An introduction to Lebesgue Measure and Fourier Analysis

Laurent W. Marcoux

December 20, 2022

Preface to the Third Edition - September 2022

We are now at the third instance of the notes. As always, there may be typos, and it is crucial that you analyse what is written and not just accept things without thinking.

Thanks to M. Esipova, M.E. Gusak, J. Vendryes, Y. Wu and D. Yang for catching some of the remaining typos/errors.

Preface to the Second Edition - July 2019

Welcome to the second edition, also known as "*Release Candidate 1*". Doubtless there are typos/(hopefully small) errors, and you are welcome to let me know if you find some. Oh, and make sure that you read the last sentence of the Preface to the First Edition...

I would like to thank K. Santone and A. Tiwary for catching a great number of typos/errors for me.

Preface to the First Edition - January 2018

These notes are a work in progress, and - this being the "first edition" - they are replete with typos. As of April 10, 2018, I have had the opportunity to look over Chapters 1 - 4. That is not to say that they are mistake-free. It is, instead, an admission that I simply haven't had the chance to look over the remaining Chapters. A student should approach these notes with the same caution he or she would approach buzz saws; *they can be very useful, but you should be thinking the whole time you have them in your hands.* Enjoy.

The reviews are in!

From the moment I picked your book up until I laid it down I was convulsed with laughter. Someday I intend reading it.

Groucho Marx

This is not a novel to be tossed aside lightly. It should be thrown with great force.

Dorothy Parker

The covers of this book are too far apart.

Ambrose Bierce

I read part of it all the way through.

Samuel Goldwyn

Reading this book is like waiting for the first shoe to drop.

Ralph Novak

Thank you for sending me a copy of your book. I'll waste no time reading it.

Moses Hadas

Sometimes you just have to stop writing. Even before you begin.

Stanislaw J. Lec

That's not writing, that's typing.

Truman Capote

Contents

1. Riemann integration

What the world needs is more geniuses with humility, there are so few of us left.

Oscar Levant

1.1. In first year, we study the Riemann integral for functions defined on intervals, taking values in $\mathbb{R}$. The same circle of ideas can be greatly extended, as we shall now see.

Note. Much of the theory contained in these notes applies equally well to the setting of real- or complex-valued functions and vector spaces. When writing a statement which remains valid in either context, we shall simply write $\mathbb{K}$ to denote the base field. In other words, we shall always have $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

1.2. Definition. *Let $\mathfrak{X}$ be a vector space over the field $\mathbb{K}$. A* ***seminorm*** *on $\mathfrak{X}$ is a function $\nu : \mathfrak{X} \to \mathbb{R}$ satisfying*

(i) $\nu(x) \geq 0$ *for all* $x \in \mathfrak{X}$.
(ii) $\nu(\kappa x) = |\kappa|\ \nu(x)$ *for all* $\kappa \in \mathbb{K}$ *and* $x \in \mathfrak{X}$.
(iii) $\nu(x + y) \leq \nu(x) + \nu(y)$ *for all* $x, y \in \mathfrak{X}$.

If $\nu(x) = 0$ implies that $x = 0$, we say that ν is a ***norm*** *on $\mathfrak{X}$, and refer to the ordered pair $(\mathfrak{X}, \nu)$ as a* ***normed linear space*** (NLS).

1.3. Remarks.

(a) Typically we shall denote seminorms by ν or μ. When the seminorm is known to be a norm, we shall more often write $\|\cdot\|$.
(b) Let ν be a seminorm on a vector space $\mathfrak{X}$. Let $z \in \mathfrak{X}$ denote the zero vector. It follows from condition (ii) above that

$$\nu(z) = \nu(0z) = 0\,\nu(z) = 0.$$

In other words, if $\|\cdot\|$ is a *norm* on $\mathfrak{X}$ and $x \in \mathfrak{X}$, then $\|x\| = 0$ *if and only if* $x = 0$.
(c) We leave it as an exercise for the reader to prove that condition (c) implies that $|\nu(x) - \nu(y)| \leq \nu(x - y)$ for all $x, y \in \mathfrak{X}$.
(d) In a standard abuse of terminology, we often refer to the *normed linear space* $\mathfrak{X}$, *equipped with the norm* $\|\cdot\|$.

1.4. Example. Let X be a compact, Hausdorff topological space. Let $\mathfrak{X} = \mathcal{C}(X, \mathbb{K}) := \{f : X \to \mathbb{K} | f \text{ is continuous}\}$. Given $y \in X$, define

$$\begin{array}{rccc} \nu_y : & \mathcal{C}(X, \mathbb{K}) & \to & \mathbb{R} \\ & f & \mapsto & |f(y)|. \end{array}$$

It is left as an exercise for the reader to verify that ν_y is a seminorm on $\mathcal{C}(X, \mathbb{K})$ for all $y \in \mathbb{K}$.

More generally, if $\Omega \subseteq X$, the map $\nu_\Omega : \mathcal{C}(X, \mathbb{K}) \to \mathbb{R}$ defined by

$$\nu_\Omega(f) := \sup_{x \in \Omega} |f(x)|$$

defines a seminorm on $\mathcal{C}(X, \mathbb{K})$.

We leave it to the reader to verify that ν_Ω is a norm if and only if Ω is dense in X. In that case, $\nu_\Omega = \nu_X$. This norm is *typically* denoted by $\| \cdot \|_\infty$, and is referred to as the **supremum norm** or "**sup norm**"on X. Having said that, in this course we shall be dealing with certain Banach spaces of equivalence classes of functions, equipped with an "essential supremum norm", and this norm is also typically denoted by $\| \cdot \|_\infty$. In order to distinguish between the two, in these notes we shall denote the supremum norm on $\mathcal{C}(X, \mathbb{K})$ by $\| \cdot \|_{\sup}$.

1.5. The norm $\| \cdot \|$ on a NLS $\mathfrak{X}$ gives rise to a metric via the formula:

$$\begin{array}{rccc} d : & \mathfrak{X} \times \mathfrak{X} & \to & \mathbb{R} \\ & (x, y) & \mapsto & \|x - y\|. \end{array}$$

We refer to this as the **metric induced by the norm**. (That this is indeed a metric is left as an exercise to the reader.) When referring to metric properties of $\mathfrak{X}$, it is understood that we are referring to the metric induced by the norm, unless it is explicitly stated otherwise.

1.6. Definition. *A normed linear space $(\mathfrak{X}, \| \cdot \|)$ is said to be a* ***Banach space*** *if $(\mathfrak{X}, d)$ is a complete metric space, where d is the metric on $\mathfrak{X}$ induced by the norm.*

1.7. Examples.

(a) The motivating example is $\mathfrak{X} = \mathbb{K}$ itself, where the norm is given by the absolute value function. Since $(\mathbb{K}, | \cdot |)$ is complete, it is a Banach space.

Of course, $\mathbb{C}$ is a one-dimensional Banach space over $\mathbb{C}$, and a two-dimensional Banach space over $\mathbb{R}$.

(b) Let $N \geq 1$ be an integer. For $x = (x_n)_{n=1}^N \in \mathbb{K}^N$, we define three functions:
- $\|x\|_1 := |x_1| + |x_2| + \cdots + |x_N|$;
- $\|x\|_\infty := \max(|x_1|, |x_2|, \ldots, |x_N|)$; and
- $\|x\|_2 := \left(\sum_{n=1}^N |x_n|^2\right)^{\frac{1}{2}}$.

It is a routine exercise that $\| \cdot \|_1$ and $\| \cdot \|_\infty$ define norms on $\mathbb{K}^N$, and that $\mathbb{K}^N$ becomes a Banach space when equipped with either of these norms.

It is a slightly more interesting exercise (left to the reader) to show that $(\mathbb{K}^N, \| \cdot \|_2)$ is also a Banach space. The standard proof that $\| \cdot \|_2$ satisfies the triangle inequality as a function on $\mathbb{K}^N$ relies on the Cauchy-Schwarz Inequality. See Chapter 7 for more details.

(c) More generally, if $N \geq 1$ is an integer and $1 < p < \infty$, we may define $\|\cdot\|_p : \mathbb{K}^N \to \mathbb{R}$ via

$$\|(x_n)_{n=1}^N\|_p = \left(\sum_{k=1}^N |x_k|^p\right)^{\frac{1}{p}}$$

to obtain a norm on $\mathbb{K}^N$. The proof of the triangle inequality is somewhat more delicate than in the above cases. We shall return to this in Chapter 6, and in the assignments.

It is an interesting fact (which the reader is encouraged to prove) that for $x \in \mathbb{K}^n$,

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

(d) We saw in Example 1.4 that $\mathcal{C}([0,1],\mathbb{K})$ is a NLS when equipped with the norm

$$\|f\|_{\sup} = \sup_{x\in[0,1]} |f(x)| = \max_{x\in[0,1]} |f(x)|.$$

It is clear that convergence of a sequence $(f_n)_n$ in this norm to a function $f \in \mathcal{C}([0,1],\mathbb{K})$ is simply uniform convergence of $(f_n)_{n=1}^\infty$ to f, as studied in first-year calculus.

Once again, we leave it as an exercise for the reader to show that $(\mathcal{C}([0,1],\mathbb{K}), \|\cdot\|_{\sup})$ is complete, and is thus a Banach space.

(e) With $\mathfrak{X} = \mathcal{C}([0,1],\mathbb{K})$ as above, define

$$\|f\|_1 := \int_0^1 |f(x)|dx.$$

Then $(\mathcal{C}([0,1],\mathbb{K}), \|\cdot\|_1)$ is a NLS, but it is *not* a Banach space. The details are yet again left to the reader.

(f) Let $(\mathfrak{X}, \|\cdot\|_{\mathfrak{X}})$ and $(\mathfrak{Y}, \|\cdot\|_{\mathfrak{Y}})$ be normed linear spaces over the field $\mathbb{K}$. Let $T : \mathfrak{X} \to \mathfrak{Y}$ be a $\mathbb{K}$-linear map. Consider

$$\|T\| := \sup\{\|Tx\|_{\mathfrak{Y}} : x \in \mathfrak{X},\ \|x\|_{\mathfrak{X}} \leq 1\}.$$

Set

$$\mathcal{B}(\mathfrak{X},\mathfrak{Y}) := \{T : \mathfrak{X} \to \mathfrak{Y} \mid T \text{ is linear and } \|T\| < \infty\}.$$

As we shall see in the assignments, $\mathcal{B}(\mathfrak{X},\mathfrak{Y})$ is a vector space over $\mathbb{K}$, and $\|\cdot\| : \mathcal{B}(\mathfrak{X},\mathfrak{Y}) \to \mathbb{R}$ defines a norm on $\mathcal{B}(\mathfrak{X},\mathfrak{Y})$, referred to as the **operator norm** on $\mathcal{B}(\mathfrak{X},\mathfrak{Y})$. One can show that $(\mathcal{B}(\mathfrak{X},\mathfrak{Y}), \|\cdot\|)$ is complete if and only if $(\mathfrak{Y}, \|\cdot\|_{\mathfrak{Y}})$ is complete.

1.8. Definition. *Let $(\mathfrak{X}, \|\cdot\|)$ be a Banach space, $a < b \in \mathbb{R}$, and $f : [a, b] \to \mathfrak{X}$ be a function.*

*A **partition** of $[a, b]$ is a finite set*

$$P := \{a = p_0 < p_1 < \cdots < p_N = b\}$$

*for some integer $N \geq 1$. The set of all partitions of $[a, b]$ is denoted by $\mathcal{P}[a, b]$. Given P as above, a finite set $P^* = \{p_k^* : 1 \leq k \leq N\}$ satisfying $p_{k-1} \leq p_k^* \leq p_k$, $1 \leq k \leq N$ is said to be a set of **test values** for the partition. We then define the corresponding **Riemann sum***

$$S(f, P, P^*) := \sum_{k=1}^{N} f(p_k^*)(p_k - p_{k-1}).$$

1.9. Remarks.

(a) When $(\mathfrak{X}, \|\cdot\|) = (\mathbb{R}, |\cdot|)$, then this is the usual Riemann sum that one studies in first-year calculus. In particular, in the case where $0 \leq f(x)$ for all $x \in [a, b]$, we see that $S(f, P, P^*)$ estimates the area under the curve $y = f(x)$, $x \in [a, b]$.

(b) Observe that in general,

$$\frac{1}{b-a} S(f, P, P^*) = \sum_{k=1}^{N} \lambda_k f(p_k^*),$$

where $\lambda_k := \dfrac{p_k - p_{k-1}}{b-a}$, $1 \leq k \leq N$. Clearly $\lambda_k \geq 0$ for all k, while $\sum_{k=1}^{N} \lambda_k = 1$.

Thus $\dfrac{1}{b-a} S(f, P, P^*)$ is a convex combination of the $f(p_k^*)$'s, and as such, "*averages*" f over $[a, b]$.

1.10. Example. Let $\mathfrak{X} = \mathcal{C}([-\pi, \pi], \mathbb{C})$, equipped with the supremum norm $\|\cdot\|_{\sup}$. Let $1 \leq n \in \mathbb{N}$ be a fixed integer and consider the function $f : [0, 1] \to \mathfrak{X}$ given by

$$[f(x)](\theta) = e^{2\pi x} \sin(n\theta) + \cos x \ \cos(n\theta), \quad \theta \in [-\pi, \pi].$$

If $P = \{0, \frac{1}{10}, \frac{1}{2}, 1\}$, and if $P^* = \{\frac{1}{50}, \frac{1}{3}, \frac{4}{5}\}$, then

$$\begin{aligned} S(f, P, P^*) = &(e^{\pi/25} \sin(n\theta) + \cos(1/50)\cos(n\theta))(\frac{1}{10} - 0) \\ &+ (e^{2\pi/3} \sin(n\theta) + \cos(1/3)\cos(n\theta))(\frac{1}{2} - \frac{1}{10}) \\ &+ (e^{8\pi/5} \sin(n\theta) + \cos(4/5)\cos(n\theta))(1 - \frac{1}{2}). \end{aligned}$$

1.11. Definition. *Let $a < b$ be real numbers and $(\mathfrak{X}, \|\cdot\|)$ be a Banach space. We say that a function $f : [a, b] \to \mathfrak{X}$ is* ***Riemann integrable*** *if there exists a vector $x_0 \in \mathfrak{X}$ such that for any $\varepsilon > 0$, there exists a partition $P \in \mathcal{P}[a, b]$ with the property that if Q is any refinement of P and Q^* is any choice of test values for Q, then*

$$\|x_0 - S(f, Q, Q^*)\| < \varepsilon.$$

1.12. Remark. We note that if such a vector x_0 as above exists, then it is unique. Indeed, suppose that $y_0 \in \mathfrak{X}$ also satisfies the condition in Definition 1.11. If $y_0 \neq x_0$, we let $\varepsilon = \|y_0 - x_0\|/2 > 0$. Choose partitions $P_1 \in \mathcal{P}[a, b]$ (resp. $P_2 \in \mathcal{P}[a, b]$) as in the definition of Riemann integrability corresponding to ε and x_0 (resp. corresponding to ε and y_0). Let $R = P_1 \cup P_2$, so that R is a common refinement of P_1 and P_2.

If Q is any refinement of R and if Q^* is any set of test values for Q, then – noting that Q is again a common refinement of P_1 and P_2 – we see that

$$\begin{aligned} 2\varepsilon &= \|y_0 - x_0\| \\ &\leq \|y_0 - S(f, Q, Q^*)\| + \|S(f, Q, Q^*) - x_0\| \\ &< \varepsilon + \varepsilon \\ &= 2\varepsilon, \end{aligned}$$

an obvious contradiction. Thus $x_0 = y_0$, and the Riemann integral is unique.

When it exists, we refer to this unique vector x_0 as the **Riemann integral** of f over $[a, b]$, and we write

$$x_0 = \int_a^b f = \int_a^b f(s) ds.$$

As is the case with the usual version of Riemann integration, the usefulness of the Cauchy Criterion below is that it allows us to verify that a given function is Riemann integrable without first having to know what its integral is.

1.13. Theorem. [**The Cauchy Criterion for Riemann integrability.**]
Let $\mathfrak{X}$ be a Banach space, $a < b$ be real numbers and $f : [a, b] \to \mathfrak{X}$ be a function. The following conditions are equivalent.

(a) *f is Riemann integrable over $[a, b]$.*
(b) *For all $\varepsilon > 0$ there exists a partition $R \in \mathcal{P}[a, b]$ with the property that if P and Q are refinements of R, and if P^* and Q^* are test values for P and Q respectively, then*

$$\|S(f, P, P^*) - S(f, Q, Q^*)\| < \varepsilon.$$

Proof.

(a) implies (b). This is a standard argument which is left as an exercise for the reader.

(b) implies (a).

For each integer $n \ge 1$, choose a partition $R_n \in \mathcal{P}[a,b]$ such that for all refinements $P, Q \supseteq R_n$ of R_n, and any associated choices of test values P^* and Q^* we have

$$\|S(f,P,P^*) - S(f,Q,Q^*)\| < \frac{1}{n}.$$

For each $N \ge 1$, set $W_N := \cup_{n=1}^N R_n$ and fix a choice W_N^* of test values for W_N. Define $x_n = S(f, W_n, W_n^*)$, $n \ge 1$.

If $n_2 \ge n_1 \ge N \ge 1$, then

$$\begin{aligned}\|x_{n_2} - x_{n_1}\| &= \|S(f, W_{n_2}, W_{n_2}^*) - S(f, W_{n_1}, W_{n_1}^*)\| \\ &< \frac{1}{n_1} \le \frac{1}{N},\end{aligned}$$

as W_{n_1} and W_{n_2} are both refinements of R_N.

From this it readily follows that $(x_n)_{n=1}^\infty$ is a Cauchy sequence in $\mathfrak{X}$. Since $\mathfrak{X}$ is a Banach space and as such a complete metric space, we find that

$$x := \lim_{n\to\infty} x_n$$

exists in $\mathfrak{X}$. There remains to show that $x = \int_a^b f$.

To that end, let $\varepsilon > 0$ and choose $N > 0$ such that

(i) $\frac{1}{N} < \frac{\varepsilon}{2}$, and

(ii) $k \ge N$ implies that $\|x - x_k\| < \frac{\varepsilon}{2}$.

If V is a refinement of W_N, then it is also a refinement of R_N, and hence for any choice V^* of test values for V,

$$\begin{aligned}\|x - S(f,V,V^*)\| &\le \|x - x_N\| + \|x_N - S(f,V,V^*)\| \\ &< \frac{1}{N} + \|S(f, W_N, W_N^*) - S(f,V,V^*)\| \\ &< \frac{\varepsilon}{2} + \frac{1}{N} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon.\end{aligned}$$

By definition, $x = \int_a^b f$, and f is Riemann integrable over $[a,b]$.

□

1.14. Theorem. *Let $(\mathfrak{X}, \|\cdot\|)$ be a Banach space and $a < b \in \mathbb{R}$. If $f : [a,b] \to \mathfrak{X}$ is continuous, then f is Riemann integrable over $[a,b]$.*

Proof. The proof is a routine adaptation of the fact that every continuous, real-valued function on $[a,b]$ is Riemann integrable.

Since f is continuous on the compact set $[a,b]$, and since $\mathfrak{X}$ is a metric space, we see that f is *uniformly* continuous on $[a,b]$. Let $\varepsilon > 0$ and choose $\delta > 0$ such that $x, y \in [a,b]$ and $|x-y| < \delta$ implies that $\|f(x) - f(y)\| < \dfrac{\varepsilon}{2(b-a)}$.

Let $R = \{a = r_0 < r_1 < r_2 < \cdots < r_N = b\} \in \mathcal{P}[a,b]$ be a partition of $[a,b]$ such that $r_j - r_{j-1} < \delta$ for all $1 \le j \le N$, and let $R^* = \{r_j^*\}_{j=1}^N$ be a set of test values for R.

Suppose that $P = \{p_i\}_{i=0}^M$ is any refinement of R, and that $P^* = \{p_i^*\}_{i=1}^M$ is a set of test values for P. Then we can find a sequence $0 = k_0 < k_1 < k_2 < \cdots < k_N := M$ such that

$$p_{k_j} = r_j, \quad 1 \le j \le N.$$

In other words,

$$P = \{a = r_0 = p_0 = p_{k_0} < p_1 < \cdots < p_{k_1} = r_1 < p_{k_1+1} < \cdots < p_{k_2} = r_2 < \cdots < p_{k_N} = r_N = b\}.$$

Now $S(f,R,R^*) = \sum_{j=1}^N f(r_j^*)(r_j - r_{j-1})$. while $S(f,P,P^*) = \sum_{i=1}^M f(p_i^*)(p_i - p_{i-1})$.

But then

$$\begin{aligned} S(f,R,R^*) &= \sum_{j=1}^N f(r_j^*)(r_j - r_{j-1}) \\ &= \sum_{j=1}^N f(r_j^*) \sum_{i=k_{j-1}+1}^{k_j} (p_i - p_{i-1}) \\ &= \sum_{j=1}^N \sum_{i=k_{j-1}+1}^{k_j} f(r_j^*)(p_i - p_{i-1}), \end{aligned}$$

while

$$S(f,P,P^*) = \sum_{j=1}^N \sum_{i=k_{j-1}+1}^{k_j} f(p_i^*)(p_i - p_{i-1}).$$

But for $k_{j-1} + 1 \le i \le k_j$, we have that $r_{j-1} \le r_j^*, p_i^* \le r_j$, and therefore $|r_j^* - p_i^*| \le r_j - r_{j-1} < \delta$. From our estimate above, we see that

$$\begin{aligned} \|S(f,R,R^*) - S(f,P,P^*)\| &= \|\sum_{j=1}^N \sum_{i=k_{j-1}+1}^{k_j} (f(r_j^*) - f(p_i^*))(p_i - p_{i-1})\| \\ &\le \sum_{j=1}^N \sum_{i=k_{j-1}+1}^{k_j} \|f(r_j^*) - f(p_i^*)\|(p_i - p_{i-1}) \\ &< \sum_{j=1}^N \sum_{i=k_{j-1}+1}^{k_j} \frac{\varepsilon}{2(b-a)}(p_i - p_{i-1}) \\ &= \frac{\varepsilon}{2(b-a)} \sum_{j=1}^N \sum_{i=k_{j-1}+1}^{k_j} (p_i - p_{i-1}) \end{aligned}$$

$$= \frac{\varepsilon}{2(b-a)} \sum_{i=1}^{M} (p_i - p_{i-1})$$
$$= \frac{\varepsilon}{2(b-a)} (p_M - p_0)$$
$$= \frac{\varepsilon}{2}.$$

As such, if Q is any other refinement of R and if Q^* is any set of test values for Q, then

$$\|S(f,Q,Q^*) - S(f,R,R^*)\| < \frac{\varepsilon}{2},$$

whence

$$\begin{aligned} \|S(f,Q,Q^*) - S(f,P,P^*)\| &\le \|S(f,Q,Q^*) - S(f,R,R^*)\| + \\ &\qquad \|S(f,R,R^*) - S(f,P,P^*)\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

By the Cauchy Criterion 1.13, f is Riemann integrable.

□

We provide a second proof of the above result in the Appendix to this section.

We now turn our attention to real-valued functions, with the intention of showing what one of the failings of the Riemann integral is, and thereby (hopefully) motivating the study of the Lebesgue integral in the forthcoming chapters.

1.15. Example. Given a subset $E \subseteq \mathbb{R}$ of the real numbers, we define the **characteristic function** (or **indicator function**) of E to be

$$\begin{aligned} \chi_E : \mathbb{R} &\to \mathbb{R} \\ s &\mapsto \begin{cases} 1 & \text{if } s \in E \\ 0 & \text{if } s \notin E. \end{cases} \end{aligned}$$

Consider $E = \mathbb{Q} \cap [0,1]$. We claim that the Riemann integral

$$\int_0^1 \chi_E(s) ds$$

does not exist.

Indeed, observe that E and $[0,1] \setminus E$ are each dense in $[0,1]$. Let

$$P = \{0 = p_0 < p_1 < p_2 < \cdots < p_N = 1\}$$

be any partition of $[0,1]$. For $1 \le n \le N$, choose $p_n^* \in \mathbb{Q} \cap [p_{n-1}, p_n]$, and choose $q_n^* \in \mathbb{Q}^c \cap [p_{n-1}, p_n]$, so that $P^* := \{p_1^*, p_2^*, \ldots, p_N^*\}$ and $Q^* := \{q_1^*, q_2^*, \ldots, q_N^*\}$ are both sets of test values for P. Then

$$\begin{aligned} S(\chi_E, P, P^*) &= \sum_{n=1}^{N} f(p_n^*)(p_n - p_{n-1}) \\ &= \sum_{n=1}^{N} 1(p_n - p_{n-1}) \\ &= p_N - p_0 \\ &= 1 - 0 \\ &= 1, \end{aligned}$$

while

$$\begin{aligned} S(\chi_E, P, Q^*) &= \sum_{n=1}^{N} f(q_n^*)(p_n - p_{n-1}) \\ &= \sum_{n=1}^{N} 0\,(p_n - p_{n-1}) \\ &= 0. \end{aligned}$$

It follows that for any choice of $0 < \varepsilon < 1$, the Cauchy Criterion fails for χ_E, and as such χ_E is not Riemann integrable over $[0,1]$.

1.16. Remark. Recall that $\mathbb{Q}$ and thus $E := \mathbb{Q} \cap [0,1]$ is denumerable. Write $E = \{q_n\}_{n=1}^{\infty}$, and define $E_n := \{q_1, q_2, \dots, q_n\}$, $n \geq 1$. With $f_n := \chi_{E_n}$, $n \geq 1$, we find that

$$0 \leq f_1 \leq f_2 \leq f_3 \leq \cdots \qquad \leq \chi_E.$$

In fact, for each $s \in [0,1]$, we find that

$$\chi_E(s) = \lim_{n\to\infty} f_n(s).$$

We say that the sequence $(f_n)_{n=1}^{\infty}$ is **increasing** and that it **converges pointwise** to χ_E.

Since each f_n is continuous except at finitely many points in the interval $[0,1]$ (in fact it is constantly equal to zero except at finitely many points), it is routine to verify that each f_n is Riemann integrable and that

$$\int_0^1 f_n(s)ds = 0,$$

$n \geq 1$. Nevertheless,

$$0 = \lim_{n\to\infty} \int_0^1 f_n(s)ds \neq \int_0^1 (\lim_{n\to\infty} f_n)(s)ds = \int_0^1 \chi_E(s)ds,$$

as the latter quantity does not exist.

We now seek to develop a more flexible and "forgiving" integral which will correct such "pathological" behaviour.

1.17. A heuristic approach. Whereas the Riemann integral partitions the domain of a function $f : [a, b] \to \mathbb{R}$ into intervals, and associates to each such partition a step function, our new approach will partition the range of f into subintervals $[y_{k-1}, y_k]$, $1 \le k \le N$. We then set $E_k = \{x \in [a, b] : f(x) \in [y_{k-1}, y_k]\}$, $1 \le k \le N$.

This allows us to estimate "$\int_a^b f$" by so-called **simple functions**

$$\sum_{k=1}^{N} y_k m E_k,$$

where mE_k denotes the "measure" (a generalization of length) of E_k. Since E_k need not be a particularly nice set (one need only consider $f = \chi_{\mathbb{Q} \cap [0,1]}$ as above), our first goal is to make sense of the "measure" or "length" of as many subsets of $\mathbb{R}$ as possible.

Appendix to Section 1.

1.18. There are more notions of integration in a Banach space than just Riemann integration. Indeed, in general, there exist more notions of integration than one can shake a stick at, even if one has strong arms, a very light stick, ample dexterity and a solid and enviable history of stick-shaking.

