorem 1.1.1), we obtain

$$\widehat{f}(\xi) = \int_{[a,b]} ce^{-i\xi x}\, d\mu(x) = ic\frac{e^{-i\xi b} - e^{-i\xi a}}{\xi}.$$

From this it is clear that $f(\xi) \to 0$ as $\xi \to \infty$ or as $\xi \to -\infty$. To pass from this special case to the general case, we rely on Lemma 2.2.4 and an approximation argument based on Lemma 2.2.2 and Corollary 2.1.3. The details are left to the reader (see Exercise 2.6). □

2.3 The third principle

Our version of Littlewood's third principle is often called Egorov's theorem. It will be stated and proved for a general measure space and we require the following definition.

Definition. We say that the measure space $(X, \mathcal{A}, \mu)$ is *finite* if $\mu(X) < \infty$.

A typical example of a finite measure space is the Lebesgue measure and the σ-algebra of Lebesgue measurable sets on some finite interval $(I, \mathcal{A}, \mu)$.

Theorem 2.3.1 (Littlewood's third principle)**.** *Let* $(X, \mathcal{A}, \mu)$ *be a finite measure space and suppose that* $\{f_j\}_{j\geq 1}$ *is a sequence of measurable functions which converges pointwise to a function* f*. For every* $\varepsilon > 0$ *there is a set* A_ε *such that* $\mu(A_\varepsilon) \leq \varepsilon$ *and such that* $\{f_j\}_{j\geq 1}$ *converges uniformly to* f *on* $X \setminus A_\varepsilon$*.*

Remark. Theorem 2.3.1 assigns the word *nearly* in Littlewood's third principle the meaning that we can get uniform convergence by discarding a set of arbitrarily small measure.

Proof. Fix a positive integer n and consider the sequence of sets

$$A_{n,k} = \bigcup_{j=k}^{\infty} \left\{x \in X \,:\, |f(x) - f_j(x)| \geq \frac{1}{n}\right\}$$

for $k \geq 1$. Note first that $A_{n,k+1} \subseteq A_{n,k}$. Moreover, since $\{f_j\}_{j\geq 1}$ converges pointwise to f it follows that $\cap_{k\geq 1} A_{n,k} = \emptyset$. Since $\mu(X) < \infty$, it clearly holds that $\mu(A_{n,1}) < \infty$. We can therefore appeal to continuity of measure for decreasing sequences (see [4, Proposition 7.1.5]) to conclude that

$$\lim_{k\to\infty} \mu(A_{n,k}) = 0. \tag{2.3.1}$$

Given a real number $\varepsilon > 0$ and an integer $n \geq 1$, it follows from (2.3.1) that there is an integer k_n such that $\mu(A_{n,k_n}) \leq 2^{-n}\varepsilon$. Define

$$A_\varepsilon = \bigcup_{n=1}^{\infty} A_{n,k_n}.$$

By countable subadditivity and our choice of k_n, we find that

$$\mu(A_\varepsilon) \leq \sum_{n=1}^{\infty} \mu(A_{n,k_n}) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.$$

If $x \in X \setminus A_\varepsilon$, then

$$|f(x) - f_j(x)| < \frac{1}{n}$$

whenever $j \geq k_n$ because if the converse estimate holds, then $x \in A_{n,k_n} \subseteq A_\varepsilon$. This demonstrates that $\{f_j\}_{j\geq 1}$ converges uniformly to f on $X \setminus A_\varepsilon$. $\square$

Theorem 2.3.1 is no longer true if we remove the assumption $(X, \mathcal{A}, \mu)$ is a finite measure space (see Exercise 2.7).

2.4 Exercises

Exercise 2.1. Show that the assumption that E has finite Lebesgue measure in Theorem 2.1.1 cannot be dropped.

Hint. Consider the set $E = \bigcup_{j\geq 1}[2j-1, 2j]$ and draw a picture.

Exercise 2.2. Prove Corollary 2.1.3.

Exercise 2.3. Fill in the details in the proof of Theorem 2.2.1.

Exercise 2.4. We will say that a function $f\colon \mathbb{R} \to \mathbb{R}$ is *smooth* if the derivatives $f^{(k)}$ exist for all $k \geq 1$.

(a) Let $\delta > 0$. Prove that the function $g_\delta\colon \mathbb{R} \to \mathbb{R}$ defined by

$$g_\delta(x) = \begin{cases} \exp\left(-\frac{\delta}{1-x^2}\right) & \text{if } -1 < x < 1, \\ 0 & \text{else,} \end{cases}$$

is smooth.

(b) Prove that for every $\varepsilon > 0$, there is a $\delta > 0$ such that

$$\int_{\mathbb{R}} \left|g_\delta - \mathbf{1}_{[-1,1]}\right| d\mu \leq \varepsilon.$$

Hint. Draw a picture (or ask your computer to do it).

(c) Establish the following stronger version of Littlewood's second principle: Let $f\colon \mathbb{R} \to \mathbb{R}$ be a Lebesgue integrable function. For every $\varepsilon > 0$ there is a compactly supported smooth function $g\colon \mathbb{R} \to \mathbb{R}$ such that

$$\int_{\mathbb{R}} |f - g| \, d\mu \leq \varepsilon.$$

Exercise 2.5. Let f be a Lebesgue integrable function. Prove that

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-i\xi x} \, d\mu(x)$$

is continuous.

Hint. Use Lemma 2.2.4 and the Dominated Convergence Theorem.

Exercise 2.6. Finish the proof of Theorem 2.2.5.

Exercise 2.7. Show that the assumption that $(X, \mathcal{A}, \mu)$ is a finite measure space cannot be dropped in Theorem 2.3.1.

Hint. Consider the Lebesgue measure on the real line and the sequence of functions $\{f_n\}_{n\geq 1}$ defined by $f_n = \mathbf{1}_{[n,n+1]}$. We informally say that this sequence of functions converge to 0 by escaping to infinity.

Bibliography

[1] Erik Bédos, *Notes on Elementary Linear Analysis*, 2.1 ed., Department of Mathematics, University of Oslo, 2022.

[2] Jean Dieudonné, *History of functional analysis*, Notas de Matemática, vol. 77, North-Holland Publishing Co., Amsterdam-New York, 1981.

[3] Gerald B. Folland, *Real analysis*, second ed., Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication.

[4] Tom L. Lindstrøm, *Spaces—an introduction to real analysis*, Pure and Applied Undergraduate Texts, vol. 29, American Mathematical Society, Providence, RI, 2017.

[5] J. E. Littlewood, *Lectures on the Theory of Functions*, Oxford University Press, 1944.