This notion, however, will prove sufficient for our purposes. In Theorem 1.14, we showed that every continuous function from a closed, bounded interval in $\mathbb{R}$ to a Banach space is Riemann integrable over that closed interval. A minor modification will show that every piecewise continuous, bounded function from a closed interval to a Banach space is Riemann integrable in the sense of Definition 1.11.

1.19. Culture. The notion of integrating in a Banach space is not simply some arcane and useless generalization of integration of real- or complex-valued functions. Let $1 \in n$ be an integer, and let $\mathcal{H} := \mathbb{C}^n$, equipped with the Euclidean norm $\|\cdot\|_2$. Consider $(\mathcal{B}(\mathbb{C}^n), \|\cdot\|)$, where $\|\cdot\|$ denotes the operator norm on $\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathbb{C}^n)$.

We leave it as an exercise for the reader to show that every linear map $T : \mathbb{C}^n \to \mathbb{C}^n$ satisfies $\|T\| < \infty$, and thus we may identify $\mathcal{B}(\mathbb{C}^n)$ with $\mathbb{M}_n(\mathbb{C})$, the space of $n \times n$ matrices with entries in $\mathbb{C}$ with respect to a fixed, orthonormal basis $\{e_k\}_{k=1}^n$ for $\mathbb{C}^n$.

Let $d_1, d_2, \ldots, d_n \in \mathbb{C}$ be n *distinct* points, and let $D \in \mathcal{B}(\mathbb{C}^n)$ be the unique linear operator whose corresponding matrix is the diagonal matrix

$$\begin{bmatrix} d_1 & 0 & & \ldots & 0 \\ & d_2 & 0 & \ldots & 0 \\ & & \ddots & & \\ & & & d_{n-1} & 0 \\ & & & & d_n \end{bmatrix}.$$

Suppose that $\varnothing \neq \Delta \subseteq \{1, 2, \ldots, n\}$.

Let $\Gamma \subseteq \mathbb{C}$ be a piecewise smooth curve in $\mathbb{C}$ for which

$$\operatorname{ind}(\Gamma, d_k) = \begin{cases} 1 & \text{if } k \in \Delta \\ 0 & \text{otherwise.} \end{cases}$$

It can be shown that

$$P := \frac{1}{2\pi i} \int_\Gamma (sI - D)^{-1} \, ds$$

is the orthogonal projection $P = \operatorname{diag}(p_1, p_2, \ldots, p_n)$, where $p_k = 1$ if $k \in \Delta$ and $p_k = 0$ otherwise.

The astute reader will have observed the striking similarity of the integral above to Cauchy's Integral Formula. This is not a coincidence. Such integrals are studied in much greater generality in the theory of Banach algebras.

As promised, here is a second proof of Theorem 1.14.

1.20. Theorem. *(Theorem 1.14 revisited). Let* $(\mathfrak{X}, \|\cdot\|)$ *be a Banach space and* $a < b \in \mathbb{R}$. *If* $f : [a, b] \to \mathfrak{X}$ *is continuous, then* f *is Riemann integrable over* $[a, b]$.

Proof. We shall in fact show that if $P_N \in \mathcal{P}([a, b])$ is a regular partition of $[a, b]$ into 2^N subintervals of equal length $\dfrac{b-a}{2^N}$, and if $P_N^* = P_N \setminus \{a\}$ is the set of test values for P_N consisting of "right-hand endpoints" of the subintervals of P_N, then the sequence $(S(f, P_N, P_N^*))_{N=1}^{\infty}$ converges in $\mathfrak{X}$ to

$$\int_a^b f(s)\, ds.$$

We begin by showing that the sequence $(S(f, P_N, P_N^*))_{N=1}^{\infty}$ is Cauchy, and therefore converges to *something* in $\mathfrak{X}$, which we temporarily designate by y.

Since f is continuous on the compact set $[a, b]$, and since $\mathfrak{X}$ is a metric space, we see that f is *uniformly* continuous on $[a, b]$. Let $\varepsilon > 0$ and choose $\delta > 0$ such that $x, y \in [a, b]$ and $|x - y| < \delta$ implies that $\|f(x) - f(y)\| < \dfrac{\varepsilon}{b-a}$.

For each $N \ge 1$, let P_N be as above, and choose $M \ge 1$ such that $\dfrac{b-a}{2^M} < \delta$. If $K \ge L \ge M$, then $P_M \subseteq P_L \subseteq P_K$; indeed, writing

$$P_L = \{a = p_0 < p_1 < \cdots < p_{2^L} = b\},$$

and

$$P_K = \{a = q_0 < q_1 < \cdots < q_{2^K} = b\},$$

and setting $p_j^* = p_j$, $1 \le j \le 2^L$ and $q_s^* = q_s$, $1 \le s \le 2^K$, we see that $p_j = q_{j\, 2^{K-L}}$ for all $0 \le j \le 2^L$, and that for $1 \le j \le 2^L$,

$$\|f(p_j^*) - f(q_s^*)\| < \frac{\varepsilon}{b-a}, \quad (j-1)\, 2^{K-L} < s \le j\, 2^{K-L}.$$

Thus

$$\begin{aligned}
\|S(f, P_L, P_L^*) - S(f, P_K, P_K^*)\| &= \left\| \sum_{j=1}^{2^L} \sum_{s=(j-1)\, 2^{K-L}+1}^{j\, 2^{K-L}} (f(p_j) - f(q_s))(q_s - q_{s-1}) \right\| \\
&\le \sum_{j=1}^{2^L} \sum_{s=(j-1)\, 2^{K-L}+1}^{j\, 2^{K-L}} \|f(p_j) - f(q_s)\|(q_s - q_{s-1}) \\
&\le \sum_{j=1}^{2^L} \sum_{s=(j-1)\, 2^{K-L}+1}^{j\, 2^{K-L}} \frac{\varepsilon}{b-a}(q_s - q_{s-1}) \\
&= \frac{\varepsilon}{b-a} \sum_{s=1}^{2^K} q_s - q_{s-1} \\
&= \frac{\varepsilon}{b-a}(b-a) = \varepsilon.
\end{aligned}$$

So if we set $y_N = S(f, P_N, P_N^*)$, $N \geq 1$, then the above argument shows that $(y_N)_{N=1}^{\infty}$ is a Cauchy sequence in $\mathfrak{X}$, and as such, admits a limit $y \in \mathfrak{X}$.

There remains to show that $y = \int_a^b f(s)\ ds$. The proof is almost identical to that above.

With $\varepsilon > 0$, choose $T \geq 1$ such that $\dfrac{b-a}{2^T} < \delta$ and also such that

$$\|y - S(f, P_T, P_T^*)\| < \varepsilon.$$

Let $R = \{a = r_0 < r_1 < \cdots < r_J = b\}$ be any refinement of $P_T = \{a = p_0 < p_1 < \cdots < p_{2^T} = b\}$. Thus there exists a sequence

$$0 = j_0 < j_1 < \cdots < j_{2^T} = J$$

such that

$$r_{j_k} = p_k, \quad 0 \leq k \leq 2^T.$$

Let R^* be any set of test values for R. If $j_{k-1} + 1 \leq s \leq j_k$, then $|p_k^* - r_s^*| \leq |p_k - p_{k-1}| = \dfrac{b-a}{2^T} < \delta$, $1 \leq k \leq 2^T$ and so

$$\begin{aligned}
\|S(f, P_T, P_T^*) - S(f, R, R^*)\| &\leq \sum_{k=1}^{2^T} \sum_{s=j_{k-1}+1}^{j_k} \|f(p_k^*) - f(r_s^*)\|(r_s - r_{s-1}) \\
&< \frac{\varepsilon}{b-a} \sum_{k=1}^{2^T} \sum_{s=j_{k-1}+1}^{j_k} (r_s - r_{s-1}) \\
&= \frac{\varepsilon}{b-a}(b-a) = \varepsilon.
\end{aligned}$$

Thus

$$\|y - S(f, R, R^*)\| \leq \|y - S(f, P_T, P_T^*)\| + \|S(f, P_T, P_T^*) - S(f, R, R^*)\| < \varepsilon + \varepsilon = 2\varepsilon.$$

This clearly shows that

$$y = \int_a^b f(s)\ ds.$$

□

Exercises for Section 1.

Exercise 1.1.

Let $\varnothing \neq X$ be a compact, Hausdorff space. Prove that for each $\varnothing \neq \Omega \subseteq X$, the function

$$\begin{array}{rccc} \nu_\Omega: & \mathcal{C}(X,\mathbb{K}) & \to & \mathbb{R} \\ & f & \mapsto & \sup_{x\in\Omega} |f(x)| \end{array}$$

defines a seminorm on $\mathcal{C}(X,\mathbb{K})$, and that it is a norm if and only if Ω is dense in X.

Exercise 1.2.

(a) Let ν be a seminorm on a vector space $\mathfrak{Y}$ over $\mathbb{K}$. Prove that

$$|\nu(x) - \nu(y)| \le \nu(x-y) \qquad \text{for all } x, y \in \mathfrak{Y}.$$

(b) Let $(\mathfrak{X}, \|\cdot\|)$ be a NLS. Prove that the map

$$\begin{array}{rccc} d: & \mathfrak{X}\times\mathfrak{X} & \to & \mathbb{R} \\ & (x,y) & \mapsto & \|x-y\| \end{array}$$

defines a metric on $\mathfrak{X}$.

Exercise 1.3.

Let $N \ge 1$ be an integer. Define three functions from $\mathbb{K}^N$ to $\mathbb{R}$ as follows: for $x = (x_n)_{n=1}^N \in \mathbb{K}^N$, we set

(a) $\|x\|_1 := |x_1| + |x_2| + \cdots + |x_N|$;
(b) $\|x\|_\infty := \max(|x_1|, |x_2|, \ldots, |x_N|)$; and
(c) $\|x\|_2 := \left(\sum_{n=1}^N |x_n|^2\right)^{\frac{1}{2}}$.

Prove that each of these functions defines a norm on $\mathbb{K}^N$.

Exercise 1.4.

Let $N \ge 1$ be an integer, and let $x \in \mathbb{K}^N$. Prove that

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

Exercise 1.5.

(a) Prove that the NLS $(\mathcal{C}([0,1],\mathbb{K}), \|\cdot\|_{\sup})$ is complete, and that it is therefore a Banach space.
(b) Recall that $\|f\|_1 := \int_0^1 |f(x)| dx$ defines a norm on $\mathcal{C}([0,1],\mathbb{K})$. Prove that $(\mathcal{C}([0,1],\mathbb{K}), \|\cdot\|_1)$ is not complete.

Exercise 1.6.

Let $\mathfrak{X}$ be a Banach space, $a < b$ be real numbers and $f : [a,b] \to \mathfrak{X}$ be a function. Suppose that f is Riemann integrable over $[a,b]$.

Prove that for all $\varepsilon > 0$ there exists a partition $R \in \mathcal{P}[a,b]$ with the property that if P and Q are refinements of R, and if P^* and Q^* are test values for P and Q respectively, then

$$\|S(f,P,P^*) - S(f,Q,Q^*)\| < \varepsilon.$$

Exercise 1.7.

Let $\mathfrak{X}$ be a Banach space, $a < b$ be real numbers, and $g : [a,b] \to \mathfrak{X}$ be a bounded, piecewise-continuous function. Show that g is Riemann integrable over $[a,b]$.

Exercise 1.8.

Prove the claim from Remark 1.16, namely: $\int_0^1 f_n(s)\,ds = 0$ for all $n \geq 1$, where $\mathbb{Q} \cap [0,1] = \{q_n\}_{n=1}^{\infty}$, and where f_n is the characteristic function of $E_n := \{q_1, q_2, \ldots, q_n\}$, $n \geq 1$.

Exercise 1.9. This question will be used in Chapter 9.

Let $(\mathfrak{X}, \|\cdot\|)$ be a Banach space, and let $a < b \in \mathbb{R}$. Let $g : [a,b] \to \mathbb{K}$ and $f : [a,b] \to \mathfrak{X}$ be continuous, and set $\|f\|_{\sup} := \sup\{\|f(x)\| : x \in [a,b]\}$. Observe that $\|f\|_{\sup}$ is finite as $x \mapsto \|f(x)\|$ is continuous on the compact set $[a,b]$.

Then

$$\| \int_a^b f(x)g(x)dx\| \leq \|f\|_{\sup} \int_a^b |g(x)|dx.$$

2. Lebesgue outer measure

If you want to know what God thinks of money, just look at the people he gave it to.

Dorothy Parker

2.1. Our goal in this section is to define a "measure of length" for as many subsets of $\mathbb{R}$ as possible. We would like our new notion to agree with our intuition in the cases we know; for example, it seems reasonable to ask that our generalized notion of "length" of a finite interval (a,b) should be $(b-a)$ when $a < b$ in $\mathbb{R}$. We shall therefore use this as our starting point, and we shall use this intuition to extend our notion of "length" to a greater variety of sets by approximation.

For $a \le b \in \mathbb{R}$, we define the **length** of (a,b) to be $b-a$, and we write

$$\ell((a,b)) := b - a.$$

We also set $\ell(\varnothing) = 0$, and $\ell((-\infty, b)) = \ell((a,\infty)) = \ell((-\infty,\infty)) = \infty$ for all $a, b \in \mathbb{R}$. In this way we have defined $\ell(I)$ whenever I is an open interval in $\mathbb{R}$.

2.2. Definition. *Let $E \subseteq \mathbb{R}$. A countable collection $\{I_n\}_{n=1}^\infty$ of open intervals is said to be a* ***cover of*** *E* ***by open intervals*** *if $E \subseteq \cup_{n=1}^\infty I_n$.*

For each subset E of $\mathbb{R}$, we define a quantity $m^*E \in [0,\infty] := [0,\infty) \cup \{\infty\}$ as follows:

$$m^*E := \inf\{\sum_{n=1}^\infty \ell(I_n) : \{I_n\}_{n=1}^\infty \text{ a cover of } E \text{ by open intervals}\}.$$

In order to help make the text more readable, and since these are the only covers of sets we will consider in these notes, we shall abbreviate the expression "*cover of E by open intervals*" to "*cover of E*". For any set X, we denote by $\mathfrak{P}(X) = \{Y : Y \subseteq X\}$ the **power set** of X.

2.3. Definition. *Let $\varnothing \ne X$ be a set. An* ***outer measure*** *μ on X is a function*

$$\mu : \mathfrak{P}(X) \to [0,\infty]$$

which satisfies

(a) $\mu\varnothing = 0$;
(b) *if $E \subseteq F \subseteq X$, then $\mu E \le \mu F$. We say that μ is* ***monotone increasing;*** *and*
(c) *if $F_n \subseteq X$ for all $n \ge 1$, then*

$$\mu(\cup_{n=1}^\infty F_n) \le \sum_{n=1}^\infty \mu(F_n).$$

It is worth noting that by virtue of (b), condition (c) is equivalent to condition:

(d) if $E, F_1, F_2, F_3, \ldots \subseteq X$ and if $E \subseteq \cup_{n=1}^{\infty} F_n$, then

$$\mu E \leq \sum_{n=1}^{\infty} \mu F_n.$$

Condition (c) (or equivalently (d)) is generally referred to as the **countable subadditivity** or **σ-subadditivity** of μ.

2.4. Proposition. *The function m^* defined in Definition 2.2 is an outer measure on $\mathbb{R}$.*

Proof.

(a) Let $E = \varnothing$. With $I_n = \varnothing$, $n \geq 1$, it is clear that $\{I_n\}_{n=1}^{\infty}$ is a cover of E, and so

$$0 \leq m^* \varnothing \leq \sum_{n=1}^{\infty} \ell(I_n) = \sum_{n=1}^{\infty} 0 = 0.$$

Thus $m^* \varnothing = 0$.

(b) Let $E \subseteq F \subseteq \mathbb{R}$.

If $\{I_n\}_{n=1}^{\infty}$ is a cover of F, then it is also a cover of E. It follows immediately from the definition that

$$m^* E \leq m^* F.$$

(c) Suppose that $\{E_n\}_{n=1}^{\infty}$ is a countable collection of subsets of $\mathbb{R}$. We wish to prove that $m^*(\cup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} m^* E_n$.

If $\sum_{n=1}^{\infty} m^* E_n = \infty$, then we have nothing to prove. Thus we consider the case where $\sum_{n=1}^{\infty} m^* E_n < \infty$. Set $E := \cup_{n=1}^{\infty} E_n$.

Let $\varepsilon > 0$ and for each $n \geq 1$, choose a cover $\{I_k^{(n)}\}_{k=1}^{\infty}$ of E_n such that

$$\sum_{k=1}^{\infty} \ell(I_k^{(n)}) < m^* E_n + \frac{\varepsilon}{2^n}.$$

Then $\{I_k^{(n)}\}_{k,n=1}^{\infty}$ is a cover of E, and so

$$\begin{aligned} m^* E &\leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \ell(I_k^{(n)}) \\ &\leq \sum_{n=1}^{\infty} (m^* E_n + \frac{\varepsilon}{2^n}) \\ &= \left(\sum_{n=1}^{\infty} m^* E_n \right) + \varepsilon. \end{aligned}$$

(Here we have used the fact that if $0 \leq a_n \in \mathbb{R}$, $n \geq 1$, and if $\sigma : \mathbb{N} \to \mathbb{N}$ is any permutation, then $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\sigma(n)}$.)

Since $\varepsilon > 0$ was arbitrary,

$$m^* E \leq \sum_{n=1}^{\infty} m^* E_n.$$

□

2.5. Corollary. *Let $E \subseteq \mathbb{R}$ be a countable set. Then $m^*E = 0$.*

Proof. Suppose that E is denumerable, say $E = \{x_n\}_{n=1}^\infty$.

Let $\varepsilon > 0$ and for each $n \ge 1$, set $I_n = (x_n - \frac{\varepsilon}{2^{n+1}}, x_n + \frac{\varepsilon}{2^{n+1}})$. Then $\{I_n\}_{n=1}^\infty$ is a cover of E, and therefore

$$0 \le m^*E \le \sum_{n=1}^\infty \ell(I_n) = \sum_{n=1}^\infty \frac{\varepsilon}{2^n} = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we have $m^*E = 0$.

The case where E is finite is left as an exercise.

□

2.6. Corollary. *The outer measure of the rational numbers is 0; i.e. $m^*\mathbb{Q} = 0$.*

We have defined outer measure m^*E for *any* subset E of $\mathbb{R}$, and we have done this based upon an intuitive notion of what the length of an open interval (a, b) should be, namely $b - a$. At first glance, it seems obvious that $m^*(a, b) = \ell(b - a) = b - a$. But upon reflection, we see that this is *not* how $m^*(a, b)$ is defined. This leaves us with an interesting problem: how does our notion of measure $m^*(a, b)$ of an interval compare with this notion of length? On the one hand, it is clear that $m^*(a, b) \le \ell(a, b) = b - a$, since $I_1 := (a, b)$ and $I_n = \varnothing$, $n \ge 2$ yields a cover of (a, b). On the other hand, the notion of outer measure of (a, b) requires us to consider *all* covers of (a, b) by intervals, not only the obvious cover by the interval (a, b) itself. We now turn to this problem. It will prove useful to first consider the outer measure of closed, bounded intervals $[a, b]$, as these are compact. Because of this, we will be able to replace general covers of $[a, b]$ (by open intervals) with *finite* covers of $[a, b]$ (by open intervals).

2.7. Proposition. *Let $a < b \in \mathbb{R}$. Then*

(a) *$m^*[a, b] = b - a$, and therefore*
(b) *$m^*(a, b] = m^*[a, b) = m^*(a, b) = b - a$.*

Proof.

(a) Let $\varepsilon > 0$ and note that with $I_1 = (a - \frac{\varepsilon}{2}, b + \frac{\varepsilon}{2})$ and $I_n = \varnothing$, $n \ge 2$, the collection $\{I_n\}_{n=1}^\infty$ is a cover of $[a, b]$ by open intervals and thus

$$m^*[a, b] \le \sum_{n=1}^\infty \ell(I_n) = \ell(I_1) = (b - a) + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $m^*[a, b] \le b - a$.

We now turn the question of obtaining a lower bound for $m^*[a, b]$.

Suppose that $\{I_n\}_{n=1}^\infty$ is an cover of $[a, b]$ by open intervals. (Note: without loss of generality, we may assume that $I_n \ne \varnothing$, $1 \le n < \infty$.) We must show that $\sum_{n=1}^\infty \ell(I_n) \ge b - a$.

Since $[a,b]$ is compact, we can find a finite subcover $\{I_1, I_2, \dots, I_N\}$ of $[a,b]$. If $\ell(I_n) = \infty$ for some $1 \le n \le N$, then the inequality

$$\sum_{n=1}^{\infty} \ell(I_n) \ge \sum_{n=1}^{N} \ell(I_n) \ge b - a$$

trivially holds. Thus we shall assume that $\ell(I_n) < \infty$ for all $1 \le n \le N$, and we may then write $I_n = (a_n, b_n)$, $1 \le n \le N$.

Since $a \in [a,b] \subseteq \cup_{n=1}^{N} I_n$, we can find $1 \le n_1 \le N$ such that $a \in I_{n_1} = (a_{n_1}, b_{n_1})$. If $b_{n_1} > b$, we stop.

Otherwise, $a < b_{n_1} \le b$, so $b_{n_1} \in [a,b] \subseteq \cup_{n=1}^{N} I_n$, and we can find $n_2 \in \{1, 2, \dots, N\} \setminus \{n_1\}$ such that $b_{n_1} \in (a_{n_2}, b_{n_2})$. If $b_{n_2} > b$, we stop.

Otherwise, $a_{n_1} < a < b_{n_1} < b_{n_2} \le b$, so $b_{n_2} \in [a,b] \subseteq \cup_{n=1}^{N} I_n$, and we can find $n_3 \in \{1, 2, \dots, N\} \setminus \{n_1, n_2\}$ such that $b_{n_2} \in (a_{n_3}, b_{n_3})$. If $b_{n_3} > b$, we stop.

Eventually this process must end, since we have only $N < \infty$ intervals, and each stage n_k is chosen from among $\{1, 2, \dots, N\} \setminus \{n_1, n_2, \dots, n_{k-1}\}$. Suppose therefore that $1 \le M \le N$ is the minimal integer such that $b_{n_M} > b$. Then

$$[a,b] \subseteq \cup_{k=1}^{M} (a_{n_k}, b_{n_k}).$$

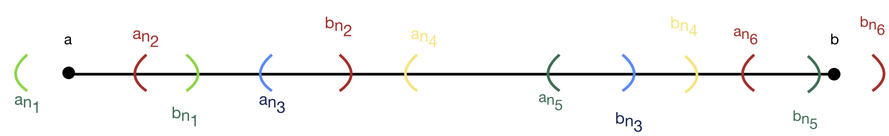


FIGURE 1. AN EXAMPLE WHERE $M = 6$.

Now,

$$\begin{aligned}
\sum_{n=1}^{\infty} \ell(I_n) &\ge \sum_{n=1}^{N} \ell(I_n) \\
&\ge \sum_{k=1}^{M} \ell((a_{n_k}, b_{n_k})) \\
&= (b_{n_1} - a_{n_1}) + (b_{n_2} - a_{n_2}) + \cdots + (b_{n_M} - a_{n_M}) \\
&= b_{n_M} + (b_{n_{M-1}} - a_{n_M}) + (b_{n_{M-2}} - a_{n_{M-1}}) + \cdots + (b_{n_1} - a_{n_2}) - a_{n_1} \\
&> b_{n_M} - a_{n_1} \\
&> b - a.
\end{aligned}$$

Since $\{I_n\}_{n=1}^{\infty}$ was an arbitrary cover of $[a,b]$, it follows that

$$m^*[a,b] \ge b - a.$$

Combining this with the reverse inequality above, we conclude that

$$m^*[a,b] = b - a.$$

(b) Consider the interval $(a,b]$.

For all $0 < \varepsilon < b-a$, we have that $[a+\varepsilon, b] \subseteq (a,b] \subseteq [a,b]$, and thus monotonicity of Lebesgue outer measure implies that

$$(b-a) - \varepsilon = m^*[a+\varepsilon, b] \le m^*(a,b] \le m^*[a,b] = b-a.$$

Since ε was arbitrary (subject to $0 < \varepsilon < b-a$), we conclude that $m^*(a,b] = b-a$.

The remaining cases are similar, and are left as an exercise for the reader.

□

2.8. Corollary. *Let $a, b \in \mathbb{R}$. Then*

$$m^*(-\infty, b) = m^*(-\infty, b] = m^*(a, \infty) = m^*[a, \infty) = m^*\mathbb{R} = \infty.$$

Proof. Consider the interval $(-\infty, b)$.

By monotonicity of Lebesgue outer measure, for each $n \ge 1$,

$$n = m^*[b-n, b) \le m^*(-\infty, b),$$

and thus $m^*(-\infty, b) = \infty$.

The remaining cases are similar, and are left as an exercise for the reader.

□

2.9. Corollary. *Let $a < b \in \mathbb{R}$. Then $[a,b]$ is uncountable.*

Proof. This follows immediately from Proposition 2.7 and Corollary 2.5.

□

2.10. Definition. *Let μ be an outer measure on $\mathbb{R}$. We say that μ is **translation invariant** if for all $E \subseteq \mathbb{R}$ and all $\kappa \in \mathbb{R}$, we have that*

$$\mu(E+\kappa) = \mu E,$$

where $E + \kappa := \{x + \kappa : x \in E\}$.

2.11. Proposition. *Lebesgue outer measure is translation-invariant on $\mathbb{R}$.*

Proof. Observe that $\{I_n\}_{n=1}^\infty$ is a cover of E (by open intervals) if and only if $\{I_n + \kappa\}_{n=1}^\infty$ is a cover of $E + \kappa$ (by open intervals).

Moreover, for any interval $I = (a,b)$, where $a < b$, we have that

$$\ell(I) = b - a = (b+\kappa) - (a+\kappa) = \ell(I+\kappa),$$

while if I is of the form $(-\infty, b)$, (a, ∞) or $(-\infty, \infty)$, then $\ell(I) = \infty = \ell(I+\kappa)$.

Thus

$$\begin{aligned} m^*E &= \inf\{\sum_{n=1}^\infty \ell(I_n) : \{I_n\}_{n=1}^\infty \text{ a cover of } E \text{ by open intervals}\} \\ &= \inf\{\sum_{n=1}^\infty \ell(I_n+\kappa) : \{I_n+\kappa\}_{n=1}^\infty \text{ a cover of } E+\kappa \text{ by open intervals}\} \\ &= m^*(E+\kappa). \end{aligned}$$

□

2.12. Countable subaddivity of Lebesgue outer measure guarantees that if E_n is a subset of $\mathbb{R}$ for all $n \geq 1$ and if $E = \cup_{n=1}^{\infty} E_n$, then

$$m^* E \leq \sum_{n=1}^{\infty} m^* E_n.$$

Of course, we can not expect equality – for example, we might have $E_n = [0,1]$ for all $n \geq 1$, in which case $E := \cup_{n=1}^{\infty} E_n = [0,1]$ and

$$m^* E = m^*[0,1] = 1 < \infty = \sum_{n=1}^{\infty} m^* E_n.$$

Clearly the issue in this example is the fact that the sets E_n are not disjoint. If $A = [0,2] \cup [7,11]$, then our intuition tells us that if outer measure is going to be a reasonable notion of "length", then we should have $m^* A = 2 + 4 = 6$. In fact, we can prove directly that this is the case (we shall see a more general version of this in the Assignments).

From this, we might be tempted to conjecture that given a *disjoint* collection $\{E_n\}_{n=1}^{\infty}$ of subsets of $\mathbb{R}$, we have that

$$m^*(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m^* E_n.$$

As we shall now discover, this fails spectacularly. The failure of this equality leads us to the notion of *non-measurable* sets, which will be the subject of the next section of the notes.

2.13. Theorem. *There does not exist a translation-invariant outer measure μ on $\mathbb{R}$ satisfying the conditions:*

(a) $\mu(\mathbb{R}) > 0$*;*
(b) $\mu[0,1] < \infty$*; and*
(c) μ *is* σ***-additive****: that is, if $\{E_n\}_{n=1}^{\infty}$ is a countable collection of disjoint subsets of $\mathbb{R}$, then with $E := \cup_{n=1}^{\infty} E_n$,*

$$\mu E = \sum_{n=1}^{\infty} \mu E_n.$$

As a consequence, we see that Lebesgue outer measure m^* is ***not*** *σ-additive.*

Proof. We shall argue by contradiction. Suppose, to the contrary, that there exists such an outer measure μ.

Step 1. We consider the relation $\sim$ on $\mathbb{R}$ defined by $x \sim y$ if $y - x \in \mathbb{Q}$. We leave it as an exercise for the reader to show that $\sim$ is indeed an *equivalence relation*. For each $x \in \mathbb{R}$, denote the equivalence class of x under this relation by $[x]$. Clearly $[x] = \{x + q : q \in \mathbb{Q}\}$.

Let $\mathcal{F} := \{[x] : x \in \mathbb{R}\} = \mathbb{R}/\sim$ denote the set of all equivalence classes of elements in $\mathbb{R}$. Observe that

(i) for any $x \in \mathbb{R}$, $[x] = \mathbb{Q} + x$ is dense in $\mathbb{R}$ (since $\mathbb{Q}$ is dense in $\mathbb{R}$), and

(ii) given any $x, y \in \mathbb{R}$, $[x] = [y]$ if and only if $y - x \in \mathbb{Q}$.

For each $F := [x] = \mathbb{Q} + x \in \mathcal{F}$, we have that F is dense in $\mathbb{R}$. This allows us to choose a *unique* representative $x_F \in [x] \cap [0,1]$ of that equivalence class. Note that to do this simultaneously for all $F \in \mathcal{F}$ requires the Axiom of Choice!

Of course, $F = x_F + \mathbb{Q}$.

STEP 2. As is well-known, every equivalence relation on a set partitions that set into disjoint equivalence classes, and so

$$\mathbb{R} = \uplus\{F = [x_F] : F \in \mathcal{F}\}.$$

Let us define the set $\mathbb{V} := \{x_F : F \in \mathcal{F}\}$. Then

$$\mathbb{R} = \cup\{\mathbb{V} + q : q \in \mathbb{Q}\}.$$

We shall refer to $\mathbb{V}$ as **Vitali's set**, in honour of G. Vitali, who first "constructed" it. As we shall see, it is interesting enough to merit this special notation.

STEP 3. We claim that $p \neq q \in \mathbb{Q}$ implies that $\mathbb{V} + p \cap \mathbb{V} + q = \varnothing$, so that the sets $\{\mathbb{V} + q\}_{q \in \mathbb{Q}}$ are in fact pairwise disjoint. Furthermore, we show that $\mu\mathbb{V} > 0$. Indeed, suppose that $p \neq q \in \mathbb{Q}$, and that $z \in (\mathbb{V} + p) \cap (\mathbb{V} + q)$. Then there exist $F_1, F_2 \in \mathcal{F}$ such that

$$z = x_{F_1} + p = x_{F_2} + q.$$

It follows that $x_{F_2} - x_{F_1} = p - q \in \mathbb{Q}$, so that $[x_{F_1}] = [x_{F_2}]$. But $\mathbb{V}$ contained a *unique* representative from each equivalence class, and therefore $F_1 = F_2$, whence $x_{F_1} = x_{F_2}$ and so $p = q$, a contradiction.

Thus

$$\mathbb{R} = \uplus\{\mathbb{V} + q : q \in \mathbb{Q}\}.$$

Write $\mathbb{Q} = \{q_n\}_{n=1}^{\infty}$, which we may do as it is denumerable. Since μ is translation invariant and σ-additive,

$$0 < \mu(\mathbb{R}) = \mu(\uplus_{n=1}^{\infty}(\mathbb{V} + q_n)) = \sum_{n=1}^{\infty} \mu(\mathbb{V} + q_n) = \sum_{n=1}^{\infty} \mu\mathbb{V}.$$

Thus $\mu\mathbb{V} > 0$, which in turn implies that $\mu\mathbb{R} = \infty$.

STEP 4. The set $\mathcal{R} := \mathbb{Q} \cap [0,1]$ is denumerable, and so we may write $\mathcal{R} = \{r_n\}_{n=1}^{\infty}$. Note that $\mathbb{V} \subseteq [0,1]$ implies that $\mathbb{V} + r_n \subseteq [0,2]$ for all $n \geq 1$.

Observe that $[0,2] = [0,1] \cup ([0,1] + 1)$, and thus

$$\mu[0,2] \leq \mu[0,1] + \mu([0,1] + 1) = 2\mu[0,1] < \infty.$$

Finally, our hypothesis that μ is σ-additive and translation invariant implies that

$$\infty = \sum_{n=1}^{\infty} \mu\mathbb{V} = \sum_{n=1}^{\infty} \mu(\mathbb{V} + r_n) = \mu(\uplus_{n=1}^{\infty}(\mathbb{V} + r_n)) \leq \mu[0,2] < \infty,$$

a contradiction.

This contradiction proves that an outer measure μ on $\mathbb{R}$ satisfying all of the above conditions can not exist.

□

2.14. In light of the above Theorem, we are faced with a difficult choice. We may either

(a) content ourselves with the notion of Lebesgue outer measure m^* for *all* subsets E of $\mathbb{R}$, which agrees with our intuitive notion of length for intervals, but in so doing we must sacrifice the highly-desirable property of σ-additivity; or

(b) restrict the *domain* of our function m^* to a more "tractable" family of subsets of $\mathbb{R}$, where we might in fact be able to prove that the restriction of m^* to this family is σ-additive.

The standard approach is the second, and it is the one we shall adopt. The next section is devoted to describing our tractable family of sets, and to proving the σ-additivity of the restriction of m^* to this collection.

Appendix to Section 2.

2.15. For those with a historical bent (and there are exciting new treatments for that now), the "construction" of the set $\mathbb{V}$ defined in Step 2 of Theorem 2.13 is due to the Italian mathematician **Giuseppe Vitali** [**6**]. Strictly speaking, this isn't an explicit construction. The Axiom of Choice was invoked to prove the *existence* of such a set $\mathbb{V}$.

In the next Chapter, we shall give a name to the "tractable" family of sets described in Paragraph 2.14. Indeed, we shall refer to elements of this family as "*Lebesgue measurable sets*". The set $\mathbb{V}$ under discussion is an example of a *non-measurable set* – see Exercise 2. More generally, however, a **Vitali set** is a set $\mathcal{B} \subseteq [0,1]$ for which $x \ne y \in \mathcal{B}$ implies that that $x - y \in \mathbb{R} \setminus \mathbb{Q}$; i.e. $\mathcal{B}$ contains at most one representative from each coset of $\mathbb{Q}$ in $\mathbb{R}/\mathbb{Q}$.

As we have just remarked, Vitali's proof of the existence of a non-measurable subset of $\mathbb{R}$ relied on the Axiom of Choice. In fact, it was shown by Robert Solovay [**5**] that the Axiom of Choice is *required* to prove the existence of a non-measurable set, in the sense that the assertion that every subset of $\mathbb{R}$ is measurable is consistent with the Zermelo-Fraenkel axioms (ZF) of set theory. Rumour has it that his mother never spoke to him again after that.

There are other examples of non-measurable sets, including **Bernstein sets**, of which your humble author knows nothing. Those interested might wish to consult John C. Oxtoby's *Measure and Category* [**4**] for further details.

2.16. Theorem 2.13 raises an interesting question: is the issue with σ-additivity due to the fact that we are considering *infinitely* many disjoint sets? In other words, might it still be possible to find a translation-invariant outer measure μ on $\mathbb{R}$ satisfying

(a) $\mu(\mathbb{R}) > 0$;

(b) $\mu[0,1] < \infty$; and

(c) μ is **finitely-additive**: that is, given $\{E_1, E_2, \dots, E_N\}$ disjoint subsets of $\mathbb{R}$, then with $E := \cup_{n=1}^N E_n$,

$$\mu E = \sum_{n=1}^{N} \mu E_n?$$

Suppose that such a measure μ exists. Let $\{E_n\}_{n=1}^\infty$ be a countable family of pairwise disjoint sets. Since μ is an outer measure on $\mathbb{R}$, it is countably subadditive, and so

$$\mu(\cup_{n=1}^\infty E_n) \le \sum_{n=1}^{\infty} \mu E_n.$$

By monotonicity and finite-additivity of μ, we find that for all $N \ge 1$,

$$\sum_{n=1}^{N} \mu E_n = \mu(\cup_{n=1}^N E_n) \le \mu(\cup_{n=1}^\infty E_n) \le \sum_{n=1}^{\infty} \mu E_n.$$

Taking limits as N tends to infinity, we obtain:

$$\sum_{n=1}^{\infty} \mu E_n \leq \mu(\cup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \mu E_n,$$

or equivalently, that

$$\mu(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu E_n.$$

In other words, such an outer measure μ would be countably additive, contradicting Theorem 2.13.

Exercises for Section 2.

Exercise 2.1.

Let $E \subseteq \mathbb{R}$ be a finite set. Prove that $m^*E = 0$.

Exercise 2.2.

Prove the remaining cases from Proposition 2.7 and Corollary 2.8. That is, prove that if $a < b \in \mathbb{R}$, then

$$m^*[a,b) = m^*(a,b) = b - a,$$

while

$$m^*(-\infty, b] = m^*(a, \infty) = m^*[a, \infty) = m^*\mathbb{R} = \infty.$$

Exercise 2.3.

Prove that

$$m^*([0,2] \cup [7,11]) = 6.$$

Exercise 2.4.

Let $\sim$ be the relation on $\mathbb{R}$ defined by $x \sim y$ if $y - x \in \mathbb{Q}$. Prove that $\sim$ is an equivalence relation on $\mathbb{R}$.

Exercise 2.5.

Let m^* denote Lebesgue outer measure on $\mathbb{R}$, and let $\mathbb{V} \subseteq [0,1]$ denote Vitali's set from the proof of Theorem 2.13.

Prove that

$$m^*\mathbb{V} > 0.$$

3. Lebesgue measure

I can speak Esperanto like a native.

Spike Milligan

3.1. As mentioned at the end of the previous Chapter, our strategy will be to restrict the domain of Lebesgue outer measure to a smaller collection of sets, where Lebesgue outer measure *will* be σ-additive. We shall refer to this collection as the collection of *Lebesgue measurable sets*.

We begin with Carathéodory's definition of a Lebesgue measurable set, since it is the most practical definition to use. Later we shall see that Lebesgue measurable sets are "almost" countable intersections of open sets, in a way which we shall make precise.

3.2. Definition. *A set E is $\mathbb{R}$ is said to be* ***Lebesgue measurable*** *if, for all $X \subseteq \mathbb{R}$,*

$$m^*X = m^*(X \cap E) + m^*(X \setminus E).$$

We denote by $\mathfrak{M}(\mathbb{R})$ the collection of all Lebesgue measurable sets.

3.3. Remarks. Since our attention in this course is almost exclusively focused upon Lebesgue measure, we shall allow ourselves to drop the adjective "Lebesgue" and refer only to "measurable sets".

Informally speaking, we see that a set $E \subseteq \mathbb{R}$ is measurable provided that it is a "*universal slicer*", in the sense that it "slices" every other set X into two *disjoint* sets, namely $X \cap E$ and $X \setminus E$, where Lebesgue outer measure *is additive*!

We also note that the inequality

$$m^*X \le m^*(X \cap E) + m^*(X \setminus E)$$

is free from the σ-subadditivity of Lebesgue outer measure. In checking to see whether a given set is measurable or not, it therefore suffices to verify that the reverse inequality holds for all sets $X \subseteq \mathbb{R}$.

Before we proceed to the examples, which shall obtain a result which allows us to show that the set $\mathfrak{M}(\mathbb{R})$ of Lebesgue measurable sets itself has an interesting structure.

3.4. Definition. *Let Y be a non-empty set. A collection $\Omega \subseteq \mathfrak{P}(Y)$ is said to be an* ***algebra of subsets of*** *Y if*

(a) *$Y \in \Omega$;*
(b) *$E \in \Omega$ implies that $E^c := Y \setminus E \in \Omega$; and*
(c) *if $N \ge 1$ and $E_1, E_2, \ldots, E_N \in \Omega$, then*

$$E := \cup_{n=1}^N E_n \in \Omega.$$

*We say that Ω is a σ-**algebra of subsets** of Y if Ω is an algebra of sets which satisfies the additional property:*

(d) *if $F_n \in \Omega$ for all $n \geq 1$, then*

$$F := \cup_{n=1}^{\infty} F_n \in \Omega.$$

Informally, we often say that Ω is a σ-algebra.

3.5. Theorem. *The collection $\mathfrak{M}(\mathbb{R})$ of Lebesgue measurable sets in $\mathbb{R}$ is a σ-algebra of subsets of $\mathbb{R}$.*

Proof.

(a) Let us first verify that $\mathbb{R} \in \mathfrak{M}(\mathbb{R})$.

If $X \subseteq \mathbb{R}$, then $X \cap \mathbb{R} = X$, while $X \cap \mathbb{R}^c = X \cap \varnothing = \varnothing$. Thus

$$m^*(X \cap \mathbb{R}) + m^*(X \cap \mathbb{R}^c) = m^* X + m^* \varnothing = m^* X + 0 = m^* X.$$

By definition, $\mathbb{R} \in \mathfrak{M}(\mathbb{R})$.

(b) The fact that $\mathfrak{M}(\mathbb{R})$ is closed under complementation is clear, since the definition of a measurable set is symmetric in E and $\mathbb{R} \setminus E$.

(c) Let $\{E_n\}_{n=1}^{\infty} \subseteq \mathfrak{M}(\mathbb{R})$. We must show that $E := \cup_{n=1}^{\infty} E_n \in \mathfrak{M}(\mathbb{R})$.

STEP 1. For each $N \geq 1$, consider $H_N := \cup_{n=1}^{N} E_n$. We shall argue by induction that $H_N \in \mathfrak{M}(\mathbb{R})$ for all $N \geq 1$.

For $N = 1$, this is trivially true, as $H_1 = E_1 \in \mathfrak{M}(\mathbb{R})$ by hypothesis.

Next suppose that $1 \leq M \in \mathbb{N}$ and that $H_M \in \mathfrak{M}(\mathbb{R})$. Let $X \subseteq \mathbb{R}$ be arbitrary. The induction hypothesis says that

$$m^* X = m^*(X \cap H_M) + m^*(X \setminus H_M).$$

But $E_{M+1} \in \mathfrak{M}(\mathbb{R})$, and so

$$m^*(X \setminus H_M) = m^*((X \setminus H_M) \cap E_{M+1}) + m^*((X \setminus H_M) \setminus E_{M+1}).$$

By the subadditivity of m^*,

$$\begin{aligned} m^* X &= m^*(X \cap H_M) + m^*(X \setminus H_M) \\ &= m^*(X \cap H_M) + m^*((X \setminus H_M) \cap E_{M+1})) + m^*((X \setminus H_M) \setminus E_{M+1}) \\ &\geq m^*(X \cap (H_M \cup E_{M+1})) + m^*(X \setminus (H_M \cup E_{M+1})) \\ &= m^*(X \cap H_{M+1}) + m^*(X \setminus H_{M+1}) \end{aligned}$$

Since the reverse inequality holds for any outer measure, we see that

$$m^* X = m^*(X \cap H_{M+1}) + m^*(X \setminus H_{M+1}),$$

and thus $H_{M+1} \in \mathfrak{M}(\mathbb{R})$.

By induction, $H_N \in \mathfrak{M}(\mathbb{R})$ for all $N \geq 1$.

STEP 2. Next we shall write each H_N as a *disjoint* union of sets in $\mathfrak{M}(\mathbb{R})$.

Let $F_1 := H_1$, and for $n \geq 2$, set $F_n := H_n \setminus H_{n-1}$. Clearly $F_i \cap F_j = \varnothing$ for $1 \leq i \neq j$, and $H_N = \cup_{n=1}^{N} F_n$ for all $N \geq 1$.

Let $N \geq 2$ be an integer and note that $H_{N-1}, H_N \in \mathfrak{M}(\mathbb{R})$ implies that $H_{N-1}^c, H_N^c \in \mathfrak{M}(\mathbb{R})$ by (b). By Step 1, $(H_N^c \cup H_{N-1}) \in \mathfrak{M}$, and using (b) once again,

$$F_N = H_N \setminus H_{N-1} = (H_N^c \cup H_{N-1})^c \in \mathfrak{M}(\mathbb{R}).$$

Step 3. Now we claim that if $X \subseteq \mathbb{R}$, then for each $N \geq 1$,

$$m^*(X \cap (\cup_{n=1}^N F_n)) = \sum_{n=1}^N m^*(X \cap F_n).$$

The claim is trivially true when $N = 1$.

Suppose that $1 \leq M \in \mathbb{N}$ and that

$$m^*(X \cap (\cup_{n=1}^M F_n)) = \sum_{n=1}^M m^*(X \cap F_n).$$

By the measurability of F_{M+1} and the fact that all F_j's are disjoint,

$$\begin{aligned} m^*(X \cap (\cup_{n=1}^{M+1} F_n)) &= m^*((X \cap (\cup_{n=1}^{M+1} F_n)) \setminus F_{M+1}) + \\ &\qquad m^*((X \cap (\cup_{n=1}^{M+1} F_n)) \cap F_{M+1}) \\ &= m^*(X \cap (\cup_{n=1}^M F_n)) + m^*(X \cap F_{M+1}) \\ &= \sum_{n=1}^M m^*(X \cap F_n) + m^*(X \cap F_{M+1}) \\ &\qquad \text{by the induction hypothesis} \\ &= \sum_{n=1}^{M+1} m^*(X \cap F_n). \end{aligned}$$

This completes the induction step and therefore proves our claim.

Step 4. Finally, observe that $E = \cup_{n=1}^\infty E_n = \cup_{n=1}^\infty H_n = \cup_{n=1}^\infty F_n$. We shall use this to prove that $E \in \mathfrak{M}(\mathbb{R})$.

Let $X \subseteq \mathbb{R}$. For all $N \geq 1$, $H_N \in \mathfrak{M}(\mathbb{R})$ and so

$$\begin{aligned} m^* X &= m^*(X \cap H_N) + m^*(X \setminus H_N) \\ &= m^*(X \cap (\cup_{n=1}^N F_n)) + m^*(X \setminus H_N) \\ &\geq m^*(X \cap (\cup_{n=1}^N F_n)) + m^*(X \setminus E) && \text{as } (X \setminus E) \subseteq (X \setminus H_N) \\ &= \sum_{n=1}^N m^*(X \cap F_n) + m^*(X \setminus E) && \text{by Step 3.} \end{aligned}$$

Taking limits, we see that

$$m^* X \geq \sum_{n=1}^\infty m^*(X \cap F_n) + m^*(X \setminus E).$$

Keeping in mind that m^* is σ-subadditive, we note that

$$\begin{aligned} m^*(X \cap E) &= m^*(X \cap (\cup_{n=1}^{\infty} F_n)) \\ &= m^*(\cup_{n=1}^{\infty}(X \cap F_n)) \\ &\le \sum_{n=1}^{\infty} m^*(X \cap F_n). \end{aligned}$$

Combining these last two estimates, we conclude that

$$m^* X \ge m^*(X \cap E) + m^*(X \setminus E),$$

and therefore that $E \in \mathfrak{M}(\mathbb{R})$.

□

It's high time that we produce examples of Lebesgue measurable sets. Thanks to the previous Theorem, given a subset $\mathfrak{S} \subseteq \mathfrak{M}(\mathbb{R})$, the entire σ-algebra generated by $\mathfrak{S}$ (namely the smallest σ-algebra of subsets of $\mathbb{R}$ which contains $\mathfrak{S}$ – why should this exist?) is also contained in $\mathfrak{M}(\mathbb{R})$.

3.6. Proposition.

(a) *If $E \subseteq \mathbb{R}$ and $m^* E = 0$, then $E \in \mathfrak{M}(\mathbb{R})$.*
(b) *For all $b \in \mathbb{R}$, $E := (-\infty, b) \in \mathfrak{M}(\mathbb{R})$.*
(c) *Every open and every closed set is Lebesgue measurable.*

Proof.

(a) Let $E \subseteq \mathbb{R}$ be a set with $m^* E = 0$, and let $X \subseteq \mathbb{R}$. By monotonicity of outer measure,

$$m^*(X \cap E) \le m^* E = 0,$$

and

$$m^*(X \cap E^c) \le m^* X.$$

Thus

$$m^* X = 0 + m^* X \ge m^*(X \cap E) + m^*(X \setminus E).$$

As we have seen, this is the statement that $E \in \mathfrak{M}(\mathbb{R})$.

(b) Fix $b \in \mathbb{R}$ and set $E = (-\infty, b)$. Let $X \subseteq \mathbb{R}$ be arbitrary. We must show that

$$m^* X \ge m^*(X \cap E) + m^*(X \setminus E).$$

If $m^* X = \infty$, then there is nothing to prove, and so we assume that $m^* X < \infty$. Let $\varepsilon > 0$ and let $\{I_n\}_{n=1}^{\infty}$ be a cover of X by open intervals such that

$$\sum_{n=1}^{\infty} \ell(I_n) < m^* X + \varepsilon < \infty.$$

It follows that each interval I_n has finite length, and so we may write $I_n = (a_n, b_n)$, $n \ge 1$. (As always, there is no harm in assuming that each $I_n \ne \varnothing$, otherwise we simply remove I_n from the cover.)

Set $J_n = I_n \cap E = (a_n, b_n) \cap (-\infty, b)$. Clearly each J_n, $n \ge 1$ is an open interval, possibly empty.

Let $K_n = I_n \setminus E = (a_n, b_n) \cap [b, \infty)$. Then

$$K_n \in \{\varnothing, [b, b_n), (a_n, b_n)\},$$

depending upon the values of a_n and b_n. But then we can find $c_n < d_n$ in $\mathbb{R}$ such that

$$K_n \subseteq L_n := (c_n, d_n)$$

and $\ell(L_n) - m^* K_n < \frac{\varepsilon}{2^n}$, $n \ge 1$.

In particular, for each $n \ge 1$, $I_n \subseteq J_n \cup L_n$ and

$$(\ell(J_n) + \ell(L_n)) - \ell(I_n) < \frac{\varepsilon}{2^n}.$$

Now

$$\begin{aligned} m^* X &> \left(\sum_{n=1}^{\infty} \ell(I_n)\right) - \varepsilon \\ &> \left(\sum_{n=1}^{\infty} (\ell(J_n) + \ell(L_n) - \frac{\varepsilon}{2^n})\right) - \varepsilon \\ &= \sum_{n=1}^{\infty} \ell(J_n) + \sum_{n=1}^{\infty} \ell(L_n) - \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} - \varepsilon. \end{aligned}$$

But $X \cap E \subseteq \cup_{n=1}^{\infty} J_n$ and $X \setminus E \subseteq \cup_{n=1}^{\infty} L_n$, and so

$$m^* X > m^*(X \cap E) + m^*(X \setminus E) - 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary,

$$m^* X \ge m^*(X \cap E) + m^*(X \setminus E),$$

proving that $E \in \mathfrak{M}(\mathbb{R})$.

(c) Fix $b \in \mathbb{R}$. We have just seen that $(-\infty, b) \in \mathfrak{M}(\mathbb{R})$. Since $\mathfrak{M}(\mathbb{R})$ is an algebra of sets, $E^c = [b, \infty) \in \mathfrak{M}(\mathbb{R})$ as well. But then $E_n := [b + \frac{1}{n}, \infty) \in \mathfrak{M}(\mathbb{R})$ for all $n \ge 1$, and since the latter is a σ-algebra,

$$(b, \infty) = \cup_{n=1}^{\infty} E_n \in \mathfrak{M}(\mathbb{R}) \qquad \text{for all } b \in \mathbb{R}.$$

If $a < b$, then $(a, b) = (-\infty, b) \cap (a, \infty) \in \mathfrak{M}(\mathbb{R})$. Since $\mathbb{R} \in \mathfrak{M}(\mathbb{R})$ by Theorem 3.5, we see that we have shown that every open interval lies in $\mathfrak{M}(\mathbb{R})$.

We saw in the Assignments that every open set $G \in \mathfrak{G}$ is a countable (disjoint) union of open intervals. Since $\mathfrak{M}$ is a σ-algebra, this means that every open set $G \in \mathfrak{M}$. But $\mathfrak{M}$ is also closed under complementation, and so every closed set lies in $\mathfrak{M}$ as well.

□

Now that we know that $\mathfrak{M}(\mathbb{R}) \ne \varnothing$, the following definition makes sense.

3.7. Definition. *Let m^* denote Lebesgue outer measure on $\mathbb{R}$. We define* ***Lebesgue measure*** *m on $\mathbb{R}$ to be the restriction of m^* to $\mathfrak{M}(\mathbb{R})$. That is, Lebesgue measure is the function*

$$\begin{array}{rccl} m: & \mathfrak{M}(\mathbb{R}) & \to & [0,\infty] \\ & E & \mapsto & m^*E. \end{array}$$

Let us recall that our strategy was to try to restrict the domain of m^* sufficiently to allow the restriction of m^* to this smaller collection of sets to be σ-additive. The next result shows that $\mathfrak{M}(\mathbb{R})$ may serve as such a domain.

3.8. Theorem. *Lebesgue measure is σ-additive as a function on $\mathfrak{M}(\mathbb{R})$. That is, if $E_n \in \mathfrak{M}(\mathbb{R})$ for all $n \geq 1$ and if $E_i \cap E_j = \varnothing$ for all $1 \leq i \neq j < \infty$, then*

$$m(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} mE_n.$$

Proof. We leave the proof of this to the Assignments.

□

3.9. Corollary. *There exist non-measurable sets.*

Proof. Suppose otherwise. Then every subset of $\mathbb{R}$ is Lebesgue measurable, and thus $\mathfrak{M}(\mathbb{R}) = \mathfrak{P}(\mathbb{R})$ and $m = m^*$. But in Theorem 2.13, we saw that Lebesgue outer measure m^* is not σ-additive on $\mathfrak{P}(\mathbb{R})$. This contradicts Theorem 3.8.

Thus $\mathfrak{M}(\mathbb{R}) \neq \mathfrak{P}(\mathbb{R})$.

□

More interestingly, we have the following result, which we leave to the exercises. Recall Vitali's set $\mathbb{V}$ from Theorem 2.13.

3.10. Proposition. *Vitali's set $\mathbb{V}$ is not Lebesgue measurable.*

3.11. Definition. *The σ-algebra of sets generated by the collection $\mathfrak{G} = \{G \subseteq \mathbb{R} : G \text{ is open}\}$ of all open subsets of $\mathbb{R}$ is called the* ***σ-algebra of Borel subsets*** *of $\mathbb{R}$, and is denoted by*

$$\mathfrak{Bor}(\mathbb{R}).$$

(That $\mathfrak{Bor}(\mathbb{R})$ exists follows from Exercise 1 below.)

By Theorem 3.5 and Proposition 3.6 above, we see that

$$\mathfrak{Bor}(\mathbb{R}) \subseteq \mathfrak{M}(\mathbb{R}).$$

3.12. Remarks. Note that since $\mathfrak{Bor}(\mathbb{R})$ is a σ-algebra which contains all open subsets of $\mathbb{R}$, it also contains all closed subsets of $\mathbb{R}$. In fact, we could have *defined* $\mathfrak{Bor}(\mathbb{R})$ to be the σ-algebra of subsets of $\mathbb{R}$ generated by the collection $\mathfrak{F} := \{F \subseteq \mathbb{R} : F \text{ is closed}\}$ of closed subsets of $\mathbb{R}$, and concluded that it would have contained $\mathfrak{G}$.

Given a family $\mathcal{A} \subseteq \mathfrak{P}(\mathbb{R})$ with $\varnothing, \mathbb{R} \in \mathcal{A}$, set

$$\mathcal{A}_\sigma := \{\cup_{n=1}^\infty A_n : A_n \in \mathcal{A}, n \geq 1\}$$
$$\mathcal{A}_\delta := \{\cap_{n=1}^\infty A_n : A_n \in \mathcal{A}, n \geq 1\}.$$

We refer to elements of $\mathcal{A}_\sigma$ as **$\mathcal{A}$-sigma sets** and elements of $\mathcal{A}_\delta$ as **$\mathcal{A}$-delta sets**. Observe that $\mathfrak{G}_\delta$ and $\mathfrak{F}_\sigma$ are both subsets of $\mathrm{Bor}(\mathbb{R})$.

3.13. Admittedly, our definition of a Lebesgue measurable set is not the most intuitive definition, and Carathéodory's definition is quite different from Lebesgue's original definition, which we shall now investigate.

3.14. Theorem. *Let $E \subseteq \mathbb{R}$. The following statements are equivalent.*

(a) *E is Lebesgue measurable; i.e. $E \in \mathfrak{M}(\mathbb{R})$.*

(b) *For every $\varepsilon > 0$ there exists an open set $G \supseteq E$ such that*

$$m^*(G \setminus E) < \varepsilon.$$

(c) *There exists a $\mathfrak{G}_\delta$-set H such that $E \subseteq H$ and*

$$m^*(H \setminus E) = 0.$$

In other words, *up to a set of measure zero*, every Lebesgue measurable set is a $\mathfrak{G}_\delta$-set. As we shall see in the assignments, up to a set of measure zero, every Lebesgue measurable set is a $\mathfrak{F}_\sigma$-set as well.

Proof.

(a) implies (b).

CASE 1. Suppose that $mE < \infty$.

Let $\varepsilon > 0$ and choose a cover $\{I_n\}_{n=1}^\infty$ of E by open intervals such that

$$\sum_{n=1}^\infty \ell(I_n) < mE + \varepsilon.$$

Set $G = \cup_{n=1}^\infty I_n$ so that G is open and $E \subseteq G$. Then

$$mG \leq \sum_{n=1}^\infty mI_n = \sum_{n=1}^\infty \ell(I_n) < mE + \varepsilon.$$

Also,

$$G = (G \cap E) \cup (G \setminus E) = E \,\dot\cup\, (G \setminus E),$$

and so

$$mE + m(G \setminus E) = mG < mE + \varepsilon,$$

whence

$$m^*(G \setminus E) = m(G \setminus E) < \varepsilon.$$

Case 2. Suppose that $mE = \infty$.

Let $\varepsilon > 0$, and for each $n \geq 1$, let $E_n = E \cap [-n, n]$. Note that $E_n \in \mathfrak{M}(\mathbb{R})$, as the latter is a σ-algebra of sets. Moreover, $E_n \subseteq [-n, n]$ implies that $mE_n \leq m[-n, n] = 2n < \infty$. From Case 1 above, we can find G_n and open set such that $E_n \subseteq G_n$ and $m(G_n \setminus E_n) < \frac{\varepsilon}{2^n}$, $n \geq 1$. Let $G = \cup_{n=1}^{\infty} G_n$, so that G is open, and

$$E = \cup_{n=1}^{\infty} E_n \subseteq \cup_{n=1}^{\infty} G_n = G.$$

If $x \in G \setminus E$, then $x \notin E_n$ for all $n \geq 1$, and there exists $N \geq 1$ such that $x \in G_N$. That is, $x \in G_N \setminus E_N$. Thus $G \setminus E \subseteq \cup_{n=1}^{\infty}(G_n \setminus E_n)$ (in fact equality is easily seen to hold), and so

$$m(G \setminus E) \leq \sum_{n=1}^{\infty} m(G_n \setminus E_n) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.$$

(b) implies (c).

For each $n \geq 1$, choose $G_n \subseteq \mathbb{R}$ open so that $E \subseteq G_n$, and $m^*(G_n \setminus E) < \frac{1}{n}$. Set $H := \cap_{n=1}^{\infty} G_n$, so that $E \subseteq H$, and $H \in \mathfrak{G}_\delta$.

By monotonicity of outer measure, for each $n \geq 1$ we have

$$m^*(H \setminus E) \leq m^*(G_n \setminus E) < \frac{1}{n},$$

and so

$$m^*(H \setminus E) = 0.$$

(c) implies (a).

Suppose there exists $H \in \mathfrak{G}_\delta \subseteq \mathfrak{M}(\mathbb{R})$ such that $E \subseteq H$ and $m^*(H \setminus E) = 0$. By Proposition 3.6, $m^*(H \setminus E) = 0$ implies that $H \setminus E \in \mathfrak{M}(\mathbb{R})$ with $m(H \setminus E) = 0$. But $\mathfrak{M}(\mathbb{R})$ is a σ-algebra of sets, and thus it is an algebra, from which we deduce that

$$E = H \setminus (H \setminus E) \in \mathfrak{M}(\mathbb{R}).$$

□

3.15. Example. The Cantor middle thirds set. Recall from Corollary 2.5 that if $E \subseteq \mathbb{R}$ is countable, then $m^*E = 0$. By Proposition 3.6, it follows that $E \in \mathfrak{M}(\mathbb{R})$, and thus $mE = m^*E = 0$. In other words, every countable set is Lebesgue measurable with Lebesgue measure zero.

We shall now construct an *uncountable set* C – in fact one whose cardinality is $\mathfrak{c}$, the cardinality of the real line $\mathbb{R}$ – whose measure mC is equal to zero. Since $m\mathbb{R} = \infty$, we see that the Lebesgue measure of a set is not so much a reflection of its cardinality, as much as a question of *how* the points in the set are distributed. Having said that, when the set in question is countable, the above argument shows that it is always "thinly distributed", in this analogy.

The **Cantor set** is typically obtained as the intersection of a countable family of sets, each iteratively constructed from the previous as follows:

We set $C_0 = [0, 1]$, and for $n \geq 1$, we set $C_n = \frac{1}{3}C_{n-1} \cup (\frac{2}{3} + \frac{1}{3}C_{n-1})$. Thus

- $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$;
- $C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{6}{9}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$,
- $C_3 = [0, \frac{1}{27}] \cup [\frac{2}{27}, \frac{3}{27}] \cup [\frac{6}{27}, \frac{7}{27}] \cup [\frac{8}{27}, \frac{9}{27}] \cup [\frac{18}{27}, \frac{19}{27}] \cup [\frac{20}{27}, \frac{21}{27}] \cup [\frac{24}{27}, \frac{25}{27}] \cup [\frac{26}{27}, 1]$;
- $\vdots$

The figure below shows each of the sets C_n, $0 \le n \le 7$.

FIGURE 2. AN ILLUSTRATION FROM HTTP://MATHFORUM.ORG/MATHIMAGES/INDEX.PHP/CANTOR_SET.

The **Cantor middle thirds set** is defined as the intersection of all of these sets, i.e.

$$C := \cap_{n=0}^{\infty} C_n.$$

Alternatively, beginning with $C_0 = [0, 1]$, one can think of obtaining C_1 from C_0 by removing the (open) "middle third" interval $(\frac{1}{3}, \frac{2}{3})$, resulting in the two intervals which comprise C_1 above. To obtain C_2 from C_1, one removes the (open) "middle third" of each of the two intervals in C_1, and so on. This motivates the term *middle thirds* in the above nomenclature.

It should be clear from the construction above that

(a) $C_0 \supseteq C_1 \supseteq C_2 \supseteq C_3 \supseteq \cdots \supseteq C$. Furthermore, each set C_n is closed, $n \ge 0$ and $mC_0 = 1 < \infty$. From our work in the assignments,

$$m^*C = \lim_{n\to\infty} mC_n = \lim_{n\to\infty} \frac{2^n}{3^n} = 0.$$

(b) Being closed and bounded, C is compact − hence measurable with $mC = 0$. Also, C_0 is compact and the collection $\{C_n\}_{n=0}^{\infty}$ clearly has the Finite Intersection Property. Thus $C = \cap_{n=0}^{\infty} C_n \ne \varnothing$. (We shall in fact show that C is uncountable!)

Now let us approach things from a different angle. Given $x \in [0, 1]$, consider the **ternary expansion** of x, namely

$$x = 0.x_1\, x_2\, x_3\, x_4 \dots$$

where $x_k \in \{0, 1, 2\}$ for all $k \ge 1$. As with decimal expansions, the expression above is meant to express the fact that

$$x = \sum_{k=1}^{\infty} \frac{x_k}{3^k}.$$

Non-uniqueness of this expansion is a problem here, as it is with decimal expansions. For example,

$$\frac{1}{3} = 0.0222222\cdot = 0.1000000\cdots.$$

We leave it as an exercise for the reader to show that the expansion of $x \in [0,1)$ *is* unique *except when* there exists $N \geq 1$ such that

$$x = \frac{r}{3^N} \quad \text{for some} \quad 0 < r < 3^N, \text{ where } 3 \nmid r.$$

When this is the case, we have that

$$x = 0.x_1 x_2 x_3 x_4 \cdots x_N,$$

where $x_N \in \{1, 2\}$.

If $x_N = 2$, we shall use that expression.

If $x_N = 1$, then

$$\begin{aligned} x &= 0.x_1\, x_2\, x_3 \cdots x_{N-2}\, x_{N-1}\, 1000\cdots \\ &= 0.x_1\, x_2\, x_3 \cdots x_{N-2}\, 02222\cdots, \end{aligned}$$

and we shall agree to adopt the *second* expression.

Finally, we shall use the convention that

$$1 = 0.2222\cdots.$$

With this convention, over $x \in [0,1]$ has a *unique* ternary expansion.

Now

- $x \in C_1$ if and only if $x_1 \neq 1$;
- $x \in C_2$ if and only if $x \in C_1 \cap C_2$, i.e. if and only if $x_1 \neq 1 \neq x_2$.
- $x \in C_3$ if and only if $x \in C_2 \cap C_3$, i.e., if and only if $1 \notin \{x_1, x_2, x_3\}$.

More generally, for $N \geq 1$, $x \in C_N$ if and only $1 \notin \{x_1, x_2, \ldots, x_N\}$.

From this it follows that $x \in C$ if and only if $x_n \neq 1$, $n \geq 1$. In other words,

$$C = \{x = 0.x_1\, x_2\, x_3\, x_4 \ldots : x_n \in \{0, 2\} \text{ for all } n \geq 1\}.$$

As we shall see in the assignments, the map

$$\begin{array}{rccc} \varphi : & C & \to & [0,1] \\ & x = 0.x_1\, x_2\, x_3\, x_4 \ldots & \mapsto & y = \sum_{n=1}^{\infty} \frac{y_n}{2^n}, \end{array}$$

where $y_n = \frac{x_n}{2}$, $n \geq 1$ is a surjection from C onto $[0,1]$. (It is profitable to think of y as the *binary* expansion of an arbitrary element of $[0,1]$.)

But then the cardinality $|C|$ of C is greater than or equal to $\mathfrak{c} = |[0,1]|$, the cardinality of the continuum. Since $C \subseteq \mathbb{R}$, we also have that $|C| \leq \mathfrak{c}$, and so the Schröder-Bernstein Theorem (see Theorem 3.18 below) implies that

$$|C| = \mathfrak{c}.$$

Thus C is an uncountable, measurable set whose Lebesgue measure is nonetheless equal to 0.

Appendix to Section 3.

3.16. With regards to Theorem 3.14 above, first note that the definition of Lebesgue outer measure says that for all $H \subseteq \mathbb{R}$,

$$m^* H = \inf\{\sum_{n=1}^{\infty} \ell(I_n) : \{I_n\}_{n=1}^{\infty} \text{ is a cover of } H\}.$$

Let $\varepsilon > 0$, $H \subseteq \mathbb{R}$ and choose a cover $\{I_n\}_{n=1}^{\infty}$ of H. Set $G := \cup_{n=1}^{\infty} I_n$, so that G is an open set in $\mathbb{R}$. By the σ-subadditivity and monotonicity of m^*,

$$m^* H \le m^* G \le \sum_{n=1}^{\infty} m^*(I_n) = \sum_{n=1}^{\infty} \ell(I_n) < m^* H + \varepsilon.$$

It is important to realize that this *does not* say that

$$m^*(G \setminus H) < \varepsilon.$$

For a trivial counterexample, one might take $H = [0, \infty)$ and $G = \mathbb{R}$. Then $m^* H = m^* G$, and so $m^* H \le m^* G + \varepsilon$ for all $\varepsilon > 0$, and yet

$$m^*(G \setminus H) = m^*(-\infty, 0] = \infty.$$

A far more interesting example comes from Vitali's set $\mathbb{V}$. Recall that $\mathbb{V} \subseteq [0,1]$ is non-measurable, and we have seen (see Exercise 5) that it has positive, but *finite* Lebesgue outer measure. Indeed, by monotonocity of Lebesgue outer measure,

$$0 < m^* \mathbb{V} \le 1.$$

The above construction shows that for each $n \ge 1$, we can find $G_n \subseteq \mathbb{R}$ open such that $\mathbb{V} \subseteq G_n$ and

$$0 \le m^* G_n - m^* \mathbb{V} < \frac{1}{n}.$$

The non-measurability of $\mathbb{V}$, combined with Theorem 3.14, shows that there exists $\varepsilon_0 > 0$ such that if $G \subseteq \mathbb{R}$ is open and $\mathbb{V} \subseteq G$, then

$$m^*(G \setminus \mathbb{V}) \ge \varepsilon_0.$$

In particular,

$$m^*(G_n \setminus \mathbb{V}) \ge \varepsilon_0 \qquad \text{for all } n \ge 1.$$

Oh my. Wicked. Very, very wicked indeed.

3.17. Carathéodory's original definition of a Lebesgue measurable set E (see Definition 3.2) asks that for *every* $X \subseteq \mathbb{R}$, we must have

$$m^* X = m^*(X \cap E) + m^*(X \setminus E).$$

As it turns out (this will be an assignment question), if $A \subseteq \mathbb{R}$ is Lebesgue measurable with $mA < \infty$, and if $E \subseteq A$, then the following are equivalent:

(a) E is Lebesgue measurable.
(b) $m^* A = m^*(A \cap E) + m^*(A \setminus E)$.

In other words, instead asking that m^* be additive with respect to the decomposition $X = E \cup (X \setminus E)$ for *every* set X that contains E, it suffices to ask that this condition holds in the single, solitary case where $X = A$!

In particular, this shows that if E is bounded (i.e. there exists $M > 0$ such that $E \subseteq [-M, M]$), then E is Lebesgue measurable if and only if

$$m^*(E) + m^*([-M, M] \setminus E) = 2M.$$

Suddenly Carathéodory's definition of a Lebesgue measurable set doesn't seem so bad!

It might be worthwhile to remind the reader of the Schröder-Bernstein Theorem and of its proof.

3.18. Theorem. ***The Schröder-Bernstein Theorem.*** *Let A and B be sets. If $|A| \le |B|$ and $|B| \le |A|$, then $|A| = |B|$.*

Proof.

Step 1. If Z is any set and $\varphi : \mathcal{P}(Z) \to \mathcal{P}(Z)$ is increasing in the sense that $X \subseteq Y \subseteq Z$ implies that $\varphi(X) \subseteq \varphi(Y)$, then φ has a **fixed point**; that is, there exists $T \subseteq X$ such that $\varphi(T) = T$.

Indeed, let $T = \cup\{X \subseteq Z : X \subseteq \varphi(X)\}$. If $X \subseteq Z$ and $X \subseteq \varphi(X)$, then $X \subseteq T$ and so $\varphi(X) \subseteq \varphi(T)$. That is, $X \subseteq Z$ and $X \subseteq \varphi(X)$ implies $X \subseteq \varphi(T)$, and thus

$$T = \cup\{X \subseteq Z : X \subseteq \varphi(X)\} \subseteq \varphi(T).$$

But then $\varphi(T) \subseteq Z$ and $\varphi(T) \subseteq \varphi(\varphi(T))$, so that $\varphi(T)$ is one of the sets appearing in the definition of T - i.e. $\varphi(T) \subseteq T$.

Together, these imply that $\varphi(T) = T$. (We remark that it is entirely possible that $T = \varnothing$.)

Step 2. Given sets A, B as above and injections $\kappa : A \to B$ and $\lambda : B \to A$, define

$$\begin{array}{rccc} \varphi : & \mathcal{P}(A) & \to & \mathcal{P}(A) \\ & X & \mapsto & A \setminus \lambda[B \setminus \kappa(X)]. \end{array}$$

Suppose $X \subseteq Y \subseteq A$. Then $\kappa(X) \subseteq \kappa(Y)$.

Hence

$$\begin{aligned} B \setminus \kappa(X) &\supseteq B \setminus \kappa(Y), \text{ so} \\ \lambda(B \setminus \kappa(X)) &\supseteq \lambda(B \setminus \kappa(Y)), \text{ which implies} \\ A \setminus \lambda(B \setminus \kappa(X)) &\subseteq A \setminus \lambda(B \setminus \kappa(Y)), \text{ which in turn implies} \\ \varphi(X) &\subseteq \varphi(Y). \end{aligned}$$

Step 3. By Steps One and Two, there exists $T \subseteq A$ such that $T = \varphi(T) = A \setminus \lambda[B \setminus \kappa(T)]$.

Define

$$\begin{aligned} f: \quad A &\to B \\ a &\mapsto \begin{cases} \kappa(a) & \text{if } a \in T, \\ \lambda^{-1}(a) & \text{if } a \in A \setminus T. \end{cases} \end{aligned}$$

Observe that λ is a bijection between $B \setminus \kappa(T)$ and $A \setminus T$, and that κ is a bijection between T and $\kappa(T)$, so that f is a bijection between A and B.

$\square$

Exercises for Section 3.

Exercise 3.1.
Let $E \in \mathfrak{M}(\mathbb{R})$, so that E is a Lebesgue measurable set. Let $\kappa \in \mathbb{R}$, and set $E + \kappa := \{x + \kappa : x \in E\}$. Prove that $E + \kappa \in \mathfrak{M}(\mathbb{R})$, and that $mE = m(E + \kappa)$.

Thus Lebesgue measure m is translation-invariant.

Exercise 3.2.
Let $\mathbb{V} \subseteq [0, 1]$ be Vitali's set from Theorem 2.13. Prove that $\mathbb{V}$ is not Lebesgue measurable.

Exercise 3.3.
Let $\mathfrak{S} \subseteq \mathfrak{M}(\mathbb{R})$.

Prove that there exists a σ-algebra $\mathfrak{N} \subseteq \mathfrak{M}(\mathbb{R})$ of subsets which contains $\mathfrak{S}$ and with the property that if $\mathfrak{K}$ is *any* σ-algebra of measurable sets which contains $\mathfrak{S}$, then $\mathfrak{N} \subseteq \mathfrak{K}$. We say that $\mathfrak{N}$ is the **σ-algebra generated** by $\mathfrak{S}$.

Exercise 3.4.
Let $x \in [0, 1)$ and consider the ternary expansion of x given by

$$x = 0.x_1\, x_2\, x_3\, x_4 \ldots.$$

Show that the expansion *is* unique *except when* there exists $N \geq 1$ such that

$$x = \frac{r}{3^N} \quad \text{for some} \quad 0 < r < 3^N, \text{ where } 3 \nmid r.$$

Exercise 3.5.
Prove that the cardinality $|\mathfrak{M}(\mathbb{R})|$ of the collection of Lebesgue measurable sets is equal to that of the collection $\mathfrak{P}(\mathbb{R}) \setminus \mathfrak{M}(\mathbb{R})$ of non-measurable sets.

Exercise 3.6. Assignment
Prove that every open subset $G \subseteq \mathbb{R}$ is a countable union of disjoint, open intervals.

Exercise 3.7. Assignment
Prove that Lebesgue measure is σ-additive on $\mathfrak{M}(\mathbb{R})$.

Exercise 3.8. Assignment
Let $E \subseteq \mathbb{R}$. Prove that the following statements are equivalent.

(a) E is Lebesgue measurable.

(b) For every $\varepsilon > 0$ there exists a closed set $F \subseteq E$ such that

$$m^*(E \setminus F) < \varepsilon.$$

(c) There exists an $\mathfrak{F}_\sigma$-set H such that $H \subseteq E$ and

$$m^*(E \setminus H) = 0.$$

Exercise 3.9. Assignment Question.

Suppose that $\{E_n\}_{n=1}^{\infty}$ is an *increasing* sequence of Lebesgue measurable sets; i.e.

$$E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots.$$

Let $E = \cup_{n=1}^{\infty} E_n$, so that $E \in \mathcal{L}(\mathbb{R})$, as the latter is a σ-algebra. Prove that

$$mE = \lim_{n \to \infty} mE_n.$$

4. Lebesgue measurable functions

Don't ever wrestle with a pig. You'll both get dirty, but the pig will enjoy it.

Cale Yarborough

4.1. Let $H \subseteq \mathbb{R}$. We denote by $\mathfrak{M}(H)$ the collection of all Lebesgue measurable subsets of H. When $E \in \mathfrak{M}(\mathbb{R})$, it follows that

$$\mathfrak{M}(E) = \{F \cap E : F \in \mathfrak{M}(\mathbb{R})\}.$$

We leave it as an exercise for the reader to prove that $\mathfrak{M}(E)$ is a σ-algebra of subsets of E.

4.2. Definition. *Let (X, d) be a metric space and $E \in \mathfrak{M}(\mathbb{R})$. A function $f : E \to X$ is said to be* ***Lebesgue measurable*** *if*

$$f^{-1}(G) := \{x \in E : f(x) \in G\} \in \mathfrak{M}(E)$$

for every open set $G \subseteq X$.

We denote the set of Lebesgue measurable X-valued functions on E by $\mathcal{L}(E, X)$.

It is easily verified that this is equivalent to asking that $f^{-1}(F) \in \mathfrak{M}(E)$ for every closed subset F of X.

In the definition above, we have insisted that the domain of the function be a measurable set. Part of the reason for this is that (at the very least) we would want the constant functions to be measurable, and this happens if and only if the domain of our function is measurable.

As was the case with (Lebesgue) measurable sets, in light of the fact that we are dealing almost exclusively with Lebesgue measure in these notes, we drop the adjective "Lebesgue" and henceforth refer simply to *measurable functions.*

4.3. Proposition. *Let $E \subseteq \mathbb{R}$ be a measurable set. Then every continuous function $f : E \to X$ is measurable.*

Proof. To see this, note that the continuity of f implies that $f^{-1}(G)$ is a (relatively) open subset of E for all open sets $G \subseteq \mathbb{R}$. But a set $L \subseteq E$ is relatively open provided that $L = U \cap E$, where $U \subseteq \mathbb{R}$ is open. In our case, once we choose $U_0 \subseteq \mathbb{R}$ open such that $f^{-1}(G) = U_0 \cap E$, it follows from Theorem 3.5 and Proposition 3.6 that $U_0 \in \mathfrak{M}(\mathbb{R})$, whence $f^{-1}(G) \in \mathfrak{M}(E)$.

Thus f is measurable.

□

4.4. Example. Let $E \in \mathfrak{M}(\mathbb{R})$ and $H \subseteq E$. Consider the **characteristic function** or **indicator function** of H, namely

$$\begin{array}{rccc} \chi_H : & E & \to & \mathbb{R} \\ & x & \mapsto & \begin{cases} 0 & \text{if } x \in E \setminus H \\ 1 & \text{if } x \in H. \end{cases} \end{array}$$

If $G \subseteq \mathbb{R}$ is open, then

$$\chi_H^{-1}(G) = \begin{cases} \varnothing & \text{if } G \cap \{0,1\} = \varnothing \\ H & \text{if } G \cap \{0,1\} = \{1\} \\ E \setminus H & \text{if } G \cap \{0,1\} = \{0\} \\ E & \text{if } G \cap \{0,1\} = \{0,1\}. \end{cases}$$

It follows that χ_H *is Lebesgue measurable if and only if* H *is a Lebesgue measurable set.*

4.5. Proposition. *Let* $E \in \mathfrak{M}(\mathbb{R})$*, and suppose that* (X, d_X) *and* (Y, d_Y) *are metric spaces. Suppose furthermore that* $f : E \to X$ *is measurable and that* $g : X \to Y$ *is continuous. Then* $g \circ f : E \to Y$ *is measurable.*

Proof. Let $G \subseteq Y$ be open. Then $g^{-1}(G) \subseteq X$ is open, as g is continuous. But f is measurable, so

$$(g \circ f)^{-1}(G) = f^{-1}(g^{-1}(G)) \in \mathfrak{M}(E),$$

proving that $g \circ f$ is measurable.

□

4.6. Example. Let $E \subseteq \mathbb{R}$ be a measurable set and let $f : E \to \mathbb{K}$ be a measurable function. If $g : \mathbb{K} \to \mathbb{R}$ is the function $g(x) = |x|$, $x \in \mathbb{K}$, then g is continuous and so $g \circ f = |f|$ is measurable.

It is perhaps worth noting that the converse to this is false.

4.7. Proposition. *Let* $E \in \mathfrak{M}(\mathbb{R})$ *and* $f, g : E \to \mathbb{K}$ *be functions. The following statements are equivalent.*

(a) f *and* g *are measurable.*

(b) *The map*

$$\begin{array}{rccc} h : & E & \to & (\mathbb{K}^2, \|\cdot\|_2) \\ & x & \mapsto & (f(x), g(x)) \end{array}$$

is measurable.

Proof.

(a) implies (b).

Let $A, B \subseteq \mathbb{K}$ be open sets. Then $A \times B = \{(a, b) : a \in A, b \in B\}$ is open in $\mathbb{K}^2$, and furthermore, every open set $G \subseteq \mathbb{C}$ is a countable union of sets of this form (see the exercises).

Let $G \subseteq \mathbb{K}^2$ be open, and choose open sets A_n, B_n in $\mathbb{K}$ such that $G = \cup_{n=1}^{\infty} A_n \times B_n$. Then

$$h^{-1}(G) = \cup_{n=1}^{\infty} h^{-1}(A_n \times B_n) = \cup_{n=1}^{\infty} f^{-1}(A_n) \cap g^{-1}(B_n).$$

But each $f^{-1}(A_n) \cap g^{-1}(B_n)$ is measurable, being the intersection of measurable sets. Since $\mathfrak{M}(E)$ is a σ-algebra, $h^{-1}(G) \in \mathfrak{M}(E)$.

Thus h is measurable.

(b) implies (a).

Suppose next that h is measurable. The maps $\pi_i : \mathbb{K}^2 \to \mathbb{K}$, $i = 1, 2$ defined by $\pi_1(w, z) = w$ and $\pi_2(w, z) = z$ are continuous. By Proposition 4.5, $f = \pi_1 \circ h$ and $g = \pi_2 \circ h$ are measurable.

□

4.8. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and let $f, g : E \to \mathbb{K}$ be measurable. Then*

(a) *the constant functions are all measurable;*
(b) *$f + g$ is measurable;*
(c) *fg is measurable; and*
(d) *if for all $x \in E$ we have that $g(x) \neq 0$, then f/g is measurable.*

In particular, $\mathcal{L}(E, \mathbb{K})$ is an algebra.

Proof.

(a) Let $\kappa \in \mathbb{K}$ be fixed and suppose that $\varphi : E \to \mathbb{K}$ is the constant function $\varphi(x) = \kappa$, $x \in E$. If $G \subseteq \mathbb{K}$ is open, then either $\kappa \in G$, in which case $\varphi^{-1}(G) = E$ or $\kappa \notin G$, in which case $\varphi^{-1}(G) = \varnothing$. Either way, $\varphi^{-1}(G)$ is measurable, whence φ is measurable.

Next we shall show that $\mathcal{L}(E, \mathbb{K})$ is closed under sums, products and quotients. Let $h : E \to \mathbb{K}^2$ be the function $h(x) = (f(x), g(x))$. As we saw in Proposition 4.7, h is measurable.

(b) Let $\sigma : \mathbb{K}^2 \to \mathbb{K}$ be the function $\sigma(x, y) = x + y$. Then σ is continuous and so by Proposition 4.5,
$$f + g = \sigma \circ h$$
is measurable.

(c) Let $\mu : \mathbb{K}^2 \to \mathbb{K}$ be the function $\mu(x, y) = xy$. Then μ is continuous and so by Proposition 4.5,
$$fg = \mu \circ h$$
is measurable.

(d) Let $\delta : \mathbb{K} \times (\mathbb{K} \setminus \{0\}) \to \mathbb{K}$ be the function $\delta(x, y) = \dfrac{x}{y}$. Then δ is continuous and so by Proposition 4.5,
$$\frac{f}{g} = \delta \circ h$$
is measurable.

Thus the set $\mathcal{L}(E, \mathbb{K})$ forms a subalgebra of the algebra $\mathbb{K}^E$ of all functions from E into $\mathbb{K}$, and the constant function $\varphi(x) = 1$ for all $x \in E$ clearly serves as the identity of this algebra.

□

4.9. Remark. Note that $\mathbb{C}$ is a metric space where $d(w, z) := |w - z|$, $w, z \in \mathbb{C}$. Moreover, the map

$$\begin{array}{cccc} \gamma: & \mathbb{C} & \to & \mathbb{R}^2 \\ & x + iy & \mapsto & (x, y) \end{array}$$

is a homemorphism.

Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $f : E \to \mathbb{C}$ is a function.

- If f is measurable, then $\gamma \circ f = (\operatorname{Re} f, \operatorname{Im} f)$ is measurable by Proposition 4.5. By Proposition 4.7, each of the functions $g = \operatorname{Re} f$ and $h = \operatorname{Im} f$ is measurable.
- If $g_1 = \operatorname{Re} f$ and $g_2 = \operatorname{Im} f$ are both measurable, then by Proposition 4.7, so is $h := (g_1, g_2)$. Then $f = \gamma^{-1} \circ h$ is measurable by Proposition 4.5

In other words, *a complex-valued function is measurable if and only if its real and imaginary parts are measurable.* In light of this, we shall prove a number of results for real-valued measurable functions (allowing us to bypass some routine technical details), and leave it to the reader to formulate and prove the corresponding results for complex-valued functions. But first, we stop for a result which will prove useful in the second half of the notes. Recall that $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$.

4.10. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $f : E \to \mathbb{C}$ is measurable. There exists a measurable function $u : E \to \mathbb{T}$ such that*

$$f = u \cdot |f|.$$

Proof. Since $\{0\} \subseteq \mathbb{C}$ is closed and f is measurable, $K := f^{-1}(\{0\}) \in \mathfrak{M}(E)$. From Example 4.4, we have that χ_K is a measurable function, and therefore $f + \chi_K$ is measurable as well. Note also that $x \in E$ implies that $(f + \chi_K)(x) \neq 0$.

Set $u = \dfrac{f + \chi_K}{|f + \chi_K|}$. Then u is measurable by Proposition 4.8, and clearly $f = u \cdot |f|$.

□

At the moment, given a set $E \in \mathfrak{M}(\mathbb{R})$, in order to verify that a function $f : E \to \mathbb{R}$ is measurable, we must check that $f^{-1}(G)$ is measurable for all $G \subseteq \mathbb{R}$ open, or equivalently that $f^{-1}(F)$ is measurable for all $F \subseteq \mathbb{R}$ closed. Given how many different open (resp. closed) subsets of $\mathbb{R}$ there are – this threatens to be an Herculean, so as not to say a Sisyphean task. The following result makes the process much more manageable.

4.11. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and let $f : E \to \mathbb{R}$ be a function. The following statements are equivalent.*

(a) *f is measurable.*
(b) *$f^{-1}((a,\infty)) \in \mathfrak{M}(E)$ for all $a \in \mathbb{R}$.*
(c) *$f^{-1}((-\infty,b]) \in \mathfrak{M}(E)$ for all $b \in \mathbb{R}$.*
(d) *$f^{-1}((-\infty,b)) \in \mathfrak{M}(E)$ for all $b \in \mathbb{R}$.*
(e) *$f^{-1}([a,\infty)) \in \mathfrak{M}(E)$ for all $a \in \mathbb{R}$.*

Proof.

(a) implies (b). Since (a,∞) is open for all values of $a \in \mathbb{R}$, this is an immediate consequence of the definition of measurability.

(b) implies (c). Let $b \in \mathbb{R}$. By hypothesis, we have that $f^{-1}((b,\infty)) \in \mathfrak{M}(E)$. But then

$$f^{-1}((-\infty,b]) = E \setminus f^{-1}((b,\infty)) \in \mathfrak{M}(E),$$

since the latter is a σ-algebra of sets, by Exercise 1.

(c) implies (d). For each integer $n \geq 1$, $f^{-1}((-\infty, b - \frac{1}{n}]) \in \mathfrak{M}(E)$ by hypothesis, and thus

$$f^{-1}((-\infty,b)) = \cup_{n=1}^{\infty} f^{-1}((-\infty, b - \frac{1}{n}]) \in \mathfrak{M}(E).$$

(d) implies (e). The proof is similar to that of (b) implies (c), and is left as an exercise.

(e) implies (a). Observe first that if $a, b \in \mathbb{R}$, then

$$f^{-1}((a,\infty)) = \cup_{n=1}^{\infty} f^{-1}([a + \frac{1}{n}, \infty)) \in \mathfrak{M}(E),$$

while

$$f^{-1}((-\infty,b)) = E \setminus f^{-1}([b,\infty)) \in \mathfrak{M}(E).$$

If $a < b$, then

$$f^{-1}((a,b)) = f^{-1}((a,\infty)) \cap f^{-1}((-\infty,b)) \in \mathfrak{M}(E).$$

Thus $f^{-1}(I)$ is open for every open interval $I \subseteq \mathbb{R}$.

Finally, if $G \subseteq \mathbb{R}$ is open, then there exists a countable family of open intervals I_n such that $G = \cup_{n=1}^{\infty} I_n$. (As seen in Exercise 6, we can choose the I_n's to be disjoint, though this is not necessary here.) Since $\mathfrak{M}(E)$ is a σ-algebra,

$$f^{-1}(G) = \cup_{n=1}^{\infty} f^{-1}(I_n) \in \mathfrak{M}(E).$$

By definition, f is measurable.

□

4.12. Corollary. *Let $E \in \mathfrak{M}(\mathbb{R})$ and $f : E \to \mathbb{R}$ be a function. The following statements are equivalent.*

(a) *f is measurable.*
(b) *$f^{-1}(B) \in \mathfrak{M}(E)$ for all $B \in \mathfrak{Bor}(\mathbb{R})$.*

Proof. This is left to the Assignments.

□

4.13. Remark. Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $f : E \to \mathbb{R}$ is a function. We define

$$f^+(x) = \max(f(x), 0) \qquad x \in E$$
$$f^-(x) = \max(-f(x), 0) \qquad x \in E.$$

Note that $f = f^+ - f^-$ and that $|f| = f^+ + f^-$.

It follows from Examples 4.6 and Proposition 4.8 that if f is measurable, then so are

$$f^+ = \frac{|f| + f}{2} \qquad \text{and} \qquad f^- = \frac{|f| - f}{2}.$$

Combining this with with Remark 4.9, we see that every complex-valued measurable function is a linear combination of four non-negative, real-valued measurable functions.

4.14. We are now going to examine a number of results that deal with pointwise limits of sequences of measurable, real-valued functions. It will prove useful to include the case where the limit at a given point exists as an *extended* real number; that is, when the sequence diverges to ∞, or to $-\infty$. While this is useful for treating measure-theoretic and analytic properties of sequences of functions, there is a price to pay: the extended real numbers have poor algebraic properties. In particular, we can not add ∞ to $-\infty$, and so the class of functions we shall examine will not form a vector space.

4.15. Definition. *We define the* ***extended real numbers*** *to be the set*

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\},$$

also written $\overline{\mathbb{R}} = [-\infty, \infty]$.

By ***convention****, we shall define*

- $\alpha + \infty = \infty = \infty + \alpha$ *for all* $\alpha \in \mathbb{R} \cup \{\infty\}$*;*
- $\alpha + -\infty = -\infty = -\infty + \alpha$ *for all* $\alpha \in \mathbb{R} \cup \{-\infty\}$*;*
- $\alpha \cdot \infty = \infty \cdot \alpha = (-\infty) \cdot (-\alpha) = (-\alpha) \cdot (-\infty) = \infty$ *if* $0 < \alpha \in \mathbb{R}$ *or* $\alpha = \infty$*;*
- $\alpha \cdot \infty = \infty \cdot \alpha = (-\infty) \cdot (-\alpha) = (-\alpha) \cdot (-\infty) = -\infty$ *if* $\alpha < 0 \in \mathbb{R}$ *or* $\alpha = -\infty$*;*
- $0 = 0 \cdot \infty = \infty \cdot 0 = 0 \cdot (-\infty) = (-\infty) \cdot 0.$

Observe that we define neither $\infty - \infty$ nor $-\infty + \infty$.

4.16. Definition. *Given $H \subseteq \mathbb{R}$, we refer to a function $f : H \to \overline{\mathbb{R}}$ as an* ***extended real-valued function****.*

If $E \in \mathfrak{M}(\mathbb{R})$, an extended real-valued function $f : E \to \overline{\mathbb{R}}$ is said to be ***Lebesgue measurable*** *if*

(a) *$f^{-1}(G) \in \mathfrak{M}(E)$ for all open sets $G \subseteq \mathbb{R}$, and*
(b) *$f^{-1}(\{-\infty\})$, $f^{-1}(\{\infty\}) \in \mathfrak{M}(E)$.*

We denote the set of Lebesgue measurable extended real-valued functions on E by

$$\mathcal{L}(E, \overline{\mathbb{R}}) = \{f : E \to \overline{\mathbb{R}} : f \text{ is measurable}\}.$$

We shall often have occasion to refer to the non-negative elements of $\mathcal{L}(E, \overline{\mathbb{R}})$, and so we also define the notation

$$\mathcal{L}(E, [0, \infty]) = \{f \in \mathcal{L}(E, \overline{\mathbb{R}}) : 0 \le f(x) \text{ for all } x \in E\}.$$

Note. We remark that condition (a) above can be replaced with the condition that $f^{-1}(F) \in \mathfrak{M}(E)$ for all closed sets $F \subseteq \mathbb{R}$.

As was the case with real-valued measurable functions, in testing whether or not a given extended real-valued function is measurable or not, it suffices to check that the inverse images of certain intervals are measurable. In what follows, we write $(a, \infty]$ to mean $(a, \infty) \cup \{\infty\}$ and $[-\infty, b) = (-\infty, b) \cup \{-\infty\}$ for all $a, b \in \mathbb{R}$.

4.17. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $f : E \to \overline{\mathbb{R}}$ is a function. The following statements are equivalent.*

(a) *f if Lebesgue measurable.*
(b) *For all $a \in \mathbb{R}$, $f^{-1}((a, \infty]) \in \mathfrak{M}(E)$.*
(c) *For all $b \in \mathbb{R}$, $f^{-1}([-\infty, b)) \in \mathfrak{M}(E)$.*

Proof. The proof of this Proposition is left as an exercise for the reader.

□

4.18. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $(f_n)_{n=1}^\infty$ is a sequence of extended real-valued, measurable functions on E. The following (extended real-valued) functions are also measurable.*

(a) *$g_1 := \sup_{n \ge 1} f_n$;*
(b) *$g_2 := \inf_{n \ge 1} f_n$;*
(c) *$g_3 := \limsup_{n \ge 1} f_n$; and*
(d) *$g_4 := \liminf_{n \ge 1} f_n$.*

Proof.

(a) By Proposition 4.17 above, it suffices to prove that $g_1^{-1}((a, \infty]) \in \mathfrak{M}(E)$ for all $a \in \mathbb{R}$.

But

$$g_1^{-1}((a, \infty]) = \cup_{n=1}^\infty f_n^{-1}((a, \infty]).$$

Since each $f_n^{-1}((a, \infty]) \in \mathfrak{M}(E)$ (as each f_n is measurable), we have that $g_1^{-1}((a, \infty]) \in \mathfrak{M}(E)$.

Hence g_1 is measurable.

(b) The proof of (b) is similar to that of (a). For each $b \in \mathbb{R}$,

$$g_2^{-1}([-\infty, b)) = \cup_{n\geq 1} f_n^{-1}([-\infty, b)) \in \mathfrak{M}(E).$$

Thus g_2 is measurable.

(c) For each $N \geq 1$, $h_N := \sup_{n\geq N} f_n$ is measurable by (a) above. Clearly

$$h_1 \geq h_2 \geq h_3 \geq \cdots.$$

But then $\limsup_{n\geq 1} f_n = \lim_{n\to\infty} h_n = \inf_{n\geq 1} h_n$, and this is measurable by (b) above.

(d) The proof of this is similar to that of (c), and is left as an exercise.

□

The next result is an immediate corollary of the above Proposition.

4.19. Corollary. *Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $(f_n)_{n=1}^\infty$ is a sequence of real-valued, measurable functions on E such that $f(x) = \lim_{n\to\infty} f_n(x)$ exists as an extended real number for each $x \in E$.*

Then $f \in \mathcal{L}(E, \overline{\mathbb{R}})$; i.e. f is measurable.

4.20. Definition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and $\varphi : E \to \overline{\mathbb{R}}$ be a function. We say that φ is **simple** if $\operatorname{ran}\varphi$ is finite. Suppose that $\operatorname{ran}\varphi = \{\alpha_1 < \alpha_2 < \cdots < \alpha_N\}$, and set $E_n := \varphi^{-1}(\{\alpha_n\})$, $1 \leq n \leq N$. We shall say that*

$$\varphi = \sum_{n=1}^{N} \alpha_n \chi_{E_n}$$

*is the **standard form** of φ.*

4.21. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $\varphi : E \to \overline{\mathbb{R}}$ is a simple function with range $\operatorname{ran}\varphi = \{\alpha_1 < \alpha_2 < \cdots < \alpha_N\}$. The following statements are equivalent.*

(a) *φ is measurable.*

(b) *If $\varphi = \sum_{n=1}^N \alpha_n \chi_{E_n}$ is the standard form of φ, then $E_n \in \mathfrak{M}(E)$, $1 \leq n \leq N$.*

Proof.

(a) implies (b).

Suppose that φ is measurable. Let $1 \leq n \leq N$. If $\alpha_n \in \mathbb{R}$, then $\{\alpha_n\}$ is a closed set, and thus $E_n = \varphi^{-1}(\{\alpha_n\})$ is measurable. If $\alpha_1 = -\infty$, then $\varphi^{-1}(\{\alpha_1\}) \in \mathfrak{M}(E)$ by definition of measurability, and similarly if $\alpha_N = \infty$, then $\varphi^{-1}(\{\alpha_N\}) \in \mathfrak{M}(E)$ by definition of measurability.

(b) implies (a).

Suppose that $E_n \in \mathfrak{M}(E)$ for all $1 \leq n \leq N$. Then χ_{E_n} is measurable for each n, by Example 4.4. For any $a \in \mathbb{R}$,

$$\varphi^{-1}((a, \infty]) = \cup\{E_n : a < \alpha_n\}.$$

Thus $\varphi^{-1}((a, \infty])$ is a finite (possibly empty) union of measurable sets, and as such it is measurable.

By Proposition 4.18, φ is measurable.

□

4.22. Example. The standard form is not the only way of expressing a simple function as a linear combination of characteristic functions.

Consider $\varphi : \mathbb{R} \to \mathbb{R}$ defined as $\varphi = \chi_{\mathbb{Q}} + 9\chi_{[2,6]}$. Then $\operatorname{ran}\varphi = \{0, 1, 9, 10\}$.

Set

$$\begin{aligned} E_1 &= \varphi^{-1}(\{0\}) = \mathbb{R} \setminus (\mathbb{Q} \cup [2,6]) \\ E_2 &= \varphi^{-1}(\{1\}) = \mathbb{Q} \setminus [2,6] \\ E_3 &= \varphi^{-1}(\{9\}) = [2,6] \setminus \mathbb{Q} \\ E_4 &= \varphi^{-1}(\{10\}) = \mathbb{Q} \cap [2,6]. \end{aligned}$$

Then

$$\varphi = 0\chi_{E_1} + 1\chi_{E_2} + 9\chi_{E_3} + 10\chi_{E_4}$$

is the standard from of φ.

4.23. Definition. *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. A subset $C \subseteq \mathcal{V}$ is said to form a* ***(real) cone*** *if*

(a) $C \cap -C = \{0\}$, *where* $-C = \{-w : w \in C\}$ *and*
(b) $y, z \in C$ *and* $0 \le \kappa \in \mathbb{R}$ *imply that*

$$\kappa y + z \in C.$$

4.24. Example.

(a) Let $\mathcal{V} = \mathbb{R}^3$, and let $C = \{(x, y, z) \in \mathcal{V} : 0 \le x, y, z\}$. Then C is a cone.
(b) Let $\mathcal{V} = \mathbb{C}$ and let

$$C = \{w \in \mathbb{C} : w = re^{i\theta}, \frac{\pi}{6} \le \theta < \frac{2\pi}{6}, 0 \le r < \infty\}.$$

Then C is a cone. We mention in passing that C is not closed in $\mathbb{C}$.
(c) Let $\mathcal{V} = \mathcal{C}([0,1], \mathbb{C}) := \{f : [0,1] \to \mathbb{C} : f \text{ is continuous}\}$. Let

$$C := \{f \in \mathcal{C}([0,1], \mathbb{C}) : f(x) \ge 0 \text{ for all } x \in [0,1]\}.$$

Then C is a cone.

4.25. Remark. Let $E \in \mathfrak{M}(\mathbb{R})$ be a measurable set. We shall denote by

$$\operatorname{Simp}(E, \mathbb{R})$$

the set of all simple, real-valued, **measurable** functions on E. We leave it as an exercise for the reader to show that $\operatorname{Simp}(E, \mathbb{R})$ is an algebra, and thus a vector space over $\mathbb{R}$.

It will also be useful to adopt the following notation:

$$\operatorname{Simp}(E, [0, \infty)) := \{\varphi \in \operatorname{Simp}(E, \mathbb{R}) : 0 \le \varphi(x) \text{ for all } x \in E\}.$$

Observe that this is a real cone in $\operatorname{Simp}(E, \mathbb{R})$.

We denote by

$$\mathrm{S{\scriptsize IMP}}(E, \overline{\mathbb{R}})$$

the set of all simple, extended, **measurable** real-valued functions on E, and we set

$$\mathrm{S{\scriptsize IMP}}(E, [0, \infty]) := \{\varphi \in \mathrm{S{\scriptsize IMP}}(E, \overline{\mathbb{R}}) : 0 \le \varphi(x) \text{ for all } x \in E\}.$$

Alas, $\mathrm{S{\scriptsize IMP}}(E, \overline{\mathbb{R}})$ is not a vector space over $\mathbb{R}$, since if $\varphi \in \mathrm{S{\scriptsize IMP}}(E, \overline{\mathbb{R}})$ and $\varphi^{-1}(\{-\infty, \infty\}) \neq \varnothing$, then φ does not admit an additive inverse (recall that we have not defined $-\infty + \infty$ in $\overline{\mathbb{R}}$).

The next result will be the key to our definition of Lebesgue integrability of functions in the next section.

4.26. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and $f : E \to [0, \infty]$ be a measurable function. Then there exists an increasing sequence*

$$\varphi_1 \le \varphi_2 \le \varphi_3 \le \cdots \le f$$

of simple, real-valued functions φ_n such that

$$f(x) = \lim_{n \to \infty} \varphi_n(x) \text{ for all } x \in E.$$

Proof. The proof of this Proposition is left as an Assignment question.

□

Appendix to Section 4.

4.27. Let us now examine the extended real numbers from a somewhat different point of view. You may consider the discussion below "culture".

Definition. *Let (X, τ_X) be a topological space. A compact topological space (Y, τ_Y) is said to be a **compactification** of X if there exists a dense subset $Z \subseteq Y$ and a homeomorphism*

$$\rho : X \to Z.$$

In general, there are a great many compactifications of topological space. Some (including the **Stone-Čech compactification**) are more important than others. A full treatment of these, however, is beyond the scope of these notes. Instead, we refer the interested reader to the excellent monograph of Stephen Willard [**8**].

4.28. There are, nevertheless, two particular compactifications which are of interest to us here, and these are simple enough to briefly describe.

In fact, we have already seen one. Consider the following topology τ_2 on the set $\overline{\mathbb{R}}$ of extended real numbers. A subset $G \subseteq \overline{\mathbb{R}}$ belongs to τ_2 if an only if:

(a) $G \cap (-\infty, \infty)$ is open in $\mathbb{R}$ (with its usual topology inherited from the metric $d(x, y) = |x - y|$, $x, y \in \mathbb{R}$);
(b) $\infty \in G$ implies that there exists $a \in \mathbb{R}$ such that $(a, \infty] \subseteq G$; and
(c) $-\infty \in G$ implies that there exists $b \in \mathbb{R}$ such that $[-\infty, b) \subseteq G$.

That this is indeed a topology on $\overline{\mathbb{R}}$ is left as an exercise for the reader. It is worth observing that τ_2 is the topology on $\mathbb{R}$ generated by the sets $\{(a, \infty] : a \in \mathbb{R}\}$, in the sense that this collection forms a **subbase** for τ_2. A different subbase for τ_2 is the collection $\{[-\infty, b) : b \in \mathbb{R}\}$. Note that these are precisely the families of "test sets" which appear in Proposition 4.17.

The proof of the following result is left as yet another exercise for the reader. We emphasize that the topology on $[0, 1] \subseteq \mathbb{R}$ is the usual (relative) topology that it inherits as a subset of $\mathbb{R}$, itself equipped with the usual topology.

4.29. Theorem. $(\overline{\mathbb{R}}, \tau_2)$ *is homeomorphic to the interval* $[0, 1]$*, under a homeomorphism that sends* $\mathbb{R} \subseteq \overline{\mathbb{R}}$ *to the dense subset* $(0, 1) \subseteq [0, 1]$.

It follows that $(\overline{\mathbb{R}}, \tau_2)$ (or equivalently $[0, 1]$) is a compactification of $\mathbb{R}$. This particular compactification is often referred to as the **two-point compactification** of $\mathbb{R}$. (This motivates the subscript "2" which appears in τ_2.)

4.30. The second interesting and useful compactification of $\mathbb{R}$ we wish to consider is the so-called **one-point compactification** of $\mathbb{R}$, also known as the **Aleksandrov compactification** of $\mathbb{R}$.

Here, deferring to standard notation, we set $\alpha[\mathbb{R}] := \mathbb{R} \cup \{\infty\}$. We define a topology τ_1 on $\alpha[\mathbb{R}]$ as follows: a subset $G \subseteq \alpha[\mathbb{R}]$ belongs to τ_1 if and only if:

(a) $G \cap (-\infty, \infty)$ is open in $\mathbb{R}$ (with its usual topology inherited from the metric $d(x, y) = |x - y|$, $x, y \in \mathbb{R}$); and

(b) $\infty \in G$ implies that there exists $0 < a \in \mathbb{R}$ such that $(a, \infty) \cup (\infty, -a) \subseteq G$.

Condition (b) may be replaced by the condition that the neighbourhoods of ∞ must be of the form $\{\infty\} \cup (\mathbb{R} \setminus K)$, where $K \subseteq \mathbb{R}$ is compact (in the usual topology on $\mathbb{R}$).

Recall our notation: $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. The topology on $\mathbb{T}$ which we are considering below is the usual (relative) topology that it inherits as a subset of $\mathbb{C}$, equipped with its usual topology.

4.31. Theorem. $(\alpha[\mathbb{R}], \tau_1)$ *is homeomorphic to* $\mathbb{T}$*, under a homeomorphism that sends* $\mathbb{R} \subseteq \alpha[\mathbb{R}]$ *to the dense subset* $\mathbb{T} \setminus \{-1\}$ *of* $\mathbb{T}$*.*

It follows that $\alpha[\mathbb{R}]$ is a compactification of $\mathbb{R}$. Since it is homeomorphic to the familiar set $\mathbb{T}$, we often just think of the one-point compactification of $\mathbb{R}$ as $\mathbb{T}$ itself.

4.32. We have described these compactifications of $\mathbb{R}$ merely to place the results of this Chapter and of later Chapters that deal with Fourier Series in context. People have lived fruitful and productive lives without knowing diddly-squat about either of these compactifications. Well, if you can call that living.

4.33. Unlike the notions of continuity, of piecewise-continuity and even of Riemann integrability, measurability of functions behaves unbelievably well under pointwise limits of functions, as Proposition 4.18 shows. The fact that in many ways, Lebesgue integration respects pointwise limits is what makes it such a cogent and powerful tool. This is what we shall turn to next.

Exercises for Section 4.

Exercise 4.1.
Let $E \in \mathfrak{M}(\mathbb{R})$. Prove that $\mathfrak{M}(E)$ is a σ-algebra of sets.

Exercise 4.2.
Let (X, d) be a metric space.

(a) Let $E \in \mathfrak{M}(\mathbb{R})$ be a measurable set. Verify that a function $f : E \to X$ is measurable if and only if $f^{-1}(F) \in \mathfrak{M}(\mathbb{R})$ for every closed subset F of X.
(b) Let $H \subseteq \mathbb{R}$ be a set. Verify that a constant function $f : H \to X$ is measurable if and only if H is measurable.

Exercise 4.3.
Let $f : \mathbb{R} \to \mathbb{K}$ be a function and $g : \mathbb{K} \to \mathbb{R}$ be the absolute value function $g(x) = |x|$, $x \in \mathbb{K}$. Suppose that $g \circ f = |f|$ is measurable.
Give an example to show that f need not be measurable.

Exercise 4.4.

(a) Let $G \subseteq \mathbb{K}^2$ be an open set. Prove that there exists open sets $A_n, B_n \subseteq \mathbb{K}$, $n \geq 1$, such that
$$G = \cup_{n=1}^{\infty} A_n \times B_n.$$
(b) Suppose that $G \subseteq \mathbb{R}^2$ is a non-empty open set. Prove that there exist countably many rectangles $R_n = (a_n, b_n) \times (c_n, d_n) \subseteq \mathbb{R}^2$ such that $G = \cup_{n=1}^{\infty} R_n$.

Exercise 4.5.
Prove that the functions σ, μ and δ from Proposition 4.8 are all continuous.

Exercise 4.6.
Formulate and prove an analogue of Proposition 4.11 for complex-valued functions.

Exercise 4.7.
Complete the proof of (d) implies (e) in Proposition 4.11.

Exercise 4.8. Assignment
Prove Corollary 4.12.

Exercise 4.9.
Let $E \in \mathfrak{M}(\mathbb{R})$. Show that an extended real-valued function $f : E \to \overline{\mathbb{R}}$ is measurable if and only the following two conditions hold.

(a) $f^{-1}(F) \in \mathfrak{M}(E)$ for all closed sets $F \subseteq \mathbb{R}$, and
(b) $f^{-1}(\{-\infty\})$ and $f^{-1}(\{\infty\}) \in \mathfrak{M}(E)$.

Exercise 4.10.

Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $f : E \to \overline{\mathbb{R}}$ is a function. Prove that the following statements are equivalent.

(a) f if Lebesgue measurable.
(b) For all $\alpha \in \mathbb{R}$, $f^{-1}((\alpha, \infty]) \in \mathfrak{M}(E)$.
(c) For all $\beta \in \mathbb{R}$, $f^{-1}([-\infty, \beta)) \in \mathfrak{M}(E)$.

Exercise 4.11.

Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $(f_n)_{n=1}^{\infty}$ is a sequence of extended real-valued, measurable functions on E. Complete the proof of Proposition 4.18 by showing that

$$g := \liminf_{n \ge 1} f_n$$

is measurable.

Exercise 4.12.

Let $E \in \mathfrak{M}(\mathbb{R})$. Prove that $\textsc{Simp}(E, \mathbb{R})$ is an algebra over $\mathbb{R}$.

Exercise 4.13. Assignment

Let $E \in \mathfrak{M}(\mathbb{R})$ and $f : E \to [0, \infty]$ be a measurable function. Prove that there exists an increasing sequence

$$0 \le \varphi_1 \le \varphi_2 \le \varphi_3 \le \cdots \le f$$

of simple, real-valued functions φ_n such that

$$f(x) = \lim_{n \to \infty} \varphi_n(x) \text{ for all } x \in E.$$

Exercise 4.14.

Let E and F be measurable sets in $\mathbb{R}$ and suppose that $f : E \to \mathbb{R}$ and $g : F \to \mathbb{R}$ are functions.

(a) Define the function

$$\begin{array}{rccc} \widehat{f}: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & \begin{cases} f(x) & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}. \end{array}$$

Prove that f is measurable if and only if $\widehat{f}$ is measurable.

(b) Suppose that $E \cap F = \varnothing$. Prove that the function $h : E \cup F \to \mathbb{R}$ defined by

$$h(x) = \begin{cases} f(x) & \text{if } x \in E \\ g(x) & \text{if } x \in F \end{cases}$$

is measurable if and only if both f and g are measurable.

(c) Does the conclusion from (b) hold if $E \cap F \neq \varnothing$? Either prove that it does, or find a counterexample to show that it doesn't.

Exercise 4.15. Assignment Question.

(a) Let $f : \mathbb{R} \to [0, \infty]$ be a measurable function. Show that there exists an increasing sequence of measurable, simple functions $\varphi_n : \mathbb{R} \to [0, \infty)$ so that

$$f(x) = \lim_{n\to\infty} \varphi_n(x) \qquad \text{for all } x \in \mathbb{R}.$$

(b) Let $E \in \mathfrak{M}(\mathbb{R})$ and let $g : E \to [0, \infty]$ be a measurable function. Show that there exists an increasing sequence of measurable, simple functions $\psi_n : E \to [0, \infty)$ so that

$$g(x) = \lim_{n\to\infty} \psi_n(x) \qquad \text{for all } x \in E.$$

Hint for (a): For each $n \geq 1$, partition the interval $[0, n)$ into $n2^n$ equal subintervals $E_{k,n} = [\frac{k}{2^n}, \frac{k+1}{2^n})$, $0 \leq k < (n2^n) - 1$. Set $E_{n2^n,n} = [n, \infty]$. Use the sets $f^{-1}(E_{k,n})$, $0 \leq k \leq n2^n$ to build φ_n.

Hint for (b): This should be very short, otherwise you are doing something wrong.

Exercise 4.16.

Let $E \subseteq \mathbb{R}$ be a set of measure zero, and let $f : E \to \overline{\mathbb{R}}$ be any function whatsoever. Prove that f is measurable.

5. Lebesgue integration

I know that there are people who do not love their fellow man, and I hate people like that.

Tom Lehrer

5.1. Our approach to defining the Lebesgue integral of a measurable function f on a measurable set $E \subseteq \mathbb{R}$ will be an iterative one. We shall begin this by defining the integral of a simple, non-negative, extended real-valued function. We will then use this definition to define the integral of f when $f : E \to [0, \infty]$ is measurable, and derive a number of consequences of our definition.

Following this, we shall design our notion of Lebesgue integration so that it is linear; this will require us to impose certain conditions on the range of the functions involved.

5.2. Definition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and $\varphi \in \textsc{Simp}(E, [0, \infty])$. Let*

$$\varphi = \sum_{n=1}^{N} \alpha_n \chi_{E_n}$$

denote the standard form of φ. (Since φ is measurable, so is E_n, $1 \le n \le N$.)

We define

$$\int_E \varphi := \sum_{n=1}^{N} \alpha_n m E_n,$$

and observe that $\int_E \varphi \in [0, \infty]$.

If $F \subseteq E$ is measurable, we define

$$\int_F \varphi = \int_E \varphi \cdot \chi_F = \sum_{n=1}^{N} \alpha_n m(F \cap E_n).$$

We remind the reader that by convention, we have defined $0 \cdot \infty = 0$. Thus if $n = 1$ and $\alpha_n = 0$ and $mE_n = \infty$, or conversely if $n = N$ and $\alpha_n = \infty$ and $mE_n = 0$, then $\alpha_n m E_n = 0$, whence $\int_{E_n} \alpha_n \chi_{E_n} = 0$.

5.3. Example.

(a) Let $\varphi = 0\chi_{[4,\infty)} + 17\chi_{\mathbb{Q}\cap[0,4)} + 29\chi_{[2,4)\setminus\mathbb{Q}}$. Then

$$\begin{aligned}\int_{[0,\infty)} \varphi &= 0\, m[4, \infty) + 17\, m(\mathbb{Q} \cap [0, 4)) + 29\, m([2, 4) \setminus \mathbb{Q})\\ &= 0 \cdot \infty + 17 \cdot 0 + 29 \cdot 2\\ &= 58.\end{aligned}$$

(b) Let $C \subseteq [0,1]$ be the Cantor set from Example 3.15, and consider $\varphi = 1\,\chi_C + 2\,\chi_{[5,9]}$. Then

$$\begin{aligned}\int_{[0,6]} \varphi &= 1\,m(C \cap [0,6]) + 2\,m([5,9] \cap [0,6]) \\ &= 1 \cdot 0 + 2 \cdot m([5,6]) \\ &= 2.\end{aligned}$$

Our definition of the integral of a simple, nonnegative measurable function currently requires us to express the function in standard form. Let us now relax this condition.

5.4. Definition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and let $\varphi : E \to \overline{\mathbb{R}}$ be a simple, measurable function. Suppose that*

$$\varphi = \sum_{n=1}^{N} \alpha_n\,\chi_{H_n},$$

where $H_n \subseteq E$ is measurable and $\alpha_n \in \overline{\mathbb{R}}$, $1 \le n \le N$. Observe that we are not *requiring that the α_n's be distinct, nor that they be written in any particular order, nor that $E = \cup_{n=1}^{N} H_n$.*

We shall say that the above decomposition of φ is a ***disjoint representation*** *of φ if*

$$H_i \cap H_j = \varnothing, \quad 1 \le i \ne j \le N.$$

We emphasize that the measurability of the sets H_n, $1 \le n \le N$ is part of the definition of a disjoint representation of φ.

5.5. Lemma. *Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $\varphi, \psi : E \to \overline{\mathbb{R}}$ are simple, real-valued, measurable functions. Then there exist*

(i) $N \in \mathbb{N}$,
(ii) $\alpha_1, \alpha_2, \ldots, \alpha_N, \beta_1, \beta_2, \ldots, \beta_N \in \overline{\mathbb{R}}$, *and*
(iii) $H_1, H_2, \ldots, H_N \in \mathfrak{M}(E)$

such that $H_i \cap H_j = \varnothing$ if $1 \le i \ne j \le N$,

$$\varphi = \sum_{n=1}^{N} \alpha_n\,\chi_{H_n} \qquad \textit{and} \qquad \psi = \sum_{n=1}^{N} \beta_n\,\chi_{H_n}.$$

Remark. The key things to notice here are that the H_n's appearing in the decompositions of φ and ψ are the *same*, and the representations are *disjoint*.

Proof. Let $\varphi = \sum_{j=1}^{M_1} a_j \chi_{E_j}$ and $\psi = \sum_{k=1}^{M_2} b_k \chi_{F_k}$, where E_j, F_k are measurable subsets of E for all $1 \le j \le M_1$, $1 \le k \le M_2$, the E_j's are pairwise disjoint, and the F_k's are pairwise disjoint. (That such a decomposition exists is clear, as we may simply write φ and ψ in standard form.)

Then $\{E_j \cap F_k : 1 \le j \le M_1, 1 \le k \le M_2\}$ are disjoint, measurable sets and

$$\varphi = \sum_{j=1}^{M_1} a_j \left(\sum_{k=1}^{M_2} \chi_{E_j \cap F_k} \right) = \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} \ a_j \ \chi_{E_j \cap F_k},$$

and similarly

$$\psi = \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} \ b_k \ \chi_{E_j \cap F_k}.$$

Relabel $\{E_j \cap F_k : 1 \le j \le M_1, 1 \le k \le M_2\}$ as $\{H_n : 1 \le n \le N\}$ to complete the proof. (The α_n's and β_n's are clearly just relabelings of the a_j's and the b_k's respectively.)

□

5.6. Lemma. *Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $\varphi \in$ Simp$(E, [0, \infty])$. If $\varphi = \sum_{n=1}^N \alpha_n \chi_{H_n}$ is any disjoint representation of φ, then*

$$\int_E \varphi = \sum_{n=1}^N \alpha_n m H_n.$$

Proof. If $\cup_{n=1}^N H_n \ne E$, then we can set $H_{N+1} = E \setminus (\cup_{n=1}^N H_n)$ and $\alpha_{N+1} = 0$. As such, we assume without loss of generality that $E = \cup_{n=1}^N H_n$.

Observe that since the H_n's are mutually disjoint,

$$\operatorname{ran} \varphi \subseteq \{\alpha_n\}_{n=1}^N.$$

Of course, we might have $\alpha_i = \alpha_j$ with $1 \le i \ne j \le N$, so allow us to write

$$\operatorname{ran} \varphi = \{\beta_1 < \beta_2 < \cdots < \beta_M\},$$

for some $1 \le M \le N$, and set $E_m = \varphi^{-1}(\{\beta_m\})$, $1 \le m \le M$, so that each $E_m \in \mathfrak{M}(E)$.

Now, for $1 \le m \le M$, we have that $E_m = \uplus\{H_n : \alpha_n = \beta_m\}$. By definition,

$$\int_E \varphi = \int_E \sum_{k=1}^M \beta_k \chi_{E_k} = \sum_{k=1}^M \beta_k \, m E_k$$

Finally, $m(E_k) = \sum\{mH_n : \alpha_n = \beta_k\}$, and thus

$$\int_E \varphi = \sum_{k=1}^M \beta_k (\sum_{\alpha_j = \beta_k} m H_j) = \sum_{k=1}^M (\sum_{\alpha_j = \beta_k} \alpha_j m H_j) = \sum_{n=1}^N \alpha_n m H_n.$$

□

5.7. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$. If $\varphi, \psi \in$ Simp$(E, [0, \infty])$ and $0 \le \kappa \in \mathbb{R}$, then*

(a) $\int_E \kappa\varphi + \psi = \kappa \int_E \varphi + \int_E \psi$.
(b) *If $\varphi \le \psi$, then $\int_E \varphi \le \int_E \psi$.*

Proof. By Lemma 5.5, we can find common disjoint representations of φ and ψ, say

$$\varphi = \sum_{n=1}^N \alpha_n \chi_{H_n} \qquad \text{and} \qquad \psi = \sum_{n=1}^N \beta_n \chi_{H_n}.$$

(a) Then $\kappa\varphi + \psi = \sum_{n=1}^N (\kappa\alpha_n + \beta_n)\chi_{H_n}$ is a disjoint representation of $\kappa\varphi + \psi$, and so by Lemma 5.6,

$$\begin{aligned}\int_E (\kappa\varphi + \beta) &= \sum_{n=1}^N (\kappa\alpha_n + \beta_n) mH_n \\ &= \kappa(\sum_{n=1}^N \alpha_n mH_n) + \sum_{n=1}^N \beta_n\, mH_n \\ &= \kappa \int_E \varphi + \int_E \psi.\end{aligned}$$

(b) Suppose that $\varphi \le \psi$. Then $\alpha_n \le \beta_n$ for all $1 \le n \le N$, and so

$$\int_E \varphi = \sum_{n=1}^N \alpha_n\, mH_n \le \sum_{n=1}^N \beta_n\, mH_n = \int_E \psi.$$

□

Does part (a) of the above result hold if we consider $\kappa = \infty$?

5.8. Definition. *Recall that for $E \in \mathfrak{M}(\mathbb{R})$, we defined*

$$\mathcal{L}(E, [0,\infty]) = \{f : E \to [0,\infty] : f \text{ is measurable}\}.$$

*For $f \in \mathcal{L}(E, [0,\infty])$, we define the **Lebesgue integral** of f to be*

$$\int_E^{\text{NEW}} f = \sup\{\int_E \varphi : \varphi \in \textsc{Simp}(E, [0,\infty)), 0 \le \varphi \le f\}.$$

5.9. Remarks.

(a) We leave it as an exercise for the reader to show that the above definition is equivalent to defining

$$\int_E^{\text{NEW}} f = \sup\{\int_E \varphi : \varphi \in \textsc{Simp}(E, [0,\infty]),\ 0 \le \varphi \le f\}.$$

(The difference being that we now allow the simple functions to be *extended* real-valued and non-negative, instead of just real-valued and non-negative.)

(b) The reason for putting the superscript "NEW" in the above integral is the following. Observe that if $\varphi \in \textsc{Simp}(E, [0,\infty])$, we now have *two* definitions for the integral of φ. That is, writing $\varphi = \sum_{n=1}^N \alpha_n \chi_{H_n}$ in standard form, we have our original definition (Definition 5.2)

$$\int_E \varphi = \sum_{n=1}^N \alpha_n\, mH_n,$$

while from Definition 5.8, our new definition of the integral of φ becomes

$$\int_E^{\text{NEW}} \varphi = \sup\{\int_E \psi : \psi \in \textsc{Simp}(E, [0,\infty)), 0 \le \psi \le \varphi\}.$$

(c) It is entirely possible that $\int_E^{\text{NEW}} f = \infty$. For example, if $\varphi = \infty \cdot \chi_{[0,1]}$, then $\int_{[0,1]}^{\text{NEW}} \varphi = \infty$.

Alternatively, if $f(x) = x$, $x \in [0, \infty)$, then $\int_{[0,\infty)}^{\text{NEW}} f = \infty$. The proof of this is left as an exercise for the reader.

5.10. Let us reconcile these two definitions. Once this is done, we will no longer need to distinguish between the original and the new integral for non-negative, simple, measurable functions, and so we shall drop the superscript "NEW" for the integrals of non-negative, measurable extended-real valued functions altogether.

On the one hand, note that $\varphi \in \{\psi \in \text{SIMP}(E, [0, \infty)), 0 \leq \psi \leq \varphi\}$, and so by definition of $\int_E^{\text{NEW}} \varphi$, we have that

$$\int_E \varphi \leq \int_E^{\text{NEW}} \varphi.$$

On the other hand, if $\psi \in \text{SIMP}(E, [0, \infty))$ and $0 \leq \psi \leq \varphi$, then by Proposition 5.7 (b),

$$\int_E \psi \leq \int_E \varphi,$$

and thus

$$\int_E^{\text{NEW}} \varphi = \sup\{\int_E \psi : \psi \in \text{SIMP}(E, [0, \infty)), 0 \leq \psi \leq \varphi\} \leq \int_E \varphi.$$

This proves that

$$\int_E^{\text{NEW}} \varphi = \int_E \varphi$$

whenever φ is a non-negative, extended real-valued simple function.

Now that we have this, we shall drop the superscript NEW *and simply write*

$$\int_E f$$

for the Lebesgue integral of an element $f \in \mathcal{L}(E, [0, \infty])$.

5.11. Remark. We shall see in the Assignments that even when f is a relatively innocuous-looking function (for example $f(x) = x$ on $[0, 1]$), calculating the Lebesgue integral of f directly from the definition is an arduous task. Fortunately, Theorem 5.24 below will provide us with an alternate means of calculating the integrals of a large family of (Riemann integrable) functions, by showing that in many cases, the Lebesgue integral coincides with the Riemann integral. Of course, when the function is sufficiently nice, we may apply the Fundamental Theorem of Calculus to calculate the latter.

Sets of measure zero will play a central role in the theory that follows. The reason for this lies partly in the fact that the Lebesgue integral "ignores" these sets, in a sense which we shall now make precise.

5.12. Definition. *Let $E \in \mathfrak{M}(\mathbb{R})$. We say that a property (P) holds* ***almost everywhere*** *(**a.e.**) on E if the set*

$$B := \{x \in E : (P) \text{ does not hold}\}$$

has Lebesgue measure zero.

5.13. Example. Let $E \in \mathfrak{M}(\mathbb{R})$. Given $f, g \in \mathcal{L}(E, \overline{\mathbb{R}})$, we say that $f = g$ almost everywhere if

$$B := \{x \in E : f(x) \neq g(x)\}$$

has measure zero.

More specifically, therefore, $\chi_{\mathbb{Q}} = 0 = \chi_C$ a.e. on $\mathbb{R}$, where C is the Cantor set from Example 3.15.

5.14. Lemma. *Let $E \in \mathfrak{M}(\mathbb{R})$ and let f, g and $h : E \to [0, \infty]$ be functions. Suppose that g and h are measurable.*

(a) *Suppose furthermore that $E = X \cup Y$, where X and Y are measurable. Set $f_1 := f|_X$ and $f_2 := f|_Y$. Then f is measurable if and only if both f_1 and f_2 are measurable. When such is the case,*

$$\int_E f = \int_X f_1 + \int_Y f_2.$$

(b) *If $g \le h$, then $\int_E g \le \int_E h$.*

(c) *If $H \subseteq E$ is a measurable set, then*

$$\int_H g = \int_E g \cdot \chi_H \le \int_E g.$$

Proof. The proof of this lemma is left as a worthwhile exercise for the reader.

□

5.15. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$, and $f, g \in \mathcal{L}(E, [0, \infty])$.*

(a) *If $mE = 0$, then $\int_E f = 0$.*

(b) *If $f = g$ a.e. on E, then $\int_E f = \int_E g$.*

Proof.

(a) Let $\varphi \in \mathrm{SIMP}(E, [0, \infty))$ with $\varphi \le f$, and let $\varphi = \sum_{n=1}^N \alpha_n \chi_{H_n}$ denote the standard representation of φ, where $H_n \in \mathfrak{M}(E)$, $1 \le n \le N$.

Then

$$\begin{aligned} \int_E \varphi &= \sum_{n=1}^N \alpha_n \, mH_n \\ &\le \sum_{n=1}^N \alpha_n \, mE \\ &= \sum_{n=1}^N \alpha_n \, 0 \\ &= 0. \end{aligned}$$

Thus

$$\int_E f = \sup\{\int_E \varphi : \varphi \in \textsc{Simp}(E, [0, \infty)) : \varphi \le f\} = 0.$$

(b) Let $B = \{x \in E : f(x) \neq g(x)\}$, so that $mB = 0$. Then, using Lemma 5.14(a) as well as part (a) above, we find that

$$\begin{aligned}\int_E f &= \int_{E\setminus B} f + \int_B f \\ &= \int_{E\setminus B} g + 0 \\ &= \int_{E\setminus B} g + \int_B g \\ &= \int_E g.\end{aligned}$$

□

We now come to one of the major results in this course. In dealing with properties that hold almost everywhere on a measurable set $E \in \mathfrak{M}(\mathbb{R})$ (as will be the case in the Monotone Convergence Theorem below), we often have recourse to the following line of argument: we isolate the "bad" set K of measure zero where the property under consideration fails to hold, and the deal with the "good" set $E \setminus K$, where the property holds everywhere. Using Proposition 5.15 and Lemma 5.14, we can often "glue" these results together. This is a lot of quotation marks, which are the written equivalent of randomly flailing arms. Let us see the strategy in action, where it might make more sense.

5.16. Theorem. The Monotone Convergence Theorem.
Let $E \in \mathfrak{M}(\mathbb{R})$. Let $(f_n)_{n=1}^\infty$ be a sequence in $\mathcal{L}(E, [0, \infty])$ and suppose that for each $n \ge 1$,

$$f_n \le f_{n+1} \qquad a.e. \text{ on } E.$$

Suppose furthermore that $f : E \to [0, \infty]$ is a function and that

$$f(x) = \lim_{n\to\infty} f_n(x) \qquad a.e. \text{ on } E.$$

Then f is measurable and

$$\int_E f = \lim_{n\to\infty} \int_E f_n.$$

Proof. Step One.

First we shall show that f is measurable. Let

$$E_0 = \{x \in E : f(x) \neq \lim_{n\to\infty} f_n(x)\},$$

so that $mE_0 = 0$ (and in particular E_0 is measurable).

By Lemma 5.14, $f_n|_{E\setminus E_0}$ is measurable for all $n \ge 1$. Since $f|_{E\setminus E_0}$ is the pointwise limit of the measurable functions $f_n|_{E\setminus E_0}$ by hypothesis, it follows that $f|_{E\setminus E_0}$ is measurable by Corollary 4.19.

But $f|_{E_0}$ is also measurable, since for any $\beta \in \mathbb{R}$, $f|_{E_0}^{-1}((\beta,\infty]) \subseteq E_0$ implies that $0 \le m(f|_{E_0}^{-1}((\beta,\infty])) \le mE_0 = 0$, which in turn implies that $f|_{E_0}^{-1}((\beta,\infty])$ is measurable. (See Exercise 4.16 as well.)

By Lemma 5.14, f is measurable.

Step Two.

Next, for each $n \ge 1$, set

$$E_n := \{x \in E : f_n(x) > f_{n+1}(x)\},$$

so that $mE_n = 0$, and thus E_n is measurable. Let $B = E_0 \cup (\cup_{n=1}^\infty E_n)$. Then

$$m^* B \le \sum_{n=0}^\infty mE_n = 0,$$

and so $B \in \mathfrak{M}(E)$ and $mB = 0$. Define $H = E \setminus B$. Let $g_n = f_n|_H$ for all $n \ge 1$, and set $g = f|_H$. Arguing as in Step One, each g_n is measurable, as is g.

For $x \in H$, we have that

$$g_1(x) \le g_2(x) \le g_3(x) \le \cdots \le g(x),$$

and in fact $g(x) = \lim_{n\to\infty} g_n(x)$.

Step Three. The first two steps were only to reduce the problem to the case where the interesting properties hold everywhere. Now the real argument begins.

Since $g_n \le g_{n+1} \le g$ for all $n \ge 1$, by Lemma 5.14, we have that

$$\int_H g_n \le \int_H g_{n+1} \le \int_H g$$

for all $n \ge 1$, and thus

$$\sup_{n\ge1} \int_H g_n = \lim_{n\to\infty} \int_H g_n \le \int_H g.$$

Conversely, suppose that $\varphi \in \mathrm{Simp}(H,[0,\infty])$ and that $\varphi \le g$. Let $0 < \rho < 1$. We shall prove that

$$\int_H \rho\, \varphi = \rho \int_H \varphi \le \sup_{n\ge1} \int_H g_n.$$

Let $x \in H$. Either $\rho\, \varphi(x) < g(x)$, or $\rho\,\varphi(x) = 0$. Setting

$$H_n := \{x \in H : \rho\,\varphi(x) \le g_n(x)\} = (\rho\,\varphi - g_n)^{-1}([-\infty, 0]),$$

we see that $H_1 \subseteq H_2 \subseteq H_3 \subseteq \cdots$ is an increasing sequence of Lebesgue measurable subsets of H with $H = \cup_{n=1}^\infty H_n$.

Step Four. We claim that

$$\lim_{n\to\infty} \int_{H_n} \varphi = \int_H \varphi.$$

To see this, express $\varphi = \sum_{k=1}^N \alpha_k \chi_{J_k}$ in standard form. Then

$$\int_{H_n} \varphi = \int_H \varphi \cdot \chi_{H_n} = \sum_{k=1}^N \alpha_k m(J_k \cap H_n).$$

But for each $1 \le k \le N$, we also see that

$$J_k \cap H_1 \subseteq J_k \cap H_2 \subseteq J_k \cap H_3 \subseteq \cdots$$

is an increasing sequence of measurable sets and

$$\cup_{n=1}^\infty (J_k \cap H_n) = J_k \cap (\cup_{n=1}^\infty H_n) = J_k \cap H = J_k.$$

By the Continuity of Lebesgue Measure (see Exercise 3.9),

$$\lim_{n\to\infty} m(J_k \cap H_n) = mJ_k.$$

It follows that

$$\lim_{n\to\infty} \int_{H_n} \varphi = \lim_{n\to\infty} \sum_{k=1}^N \alpha_k m(J_k \cap H_n) = \sum_{k=1}^N \alpha_k \, mJ_k = \int_H \varphi.$$

Step Five. Hence

$$\rho \int_H \varphi = \rho \left(\lim_{n\to\infty} \int_{H_n} \varphi \right) \le \rho \left(\lim_{n\to\infty} \int_{H_n} g_n \right) \le 1 \cdot \left(\lim_{n\to\infty} \int_H g_n \right) = \sup_{n\ge1} \int_H g_n.$$

Since $\varphi \in \text{Simp}([0,\infty])$ was arbitrary (subject to the condition that $0 \le \varphi \le g$, we conclude that

$$\rho \int_H g \le \sup_{n\ge1} \int_H g_n.$$

But then $0 < \rho < 1$ was also arbitrary, and so

$$\int_H g \le \sup_{n\ge1} \int_H g_n.$$

Combining this with the reverse inequality from Step Three shows that

$$\int_H g = \sup_{n\ge1} \int_H g_n = \lim_{n\to\infty} \int_H g_n.$$

Step Six. There remains only to "glue" the above results together to get the desired statement.

By Lemma 5.14,

$$\int_E f = \int_B f + \int_H f = 0 + \int_H g = \lim_{n\to\infty} \int_H g_n = \lim_{n\to\infty} \int_B f_n + \int_H f_n = \lim_{n\to\infty} \int_E f_n.$$

□

Steps Three to Five of the above proof provide a proof of the Monotone Convergence Theorem in the case where the sequence $(f_n)_{n=1}^\infty$ is everywhere increasing and where the sequence tends to f everywhere.

5.17. Let us remind ourselves of a "pathological" sequence of Riemann integral functions we constructed in Remark 1.16. In that remark, we enumerated the set $E := \mathbb{Q} \cap [0,1] = \{q_n\}_{n=1}^{\infty}$, and set $f = \chi_{E_n}$, where $E_n := \{q_1, q_2, \dots, q_n\}$, $n \ge 1$. We observed that

$$0 \le f_1 \le f_2 \le f_3 \le \cdots \le \chi_E,$$

and that each f_n is Riemann integrable with $\int_0^1 f_n(x)dx = 0$. Moreover,

$$\chi_E(x) = \lim_{n\to\infty} f_n(x) \text{ for all } x \in [0,1],$$

and yet χ_E is *not* Riemann integrable.

We leave it as an exercise for the reader to show that each f_n is Lebesgue integrable with $\int_{[0,1]} f_n = 0$, $n \ge 1$. By the Monotone Convergence Theorem 5.16, we find that $\chi_E \in \mathcal{L}([0,1],\mathbb{R})$ and

$$\int_{[0,1]} \chi_E = \lim_{n\to\infty} \int_{[0,1]} f_n = \lim_{n\to\infty} 0 = 0.$$

This agrees with the fact that

$$0 \le \int_{[0,1]} f = \int_{[0,1]} \chi_E = mE \le m\mathbb{Q} = 0.$$

The point is that the limit function χ_E is Lebesgue integrable, even though it is not Riemann integrable.

The first half the of the following Corollary extends results from Proposition 5.7. There, we dealt with nonnegative, simple, measurable functions. We remove the requirement that the functions be simple.

5.18. Corollary. *Let $E \in \mathfrak{M}(\mathbb{R})$.*

(a) *If $f, g \in \mathcal{L}(E,[0,\infty])$ and $\kappa \ge 0$, then*

$$\int_E \kappa f + g = \kappa \int_E f + \int_E g.$$

(b) *If $(h_n)_{n=1}^{\infty}$ is a sequence in $\mathcal{L}(E,[0,\infty])$ and if $h(x) := \lim_{N\to\infty} \sum_{n=1}^{N} h_n(x)$ for all $x \in E$, then h is measurable and*

$$\int_E h = \sum_{n=1}^{\infty} \int_E h_n.$$

(c) *Let $f \in \mathcal{L}(E,[0,\infty])$. If $(H_n)_{n=1}^{\infty}$ is a sequence in $\mathfrak{M}(E)$ with $H_i \cap H_j = \varnothing$ when $1 \le i \ne j < \infty$ and $H = \cup_{n=1}^{\infty} H_n$, then*

$$\int_H f = \sum_{n=1}^{\infty} \int_{H_n} f.$$

Proof.

(a) From our work in the Assignments (see Exercise 4.15), we may choose increasing sequences $(\varphi_n)_{n=1}^\infty$ and $(\psi_n)_{n=1}^\infty$ in $\text{SIMP}(E, [0, \infty])$ such that $f(x) = \lim_{n\to\infty} \varphi_n(x)$ and $g(x) = \lim_{n\to\infty} \psi_n(x)$ for all $x \in E$.

By the Monotone Convergence Theorem,

$$\int_E f = \lim_{n\to\infty} \int_E \varphi_n$$

and

$$\int_E g = \lim_{n\to\infty} \int_E \psi_n.$$

Given $0 \le \kappa \in \mathbb{R}$, it follows that $(\kappa\varphi_n + \psi_n)_{n=1}^\infty$ is again an increasing sequence of non-negative, simple, measurable, functions converging pointwise to the function $\kappa f + g$.

Applying the Monotone Convergence Theorem 5.16 once more, we see that

$$\int_E (\kappa f + g) = \lim_{n\to\infty} \int_E (\kappa\varphi_n + \psi_n) = \lim_{n\to\infty} \kappa \int_E \varphi_n + \int_E \psi_n = \kappa \int_E f + \int_E g.$$

(b) For $N \ge 1$, set $g_N := \sum_{n=1}^N h_n$. Then each g_N is measurable (exercise), and $\int_E g_N = \sum_{n=1}^N \int_E h_n$ by part (a). Furthermore,

$$0 \le g_1 \le g_2 \le \cdots.$$

Now $h(x) = \lim_{N\to\infty} g_N(x)$ for all $x \in E$, and so h is measurable by Corollary 4.19.

By the Monotone Convergence Theorem,

$$\int_E h = \lim_{N\to\infty} \int_E g_N = \lim_{N\to\infty} \sum_{n=1}^N \int_E h_n = \sum_{n=1}^\infty \int_E h_n.$$

(c) For each $n \ge 1$, set $f_n = f \cdot \chi_{H_n}$. Then f_n is measurable for each $n \ge 1$ (exercise) and

$$f \cdot \chi_H(x) = \lim_{N\to\infty} \sum_{n=1}^N f_n(x) \qquad \text{for all } x \in E.$$

By part (b),

$$\begin{aligned} \int_H f &= \int_E f \cdot \chi_H \\ &= \int_E \sum_{n=1}^\infty f_n \\ &= \sum_{n=1}^\infty \int_E f_n \\ &= \sum_{n=1}^\infty \int_E f \cdot \chi_{H_n} \\ &= \sum_{n=1}^\infty \int_{H_n} f. \end{aligned}$$

□

5.19. Definition. The Lebesgue Integral
Let $E \in \mathfrak{M}(\mathbb{R})$ and $f \in \mathcal{L}(E, \overline{\mathbb{R}})$. We say that f is ***Lebesgue integrable*** *on E if*

$$\int_E f^+ < \infty \qquad \text{and} \qquad \int_E f^- < \infty,$$

in which case we set

$$\int_E f := \int_E f^+ - \int_E f^-.$$

We denote by $\mathcal{L}_1(E, \overline{\mathbb{R}})$ the set of all extended real-valued Lebesgue integrable functions on E, and by $\mathcal{L}_1(E, \mathbb{R})$ all real-valued Lebesgue integrable functions on E.

In Chapter 9 and later, we shall need to specify the variable with respect to which we are integrating. In analogy to the usual notation for Riemann integration, we shall write

$$\int_E f = \int_E f(s) dm(s)$$

to denote the Lebesgue integral of f with respect to the variable s.

5.20. Remarks. Let $E \in \mathfrak{M}(\mathbb{R})$.

(a) By definition, every Lebesgue integrable function on E is Lebesgue measurable.

(b) A *measurable* function f is Lebesgue integrable if and only if $|f|$ is Lebesgue integrable. This fails in general if f is not assumed to be measurable. (Consider $f = \chi_H - \chi_{[0,1] \setminus H}$, where $H \subseteq [0,1]$ is any non-measurable set. Clearly $|f|$ is Lebesgue integrable over $[0,1]$, while f is not.)

Note that this is also a distinguishing feature of Lebesgue integration versus improper Riemann integrals. For example, the function

$$f(x) = \frac{\sin x}{x}, \quad x \geq 1$$

admits an (improper) Riemann integral $\int_1^\infty f(x) dx$, but it is *not* Lebesgue integrable over $[1, \infty)$.

(c) If $f \in \mathcal{L}_1(E, \overline{\mathbb{R}})$, then

$$m(f^{-1}(\{-\infty\})) = 0 = m(f^{-1}(\{\infty\})).$$

The proof of this is left as an exercise for the reader.

(d) From (c), it follows that for any $f \in \mathcal{L}_1(E, \overline{\mathbb{R}})$, we can find an element $g \in \mathcal{L}_1(E, \overline{\mathbb{R}})$ such that $f = g$ a.e. and $g(x) \in \mathbb{R}$ for all $x \in E$. Indeed, just let $H = \{x \in E : f(x) \in \{-\infty, \infty\}\}$. By (c), $mH = 0$, and we may simply choose $g = f \cdot \chi_{E \setminus H}$.

Note that this in turn implies that

$$\int_E f = \int_E g.$$

This will prove to be more useful than it might first appear to be. One huge problem with $\mathcal{L}_1(E, \overline{\mathbb{R}})$ is that it is *not a vector space*!!! One problem lies in the fact that if $f, g \in \mathcal{L}_1(E, \overline{\mathbb{R}})$, $x \in E$ and $f(x) = \infty$, $g(x) = -\infty$, then what should $(f+g)(x)$ be?

This is not an issue insofar as $\mathcal{L}_1(E, \mathbb{R})$ is concerned. As we shall see in Chapter 6, we may establish an equivalence relation on $\mathcal{L}_1(E, \overline{\mathbb{R}})$ by setting $f \sim g$ if $f = g$ a.e. on E. We can then turn the *equivalence classes* of elements of $\mathcal{L}_1(E, \overline{\mathbb{R}})$ into a vector space in a natural way. In most texts, these equivalence classes are denoted by the same notation used to denote functions, and indeed, they are often referred to as "functions", although technically speaking they are not. Being absolute sticklers for detail, and inspired by our French heritage, we shall exercise as much caution as possible in the use of language, and will try to be as precise as humanly possible in keeping the notation and terminology straight.

(e) Suppose that $g : E \to \mathbb{C}$ is a measurable function. Let us write

$$g = (g_1 - g_2) + i(g_3 - g_4),$$

where $g_1 = (\operatorname{Re} g)^+$, $g_2 = (\operatorname{Re} g)^-$, $g_3 = (\operatorname{Im} g)^+$, and $g_4 = (\operatorname{Im} g)^-$.

We shall say that g is **Lebesgue integrable** if each of g_1, g_2, g_3 and g_4 are, in which case we define

$$\int_E g = (\int_E g_1 - \int_E g_2) + i(\int_E g_3 - \int_E g_4).$$

Of course, this is equivalent to requiring that $\operatorname{Re} g$ and $\operatorname{Im} g$ be Lebesgue integrable, in which case we define

$$\int_E g = \int_E (\operatorname{Re} g) + i \int_E (\operatorname{Im} g).$$

We denote by $\mathcal{L}_1(E, \mathbb{C})$ the set of all complex-valued Lebesgue integrable functions on E.

The perspicacious reader will observe that we have carefully avoided all notions of "extended" complex-valued functions.

5.21. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$. Suppose that $f, g \in \mathcal{L}_1(E, \mathbb{R})$, and $\kappa \in \mathbb{R}$. Then*

(a) *$\kappa f \in \mathcal{L}_1(E, \mathbb{R})$ and $\int_E \kappa f = \kappa \int_E f$.*
(b) *$f + g \in \mathcal{L}_1(E, \mathbb{R})$ and $\int_E (f+g) = \int_E f + \int_E g$.*
(c) *Finally,*

$$\left| \int_E f \right| \le \int_E |f|.$$

Proof. Recall from Corollary 5.18 that we have already shown that (a) and (b) hold in the case where $0 \le f, g$ and $\kappa \ge 0$.

(a) Let us write $f = f^+ - f^-$.

CASE 1. $\kappa = 0$.

Then $\int_E \kappa f = \int_E 0 = 0 = \kappa \int_E f$.

CASE 2. $\kappa > 0$.

Then $(\kappa f)^+ = \kappa f^+$ and $(\kappa f)^- = \kappa f^-$, so that

$$\begin{aligned}\int_E \kappa f &:= \int_E (\kappa f)^+ - \int_E (\kappa f)^- \\ &= \int_E \kappa f^+ - \int_E \kappa f^- \\ &= \kappa \int_E f^+ - \kappa \int_E f^- \\ &= \kappa \int_E f.\end{aligned}$$

Case 3. $\kappa < 0$.

Then $(\kappa f)^+ = -\kappa f^-$ and $(\kappa f)^- = -\kappa f^+$, so that

$$\begin{aligned}\int_E \kappa f &:= \int_E (\kappa f)^+ - \int_E (\kappa f)^- \\ &= \int_E (-\kappa) f^- - \int_E (-\kappa) f^+ \\ &= (-\kappa) \int_E f^- - (-\kappa) \int_E f^+ \\ &= (-\kappa)(- \int_E f) \\ &= \kappa \int_E f.\end{aligned}$$

(b) Let $h = f + g$, so that $h \in \mathcal{L}(E, \mathbb{R})$, as the latter is a vector space. Write $h = h^+ - h^-$.

Then

$$h^+, h^- \le h^+ + h^- = |h| \le |f| + |g| = f^+ + f^- + g^+ + g^-,$$

so that

$$\int_E h^+ \le \int_E f^+ + \int_E f^- + \int_E g^+ + \int_E g^- < \infty,$$

and

$$\int_E h^- \le \int_E f^+ + \int_E f^- + \int_E g^+ + \int_E g^- < \infty.$$

It follows that $h \in \mathcal{L}_1(E, \mathbb{R})$.

Furthermore, $h = f + g$ implies that $h^+ + f^- + g^- = h^- + f^+ + g^+$, whence

$$\int_E h^+ + \int_E f^- + \int_E g^- = \int_E h^- + \int_E f^+ + \int_E g^+.$$

From this it easily follows that

$$\int_E h = \int_E h^+ - \int_E h^- = \int_E f + \int_E g.$$

(c) Note that $|f| = f^+ + f^-$ is measurable, and

$$\int_E |f| = \int_E f^+ + \int_E f^- < \infty,$$

proving that $|f| \in \mathcal{L}_1(E, \mathbb{R})$.

Finally,

$$\begin{aligned}\left|\int_E f\right| &= \left|\int_E f^+ - \int_E f^-\right| \\ &\leq \left|\int_E f^+\right| + \left|\int_E f^-\right| \\ &= \int_E f^+ + \int_E f^- \\ &= \int_E |f|.\end{aligned}$$

□

5.22. At the moment, we have a number of results concerning Lebesgue integrals, but we have not explicitly calculated the Lebesgue integrals of any functions, other than simple functions. In the Assignments, you will be asked to compute the Lebesgue integral of the function $f(x) = x$, $x \in [0,1]$ *by hand*. This will lead to an unusual increase in swearing on your part, and a raising of the ole' blood pressure. Surely, you will tell yourself, there *must* be a better way!

We shall now demonstrate that in the case of *bounded* functions on closed, *bounded* intervals, Riemann integrability implies Lebesgue integrability, and indeed for a given Riemann integrable function, the Riemann and Lebesgue integrals coincide. Since we have a number of tools to calculate Riemann integrals (for example, the **Fundamental Theorem of Calculus**), this will prove to be the better way in a large number of cases.

Just a quick remark about notation: we shall continue to use $\int_a^b f$ to denote the Riemann integral of Riemann integrable function $f : [a,b] \to \mathbb{R}$ and $\int_{[a,b]} f$ to denote its Lebesgue integral.

We start with a simple but useful Lemma.

5.23. Lemma. *Let $a < b \in \mathbb{R}$. If $\varphi : [a,b] \to \mathbb{R}$ is a step function, then φ is both Riemann integrable and Lebesgue integrable, and*

$$\int_{[a,b]} \varphi = \int_a^b \varphi.$$

Proof. Let $P = \{a = p_0 < p_1 < p_2 < \cdots < p_N = b\} \in \mathcal{P}([a,b])$, and

$$\varphi = \sum_{n=1}^N \alpha_n \chi_{[p_{n-1},p_n)}.$$

Then

$$\begin{aligned}\int_{[a,b]} \varphi &= \sum_{n=1}^{N} \alpha_n \, m[p_{n-1}, p_n) \\ &= \sum_{n=1}^{N} \alpha_n (p_n - p_{n-1}) \\ &= \sum_{n=1}^{N} \int_{p_{n-1}}^{p_n} \alpha_n \\ &= \int_a^b \sum_{n=1}^{N} \alpha_n \chi_{[p_{n-1}, p_n)} \\ &= \int_a^b \varphi.\end{aligned}$$

□

5.24. Theorem. *Let $a < b \in \mathbb{R}$ and $f : [a,b] \to \mathbb{R}$ be a bounded, Riemann-integrable function. Then $f \in \mathcal{L}_1([a,b], \mathbb{R})$ and*

$$\int_{[a,b]} f = \int_a^b f.$$

That is, the Lebesgue and Riemann integrals of f over $[a,b]$ coincide.

Proof. Suppose that $|f|$ is bounded above on $[a,b]$ by $0 < M \in \mathbb{R}$. Let $g = M\chi_{[a,b]}$. Clearly $\frac{1}{2}(f+g)$ is Riemann integrable and

$$\int_a^b \frac{1}{2}(f+g) = \frac{1}{2}\int_a^b f + \frac{M(b-a)}{2}.$$

If we prove that $\frac{1}{2}(f+g) \in \mathcal{L}_1([a,b], \mathbb{R})$, then it is readily seen that $f \in \mathcal{L}_1([a,b], \mathbb{R})$ and – in light of Lemma 5.23 –

$$\int_{[a,b]} \frac{1}{2}(f+g) = \frac{1}{2}\int_{[a,b]} f + \frac{M(b-a)}{2}.$$

We have shown that by replacing f by $\frac{1}{2}(f+g)$ if necessary, we may assume from the outset that $0 \le f \le M$ on $[a,b]$.

Cast your mind back to the halcyon days when you studied Chapter 1, and more specifically to the Cauchy Criterion, Theorem 1.13. Recall that it asserts that for each $n \ge 1$ there exists a partition $R_n \in \mathcal{P}([a,b])$ such that for all refinements X and Y of R_n, and for all choices of test values X^* and Y^* for X and Y respectively,

$$|S(f, X, X^*) - S(f, Y, Y^*)| < \frac{1}{n}.$$

Let $Q_N := \cup_{n=1}^N R_n$, $1 \le N \in \mathbb{N}$, so that Q_N is a common refinement of $R_1, R_2, \ldots, R_N$. Write

$$Q_N = \{a = q_{0,N} < q_{1,N} < q_{2,N} < \cdots < q_{m_N,N} = b\}.$$

Set $H_{k,N} = [q_{k-1,N}, q_{k,N})$, $1 \le k \le m_N - 1$ and $H_{m_N,N} = [q_{m_N-1,N}, q_{m_N,N}]$.

Define

$$\alpha_{k,n} = \inf\{f(t) : t \in H_{k,n}\}, \quad 1 \le k \le m_N,$$
$$\beta_{k,n} = \sup\{f(t) : t \in H_{k,n}\}, \quad 1 \le k \le m_N,$$

and set $\varphi_N = \sum_{n=1}^{N} \sum_{k=1}^{m_N} \alpha_{k,n} \chi_{H_{k,n}}$ and $\psi_N = \sum_{n=1}^{N} \sum_{k=1}^{m_N} \beta_{k,n} \chi_{H_{k,n}}$.

Since each Q_N is a refinement of Q_{N-1}, it is not hard to see that

$$\varphi_1 \le \varphi_2 \le \varphi_3 \le \cdots \le f \le \cdots \le \psi_3 \le \psi_2 \le \psi_1.$$

Moreover, using Lemma 5.23 above, we obtain

$$\int_{[a,b]} \varphi_N = \int_a^b \varphi_N = \inf\{S(f, Q_N, Q_N^*) : Q_N^* \text{ test values for } Q_N\} \le \int_a^b f,$$

and similarly,

$$\int_{[a,b]} \psi_N = \int_a^b \psi_N = \sup\{S(f, Q_N, Q_N^{**}) : Q_N^{**} \text{ test values for } Q_N\} \ge \int_a^b f.$$

But Q_N is a refinement of R_N, and thus

$$|S(f, Q_N, Q_N^*) - S(f, Q_N, Q_N^{**})| < \frac{1}{N}$$

for all choices of test values Q_N^* and Q_N^{**} for Q_N.

It follows that $\int_{[a,b]} \varphi_N \le \int_a^b f \le \int_{[a,b]} \psi_N$ and

$$|\int_{[a,b]} \varphi_N - \int_{[a,b]} \psi_N| \le \frac{1}{N}, \quad N \ge 1.$$

Let $\varphi(x) = \sup_{n \ge 1} \varphi_n(x) = \lim_{n\to\infty} \varphi_n(x) \le f(x)$, and $\psi(x) = \inf_{n\ge 1} \psi_n(x) = \lim_{n\to\infty} \psi_n(x) \ge f(x)$, $x \in [a,b]$. Since each φ_n, ψ_n is measurable, so are φ and ψ, and by the Monotone Convergence Theorem and Lemma 5.23,

$$\begin{aligned}\int_{[a,b]} \varphi &= \lim_{n\to\infty} \int_{[a,b]} \varphi_N \\ &= \lim_{n\to\infty} \int_a^b \varphi_N \\ &= \int_a^b f \\ &= \lim_{n\to\infty} \int_a^b \psi_N \\ &= \lim_{n\to\infty} \int_{[a,b]} \psi_N \\ &= \int_{[a,b]} \psi.\end{aligned}$$

Thus $\int_{[a,b]} \psi - \varphi = 0$. But $\varphi \le \psi$, and so we must have $\varphi = \psi$ a.e. on $[a,b]$. Finally, note that $\varphi \le f \le \psi$, which in turn implies that $f = \varphi = \psi$ a.e. on $[a,b]$. Since φ is measurable, so is f, and

$$\int_{[a,b]} f = \int_{[a,b]} \varphi = \int_a^b f < \infty.$$

□

5.25. Corollary. *Let $a < b \in \mathbb{R}$ and $f : [a, b] \to \mathbb{C}$ be a bounded, Riemann-integrable function. Then $f \in \mathcal{L}_1([a, b], \mathbb{C})$ and*

$$\int_{[a,b]} f = \int_a^b f.$$

Proof. Observe that f is bounded and Riemann-integrable if and only if its real and imaginary parts are bounded and Riemann-integrable. The result now immediately follows by applying Theorem 5.24 to each of these.

□

Theorem 5.24 required some effort. Let us see that it was worth it.

5.26. Example. Let $f(x) = x$, $x \in [0, 1]$. In the Assignments, you computed the Lebesgue integral of f over $[0, 1]$ to be

$$\int_{[0,1]} x = \frac{1}{2}.$$

This was anything but easy.

Equipped with Theorem 5.24, it is child's play. The function f is clearly bounded and continuous (hence Riemann-integrable) over $[0, 1]$, and so by that Theorem,

$$\int_{[0,1]} x = \int_0^1 x\,dx = \frac{x^2}{2}\Big]_{x=0}^{x=1} = \frac{1}{2}.$$

Yes, Theorem 5.24 is worth the effort.

5.27. Example. Let $f(x) = \frac{1}{x^2}$, $x \in E := [1, \infty)$. We wish to determine $\int_{[1,\infty)} f$.

For each $n \geq 1$, set $f_n := f \cdot \chi_{[1,n]}$. Then f is measurable (because it is continuous except at one point of E), and

$$0 \leq f_1 \leq f_2 \leq \cdots,$$

with $f(x) = \lim_{n\to\infty} f_n(x)$ for all $x \geq 1$. By Theorem 5.24,

$$\int_{[1,n]} f_n = \int_1^n f_n \qquad \text{for all } n \geq 1.$$

By the Monotone Convergence Theorem,

$$\begin{aligned}\int_{[1,\infty)} f &= \lim_{n\to\infty} \int_{[1,\infty)} f_n \\ &= \lim_{n\to\infty} \int_{[1,n]} \frac{1}{x^2} \\ &= \lim_{n\to\infty} \int_1^n \frac{1}{x^2} \\ &= \lim_{n\to\infty} -\frac{1}{x}\Big]_{x=1}^{x=n} \\ &= \lim_{n\to\infty} \left(-\frac{1}{n} - (-1)\right) \\ &= 1.\end{aligned}$$

In this example, the Lebesgue integral of f returns the value of the improper Riemann integral of f over $[1,\infty)$. Two things are worth noting:

- first, it is possible for the improper Riemann integral of a measurable function $f : [1,\infty) \to \mathbb{R}$ to exist, even though the Lebesgue integral $\int_{[1,\infty)} f$ *does not exist*! We shall see an example of this in the Assignments.
- Second, we don't have the notion of an *improper* Lebesgue integral. The domain of f, $[1,\infty)$, is just another measurable set.

5.28. Example. The Monotone Convergence Theorem 5.16 states that if $(f_n)_{n=1}^\infty$ is an (almost everywhere) increasing sequence of measurable functions on a set $E \in \mathfrak{M}(\mathbb{R})$, then $f = \lim_{n\to\infty} f_n$ is measurable and

$$\int_E \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_E f_n.$$

In the absence of the adjective "increasing", we can not expect this result to hold.

For example, consider the sequence $(f_n)_{n=1}^\infty$ given by

$$\begin{aligned} f_n : \quad [1,\infty) &\to \mathbb{R} \\ x &\mapsto \begin{cases} \frac{1}{nx} & \text{if } 1 \le x \le e^n \\ 0 & \text{if } x > e^n \end{cases}. \end{aligned}$$

It is easy to verify (exercise) that the sequence $(f_n)_{n=1}^\infty$ converges *uniformly* to $f = 0$ on $[1,\infty)$.

Nevertheless, for all $n \geq 1$, f_n is easily seen to be Riemann-integrable and bounded on $[1, e^n]$, and thus

$$\begin{aligned}\int_{[1,\infty)} f_n &= \int_{[1,e^n]} \frac{1}{nx} \\ &= \int_1^{e^n} \frac{1}{nx} \\ &= \frac{\log x}{n}\Big]_{x=1}^{x=e^n} \\ &= 1 - 0 \\ &= 1.\end{aligned}$$

Hence $\lim_{n\to\infty} \int_{[1,\infty)} f_n = 1 \neq 0 = \int_{[1,\infty)} f$.

I have done a bit (but not a great deal) of research to try to determine why the next result is referred to as Fatou's Lemma instead of Fatou's *Theorem*. One possible explanation is that it can be used to prove a number of other useful results very quickly, and as such, is a "facilitator", to employ the jolly discourse of psychobabble. A second possibility (which I have not read anywhere) is that it is petty jealousy on the part of his peers. In any case, there is a different result from complex analysis known as **Fatou's Theorem**. In order to state it, we first require the notion of L_p-spaces, and so we defer its statement to the Appendix of Chapter 6.

5.29. Theorem. Fatou's Lemma.
Let $E \in \mathfrak{M}(\mathbb{R})$ and $(f_n)_{n=1}^\infty$ be a sequence in $\mathcal{L}(E, [0, \infty])$. Then

$$\int_E \liminf_n f_n \leq \liminf_n \int_E f_n.$$

Proof. For each $N \geq 1$, set $g_N = \inf\{f_n : n \geq N\}$. By Proposition 4.18, g_N is measurable for all N and

$$g_1 \leq g_2 \leq g_3 \leq \cdots.$$

By the Monotone Convergence Theorem 5.16,

$$\int_E \liminf_n f_n = \int_E \lim_{N\to\infty} g_N = \lim_{N\to\infty} \int_E g_N.$$

Now $g_N \leq f_n$ for all $n \geq N$, and so

$$\int_E g_N \leq \int_E f_n \text{ for all } n \geq N,$$

whence

$$\int_E g_N \leq \liminf_n \int_E f_n.$$

But this holds for any $N \geq 1$, and so by taking limits and using the above equality, we find that

$$\int_E \liminf_n f_n = \lim_{N\to\infty} \int_E g_N \leq \liminf_n \int_E f_n.$$

□

5.30. Example. The inequality in Fatou's Lemma can be strict.

For example, let $f_n = n\chi_{(0,\frac{1}{n}]}$, $n \ge 1$. It is clear that for any $0 \le x \le 1$, $\lim_{n\to\infty} f_n(x) = 0$. Thus

$$\int_{[0,1]} \liminf_n f_n = \int_{[0,1]} 0 = 0.$$

On the other hand,

$$\int_{[0,1]} f_n = n\, m((0, \frac{1}{n}]) = 1 \qquad \text{for all } n \ge 1,$$

and so $\liminf_n \int_{[0,1]} f_n = 1$.

5.31. Example. Let $E = [0,\infty) \in \mathfrak{M}(\mathbb{R})$, and for each $n \ge 1$, let $f_n = -\frac{1}{n}\chi_{[n,2n]}$. Then each f_n is measurable and $(f_n)_n$ converges *uniformly* to $f \equiv 0$ on E. *A fortiori*, $(f_n)_n$ converges pointwise to f.

Neverthess,

$$\int_E \liminf_n f_n = \int_E 0 = 0 > -1 = \liminf_n \int_E f_n.$$

This shows that we cannot simply drop the assumption of non-negativity of the functions $(f_n)_n$ in the hypotheses of Fatou's Lemma and hope for the same conclusion.

5.32. Suppose that $E \in \mathfrak{M}(\mathbb{R})$ and that $f, g : E \to \overline{\mathbb{R}}$ are measurable. Suppose furthermore that $0 \le |f| \le g$ a.e. on E, and that $\int_E g < \infty$, i.e. that $g \in \mathcal{L}_1(E, \overline{\mathbb{R}})$. We claim that $f \in \mathcal{L}_1(E, \overline{\mathbb{R}})$.

Indeed, let $B = \{x \in E : |f(x)| > g(x)\}$, so that B is measurable with $mB = 0$. For $x \in E \setminus B$, we have that

$$f^+(x) \le |f(x)| \le g(x) \qquad \text{and} \qquad f^-(x) \le |f(x)| \le g(x).$$

From Lemma 5.14 and Proposition 5.15 we deduce that

$$\int_E f^+ = \int_{E\setminus B} f^+ \le \int_{E\setminus B} g = \int_E g < \infty,$$

and similarly,

$$\int_E f^- = \int_{E\setminus B} f^- \le \int_{E\setminus B} g = \int_E g < \infty.$$

Hence $f = f^+ - f^- \in \mathcal{L}_1(E, \overline{\mathbb{R}})$.

The following result is also one of the major results of measure theory.

5.33. Theorem. The Dominated Convergence Theorem.
Let $E \in \mathfrak{M}(\mathbb{R})$ and $(f_n)_{n=1}^\infty$ is a sequence in $\mathcal{L}(E, \overline{\mathbb{R}})$. Suppose that there exists $g \in \mathcal{L}_1(E, \overline{\mathbb{R}})$ such that $|f_n| \le g$ a.e. on E, $n \ge 1$. Suppose furthermore that $f : E \to \overline{\mathbb{R}}$ is a function and that

$$f(x) = \lim_{n\to\infty} f_n(x) \qquad \text{a.e. on } E.$$

Then $f \in \mathcal{L}_1(E, \overline{\mathbb{R}})$ and

$$\int_E f = \lim_{n\to\infty} \int_E f_n.$$

Proof.
STEP ONE. As with the Monotone Convergence Theorem, our first goal is to isolate the "bad" set of points where the convergence of the sequence fails, or where the sequence of functions is not bounded above by g.

For each $n \ge 1$, set $E_n = \{x \in E : |f_n(x)| > g(x)\}$ and set $E_0 := \{x \in E : f(x) \ne \lim_{n\to\infty} f_n(x)\}$. Finally, set $E_\infty := \{x \in E : g(x) = \infty\}$. By hypothesis, $mE_n = 0$, $0 \le n < \infty$, while $mE_\infty = 0$ as $g \in \mathcal{L}_1(E, \overline{\mathbb{R}})$. Thus, if we set $B := E_\infty \cup (\cup_{n=0}^\infty E_n)$, then B is measurable and

$$0 \le mB \le mE_\infty + \sum_{n=0}^\infty mE_n = 0 + \sum_{n=0}^\infty 0 = 0,$$

or in other words, $mB = 0$. From this it follows that for all $n \ge 1$, $\int_B f_n = 0$. Moreover, $f|_B$ is measurable, and $\int_B f = 0$.

Note that for $x \in H := E \setminus B$, we have that $f(x) = \lim_{n\to\infty} f_n(x)$ and we also have $|f_n(x)| \le g(x)$. Given that $mB = 0$, it follows from Lemma 5.14 that $f|_H$ is measurable if and only if f is. Moreover, $mB = 0$ also implies that

$$\int_H f = \lim_{n\to\infty} \int_H f_n$$

if and only if

$$\int_E f = \lim_{n\to\infty} \int_E f_n.$$

In other words, by replacing E with H if necessary, we may assume without loss of generality that $|f_n(x)| \le g(x) < \infty$ and that $f(x) = \lim_{n\to\infty} f_n(x)$ for all $x \in E$. We shall assume that we have done this.

STEP TWO.

Note that $g - f_n \ge 0$ on E, and thus by Fatou's Lemma 5.29,

$$\begin{aligned}\int_E g - \int_E \limsup_n f_n &= \int_E \liminf_n \; g - f_n \\ &\le \liminf_n \int_E \; g - f_n \\ &= \int_E g - \limsup_n \int_E f_n,\end{aligned}$$

which – given that $f(x) = \lim_n f_n(x)$ for all $x \in E$ – is equivalent to

$$\limsup_n \int_E f_n \le \int_E f.$$

But $g + f_n \geq 0$ on E as well, and hence a second application of Fatou's Lemma yields

$$\begin{aligned}\int_E g + \int_E \liminf_n f_n &= \int_E \liminf_n (g + f_n) \\ &\leq \liminf_n \int_E (g + f_n) \\ &= \int_E g + \liminf_n \int_E f_n,\end{aligned}$$

or equivalently

$$\int_E f \leq \liminf_n \int_E f_n.$$

Putting these two inequalities together shows that

$$\int_E f = \lim_{n\to\infty} \int_E f_n.$$

□

Appendix to Section 5.

5.34. Let $E \in \mathfrak{M}(\mathbb{R})$ and $g \in \mathcal{L}_1(E, \overline{\mathbb{R}})$. We have observed that the set $B := \{x \in E : g(x) \in \{-\infty, \infty\}\}$ has measure zero. When necessary, as it was in the proof of the Lebesgue Dominated Convergence Theorem, we were able to simply "excise" this set from the domain and concentrate our attention to the set $E \setminus B$. So why introduce the extended real-numbers at all?

Convenience. Given an increasing sequence $(f_n)_{n=1}^\infty$ in $\mathcal{L}_1(E, [0, \infty))$, the pointwise limit $f(x) := \lim_{n\to\infty} f_n(x)$ need not be real-valued. By introducing the extended real numbers, we are able to treat the limit function f as simply another measurable function.

When we define the L_p-spaces in Chapter 6, we shall define each $L_p(E)$ as equivalence classes of (extended real-valued) functions. However, each such equivalence class will always admit a representative which is real-valued function. Truly, fortune smiles upon us.

Exercises for Section 5.

Exercise 5.1. Let $f(x) = x$, $x \in [0, \infty)$. Prove that $\int_{[0,\infty)} f = \infty$.

Exercise 5.2. Assignment Question.
Let $f(x) = x$, $x \in \mathbb{R}$. Calculate the Lebesgue integral $\int_{[0,1]} f$ directly from the definition – that is, do *not* appeal to Theorem 5.24.

Exercise 5.3.
Prove Lemma 5.14; that is, let $E \in \mathfrak{M}(\mathbb{R})$ and let f, g and $h : E \to [0, \infty]$ be functions. Suppose that g and h are measurable.

(a) Suppose furthermore that $E = X \cup Y$, where X and Y are measurable. Set $f_1 := f|_X$ and $f_2 := f|_Y$. Then f is measurable if and only if both f_1 and f_2 are measurable. When such is the case,
$$\int_E f = \int_X f_1 + \int_Y f_2.$$
(b) If $g \le h$, then $\int_E g \le \int_E h$.
(c) If $H \subseteq E$ is a measurable set, then
$$\int_H g = \int_E g \cdot \chi_H \le \int_E g.$$

Exercise 5.4.
Prove that if $f \in \mathcal{L}_1(E, \overline{\mathbb{R}})$, then
$$m(f^{-1}(\{-\infty\})) = 0 = m(f^{-1}(\{\infty\})).$$

Exercise 5.5.
Let $f(x) = \dfrac{\sin x}{x}$, $x \ge 1$.

(a) Prove that the improper Riemann integral $\int_1^\infty f(x)dx$ exists.
(b) Prove that the Lebesgue integral $\int_{[1,\infty)} f = \infty$ does not exist.

Exercise 5.6. For $n \ge 1$, define the function
$$\begin{array}{rccc} f_n : & [1, \infty) & \to & \mathbb{R} \\ & x & \mapsto & \begin{cases} \frac{1}{nx} & \text{if } 1 \le x \le e^n \\ 0 & \text{if } x > e^n \end{cases} . \end{array}$$
Verify that the sequence $(f_n)_{n=1}^\infty$ converges *uniformly* to $f = 0$ on $[1, \infty)$.

Exercise 5.7.
Let $E \in \mathfrak{M}(\mathbb{R})$. Suppose that $f, g \in \mathcal{L}_1(E, \mathbb{C})$, and that $\kappa \in \mathbb{C}$. Prove that

(a) the function $\kappa f + g \in \mathcal{L}_1(E, \mathbb{C})$ and
$$\int_E \kappa f + g = \kappa \int_E f + \int_E g.$$

(b) Prove that $|f| \in \mathcal{L}_1(E, \mathbb{C})$ and

$$\left| \int_E f \right| \le \int_E |f|.$$

Exercise 5.8.

Let $E \in \mathfrak{M}(\mathbb{R})$. Show that a measurable function $f : E \to \mathbb{C}$ lies in $\mathcal{L}_1(E, \mathbb{C})$ if and only if $|f| \in \mathcal{L}_1(E, \mathbb{C})$.

Show that this fails if we do not assume that f is measurable.

Exercise 5.9.

The following special case of the Dominated Convergence Theorem is easily derived from the Monotone Convergence Theorem.

Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $(f_n)_{n=1}^\infty$ is a decreasing sequence in $\mathcal{L}(E, [0, \infty])$ with $f_1 \in \mathcal{L}_1(E, [0, \infty])$. Thus

$$f_1 \ge f_2 \ge f_3 \ge \cdots \ge 0.$$

Define $f : E \to [0, \infty]$ by $f(x) = \lim_{n\to\infty} f_n(x)$, $x \in E$.

Prove that f is measurable and that

$$\int_E f = \lim_{n\to\infty} \int_E f_n.$$

Exercise 5.10.

Let $E \in \mathfrak{M}(\mathbb{R})$. If $\varphi, \psi \in \mathrm{SIMP}(E, [0, \infty])$ and $\kappa = \infty$, prove or disprove that

$$\int_E \kappa\varphi + \psi = \kappa \int_E \varphi + \int_E \psi.$$

6. L_p Spaces

I've been on food stamps and welfare. Anybody help me out? No!

Craig T. Nelson

6.1. Functional analysis is the study of normed linear spaces and the continuous linear maps between them. Amongst the most important examples of Banach spaces are the so-called L_p-spaces, and it is to these that we now turn our attention. The reader may wish to refresh his/her memory as to the definition of a *seminorm* on a vector space $\mathfrak{X}$ over $\mathbb{K}$ (Definition 1.2).

6.2. Example. Let $E \subseteq \mathbb{R}$ be a Lebesgue measurable set, and suppose that $mE > 0$. Recall that

$$\mathcal{L}_1(E, \mathbb{K}) = \{f : E \to \mathbb{K} : f \text{ is measurable and } \int_E |f| < \infty\}.$$

Define the map

$$\begin{array}{rccc} \nu_1 : & \mathcal{L}_1(E, \mathbb{K}) & \to & \mathbb{R} \\ & f & \mapsto & \int_E |f|. \end{array}$$

Observe that

- $\nu_1(f) \geq 0$ for all $f \in \mathcal{L}_1(E, \mathbb{K})$.
- $\nu_1(0) = \int_E |0| = 0$.
- If $k \in \mathbb{K}$, then $\nu_1(kf) = \int_E |kf| = |k| \int_E |f| = |k| \nu_1(f)$.
- If $f, g \in \mathcal{L}_1(E, \mathbb{K})$, then

$$\nu_1(f + g) = \int_E |f + g| \leq \int_E |f| + |g| = \int_E |f| + \int_E |g| = \nu_1(f) + \nu_1(g).$$

It follows that ν_1 defines a seminorm on $\mathcal{L}_1(E, \mathbb{K})$. But if $\varnothing \neq F \subseteq E$ is a set of measure zero (for example, $F = \{x_0\}$ for some point $x_0 \in E$), then $\chi_F \not\equiv 0$ and yet

$$\nu_1(\chi_F) = \int_E |\chi_F| = mF = 0.$$

In other words, $\nu_1(\cdot)$ does *not* define a norm on $\mathcal{L}_1(E, \mathbb{K})$.

6.3. Proposition. *Let $\mathcal{W}$ be a vector space over the field $\mathbb{K}$, and suppose that ν is a seminorm on $\mathcal{W}$. Let $\mathcal{N} := \{w \in W : \nu(w) = 0\}$. Then $\mathcal{N}$ is a linear manifold in $\mathcal{W}$ and so $\mathcal{W}/\mathcal{N}$ is a vector space over $\mathbb{K}$, whose elements we denote by $[x] := x + \mathcal{N}$. Furthermore, the map*

$$\begin{array}{rccc} \|\cdot\| : & \mathcal{W}/\mathcal{N} & \to & \mathbb{R} \\ & [x] & \mapsto & \nu(x) \end{array}$$

defines a norm on $\mathcal{W}/\mathcal{N}$.

Proof. Clearly $0 \in \mathcal{N}$ and thus $\mathcal{N} \neq \varnothing$. Suppose that $v, w \in \mathcal{N}$, and $k \in \mathbb{K}$. Then

$$\nu(kv + w) \leq \nu(kv) + \nu(w) = |k|\nu(v) + \nu(w) = 0,$$

and so $kv + w \in \mathcal{N}$, proving that $\mathcal{N}$ is a linear manifold in $\mathcal{W}$.

From elementary linear algebra theory, we know that $\mathcal{W}/\mathcal{N}$ is a vector space under the operations $[x] + [y] := [x + y]$ and $k[x] = [kx]$ for all $x, y \in \mathcal{W}$, $k \in \mathbb{K}$. We normally refer to this as the *quotient space* of $\mathcal{W}$ by $\mathcal{N}$.

To see that $\|\cdot\|$ defines a norm on $\mathcal{W}/\mathcal{N}$, we first check that this function is well-defined. Indeed, suppose that $[v] = [w]$ in $\mathcal{W}/\mathcal{N}$. Then $v - w \in \mathcal{N}$, and so $\nu(v - w) = 0$. But then $|\nu(v) - \nu(w)| \le \nu(v - w) = 0$ implies that $\nu(v) = \nu(w)$, and so $\|\cdot\|$ is well-defined, as claimed.

Now

- $\|[x]\| = \nu(x) \ge 0$ for all $[x] \in \mathcal{W}/\mathcal{N}$, and $\|[0]\| = \nu(0) = 0$.
- If $\|[x]\| = 0$, then $\nu(x) = 0$, so $x \in \mathcal{N}$ and therefore $[x] = [0]$.
- If $[x] \in \mathcal{W}/\mathcal{N}$ and $k \in \mathbb{K}$, then $\|k[x]\| = \|[kx]\| = \nu(kx) = |k|\nu(x) = |k|\,\|[x]\|$; and finally
- If $[x], [y] \in \mathcal{W}/\mathcal{N}$, then

$$\|[x] + [y]\| = \|[x + y]\| = \nu(x + y) \le \nu(x) + \nu(y) = \|[x]\| + \|[y]\|.$$

Thus $\|\cdot\|$ is a norm on $\mathcal{W}/\mathcal{N}$, which completes the proof.

□

6.4. Let us return to Example 6.2, where we determined that $\nu_1(\cdot)$ defines a seminorm on $\mathcal{L}_1(E, \mathbb{K})$.

Suppose that $g \in \mathcal{N}_1(E, \mathbb{K}) := \{f \in \mathcal{L}_1(E, \mathbb{K}) : \nu_1(f) = 0\}$. Then $\int_E |g| = 0$, and therefore $g = 0$ a.e. on E. Conversely, if $g = 0$ a.e. on E, then $\int_E |g| = 0$ and therefore $g \in \mathcal{N}_1(E, \mathbb{K})$. In other words,

$$\mathcal{N}_1(E, \mathbb{K}) = \{g \in \mathcal{L}_1(E, \mathbb{K}) : g = 0 \text{ a. e. on } E\}.$$

Thus $[g] = [h]$ in $\mathcal{L}_1(E, \mathbb{K})/\mathcal{N}_1(E, \mathbb{K})$ if and only if $g - h \in \mathcal{N}_1(E, \mathbb{K})$, which is to say that $g = h$ a.e. on E. By Proposition 6.3, the map $\|[f]\| := \nu_1(f)$ defines a norm on $L_1(E, \mathbb{K}) := \mathcal{L}_1(E, \mathbb{K})/\mathcal{N}_1(E, \mathbb{K})$.

6.5. Definition. *The space $L_1(E, \mathbb{K}) = \mathcal{L}_1(E, \mathbb{K})/\mathcal{N}_1(E, \mathbb{K})$ defined above is referred to as "L_1 of E", and it is a normed linear space.*

It is *crucial* to remember that the elements of $L_1(E, \mathbb{K})$ are *cosets* of $\mathcal{L}_1(E, \mathbb{K})$; that is to say, they are *equivalence classes* of functions which are equal a.e. on E. Given an element $[f]$ of $L_1(E, \mathbb{K})$, one can **not** speak of the value of the function at a point in E, since we are *not* dealing with functions!

Our next goal is to perform a similar construction on a family of spaces indexed by positive real numbers $1 < p < \infty$.

6.6. Definition. *Let $E \in \mathfrak{M}(\mathbb{R})$, so that E is Lebesgue measurable. Let $1 < p < \infty$ be a real number, and set*

$$\mathcal{L}_p(E, \mathbb{K}) := \{f \in \mathcal{L}(E, \mathbb{K}) : \int_E |f|^p < \infty\} = \{f \in \mathcal{L}(E, \mathbb{K}) : |f|^p \in \mathcal{L}_1(E, \mathbb{K})\}.$$

6.7. We would like to verify that $\mathcal{L}_p(E, \mathbb{K})$ is a vector space for all $1 < p < \infty$, and that $\nu_p(f) := \left(\int_E |f|^p\right)^{1/p}$ defines a seminorm on $\mathcal{L}_p(E, \mathbb{K})$. Then we can once again appeal to Proposition 6.3 to obtain a normed linear space as a quotient of $\mathcal{L}_p(E, \mathbb{K})$.

The proof of this is, however, somewhat technical, and will require a couple of auxiliary results.

6.8. Definition. *Let $1 \le p \le \infty$. We associate to p the number $1 \le q \le \infty$ as follows:*

- *If $p = 1$, we set $q = \infty$;*
- *if $1 < p < \infty$, then we set $q = \left(1 - \frac{1}{p}\right)^{-1}$; and*
- *if $p = \infty$, we set $q = 1$.*

We say that q is the ***Lebesgue conjugate*** *of p. With the convention that $1/\infty := 0$, we see that in all cases,*

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

When $1 < p < \infty$, we see that the above equation is equivalent to each of the equations

- $p(q-1) = q$ and
- $(p-1)q = p$.

While these are trivial algebraic manipulations, it will prove useful to keep them in mind in the proofs below.

6.9. Lemma. Young's Inequality. *Let $1 < p < \infty$ and denote by q the Lebesgue conjugate of p. Let $0 < a, b \in \mathbb{R}$.*

(a) *Then*

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

(b) *Equality holds in the above expression if and only if $a^p = b^q$.*

Proof. Consider the function $g : (0, \infty) \to \mathbb{R}$ given by $g(x) = \frac{1}{p}x^p + \frac{1}{q} - x$. Clearly g is differentiable on $(0, \infty)$ with $g'(x) = x^{p-1} - 1$. Thus g is clearly strictly decreasing on $(0, 1)$ and it is strictly increasing on $(1, \infty)$. Furthermore,

$$g(1) = \frac{1}{p} + \frac{1}{q} - 1 = 0.$$

In particular, $g(x) \ge 0$ for all $x \in (0, \infty)$, and $g(x) = 0$ if and only if $x = 1$.

(a) Set $x_0 := \frac{a}{b^{q-1}} > 0$. Then

$$0 \le g(x_0) = \frac{a^p}{p\, b^{(q-1)p}} + \frac{1}{q} - \frac{a}{b^{q-1}},$$

so that

$$\frac{a}{b^{q-1}} \le \frac{a^p}{pb^q} + \frac{1}{q}.$$

That is,

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

(b) From above, equality holds if and only if $g(x_0) = 0$, which happens if and only if $x_0 = 1$. But this is clearly equivalent to the condition that $a = b^{q-1}$; i.e. $a^p = b^{p(q-1)} = b^q$.

□

Recall from Proposition 4.10 that if $E \in \mathfrak{M}(\mathbb{R})$ and $f : E \to \mathbb{K}$ is a measurable function, then there exists a measurable function $u : E \to \mathbb{T}$ such that $f = u \cdot |f|$. Clearly, if $\mathbb{K} = \mathbb{R}$, then the range of u is contained in $\{-1, 1\}$. Let us denote by $\overline{u}$ the function $\overline{u}(x) = \overline{u(x)}$ for all $x \in E$.

6.10. Theorem. Hölder's Inequality. *Let $E \in \mathfrak{M}(\mathbb{R})$ and $1 < p < \infty$ be a real number, and let q denote the Lebesgue conjugate of p.*

(a) *If $f \in \mathcal{L}_p(E, \mathbb{K})$ and $g \in \mathcal{L}_q(E, \mathbb{K})$, then $fg \in \mathcal{L}_1(E, \mathbb{K})$ and*

$$\nu_1(fg) \le \nu_p(f)\ \nu_q(g),$$

where $\nu_p(f) = \left(\int_E |f|^p\right)^{1/p}$ and $\nu_q(g) = \left(\int_E |g|^q\right)^{1/q}$.

(b) *Suppose that $H := \{x \in E : f(x) \ne 0\}$ has positive measure. If*

$$f^* := \nu_p(f)^{1-p}\ \overline{u}\,|f|^{p-1},$$

then $f^* \in \mathcal{L}_q(E, \mathbb{K})$, $\nu_q(f^*) = 1$, and

$$\nu_1(ff^*) = \int_E ff^* = \nu_p(f).$$

Remark. In an unfortunate coincidence of terminology, we shall also refer to f^* as the **Lebesgue conjugate function** of f.

Proof.

We first observe that by Proposition 4.8, fg is measurable.

(a) Note that if $f = 0$ a.e or $g = 0$ a.e. on E, then $fg = 0$ a.e. on E and there is nothing to prove. It is easy to verify that given $0 < \alpha, \beta \in \mathbb{K}$, $\alpha f \in \mathcal{L}_p(E, \mathbb{K})$ and $\beta g \in \mathcal{L}_q(E, \mathbb{K})$. Suppose that we can find $\alpha_0, \beta_0 \ne 0$ such that

$$\int_E |(\alpha_0 f)(\beta_0 g)| \le \nu_p(\alpha_0 f)\nu_q(\beta_0 g).$$

By dividing both sides of this inequality by $|\alpha_0|\,|\beta_0|$, we clearly have that

$$\int_E |fg| \le \nu_p(f)\nu_q(g).$$

This shows that, by choosing $\alpha_0 = \nu_p(f)^{-1}$ and $\beta_0 = \nu_q(g)^{-1}$, we may assume without loss of generality that $\nu_p(f) = 1 = \nu_q(g)$.

Now, by Young's Inequality, Lemma 6.9, we have that

$$|fg| \le \frac{|f|^p}{p} + \frac{|g|^q}{q},$$

and so

$$\begin{aligned}
\nu_1(fg) &= \int_E |fg| \\
&\le \frac{1}{p} \int_E |f|^p + \frac{1}{q} \int_E |g|^q \\
&= \frac{1}{p} \nu_p(f)^p + \frac{1}{q} \nu_q(g)^q \\
&= \frac{1}{p} + \frac{1}{q} \\
&= 1 \\
&= \nu_p(f)\, \nu_q(g).
\end{aligned}$$

This completes the first half of the proof.

(b) Now let f^* be defined as above. Then f^* is measurable, being the product of measurable functions, and recalling that $(p-1)q = p$, we have

$$\begin{aligned}
\nu_q(f^*)^q &= \int_E |f^*|^q \\
&= \int_E \left(\nu_p(f)^{1-p} |f|^{p-1}\right)^q \\
&= \nu_p(f)^{(1-p)q} \int_E |f|^{(p-1)q} \\
&= \nu_p(f)^{-p} \int_E |f|^p \\
&= \nu_p(f)^{-p}\ \nu_p(f)^p \\
&= 1.
\end{aligned}$$

Finally,

$$\nu_1(ff^*) = \int_E |ff^*| = \int_E ff^* = \nu_p(f)^{1-p} \nu_p(f)^p = \nu_p(f).$$

□

6.11. Theorem. Minkowski's Inequality.

Let $E \in \mathfrak{M}(\mathbb{R})$ be a measurable set and $1 < p < \infty$. If $f, g \in \mathcal{L}_p(E, \mathbb{K})$, then $f + g \in \mathcal{L}_p(E, \mathbb{K})$ and

$$\nu_p(f+g) \le \nu_p(f) + \nu_p(g).$$

Proof. That $f + g$ is measurable is clear, as each of f and g is. Observe that for any $0 \le a, b$ we have that

$$(a+b)^p \le (2 \max(a,b))^p \le 2^p (a^p + b^p).$$

Thus

$$|f+g|^p \le (|f|+|g|)^p \le 2^p(|f|^p+|g|^p).$$

It follows that

$$\nu_p(f+g)^p = \int_E |f+g|^p \le 2^p(\nu_p(f)^p + \nu_p(g)^p) < \infty,$$

and therefore $f+g \in \mathcal{L}_p(E,\mathbb{K})$. Set $h = f+g$, and let h^* denote the conjugate function of h. Then $h^* \in \mathcal{L}_q(E,\mathbb{K})$, $\nu_q(h) = 1$ and $\nu_1(h\cdot h^*) = \nu_p(h)$.

From this and Hölder's Inequality we deduce that

$$\begin{aligned}\nu_p(f+g) &= \nu_p(h)\\ &= \nu_1(h\cdot h^*)\\ &= \nu_1((f+g)h^*)\\ &\le \nu_1(f\cdot h^*) + \nu_1(g\cdot h^*)\\ &\le \nu_p(f)\nu_q(h^*) + \nu_p(g)\nu_q(h^*)\\ &= \nu_p(f)+\nu_p(g).\end{aligned}$$

□

6.12. Corollary. *Let $E \in \mathfrak{M}(\mathbb{R})$, and $1 < p < \infty$. Then $\mathcal{L}_p(E,\mathbb{K})$ is a vector space, and $\nu_p(\cdot)$ defines a seminorm on $\mathcal{L}_p(E,\mathbb{K})$.*

Proof. Clearly $\mathcal{L}_p(E,\mathbb{K}) \subseteq \mathcal{L}(E,\mathbb{K})$ by definition, and since the latter is a vector space, it suffices that we prove that $\mathcal{L}_p(E,\mathbb{K}) \ne \varnothing$, and that $f,g \in \mathcal{L}_p(E,\mathbb{K})$ and $k \in \mathbb{K}$ implies that $kf+g \in \mathcal{L}_p(E,\mathbb{K})$.

Let $\zeta : E \to \mathbb{K}$ be the zero function $\zeta(x) = 0$ for all $x \in E$. Clearly $\zeta \in \mathcal{L}_p(E,\mathbb{K})$ and hence $\mathcal{L}_p(E,\mathbb{K}) \ne \varnothing$.

If $f \in \mathcal{L}_p(E,\mathbb{K})$ and $k \in \mathbb{K}$, then kf is measurable by Proposition 4.8 and

$$\int_E |kf|^p = |k|^p \int_E |f|^p < \infty,$$

so that $kf \in \mathcal{L}_p(E,\mathbb{K})$. If $g \in \mathcal{L}_p(E,\mathbb{K})$ as well, then by Minkowski's inequality, $kf+g \in \mathcal{L}_p(E,\mathbb{K})$. Thus $\mathcal{L}_p(E,\mathbb{K})$ is a vector space.

With $f,g \in \mathcal{L}_p(E,\mathbb{K})$ and $k \in \mathbb{K}$ as above,

- $\nu_p(f) = \int_E |f|^p \ge 0$, and $\nu_p(\zeta) = \int_E \zeta^p = \int_E 0 = 0$.
- $$\nu_p(kf) = \left(\int_E |kf|^p\right)^{1/p} = \left(|k|^p \int_E |f|^p\right)^{1/p} = |k|\nu_p(f).$$
- $\nu_p(f+g) \le \nu_p(f)+\nu_p(g)$ by Minkowski's inequality.

Thus $\nu_p(\cdot)$ defines a seminorm on $\mathcal{L}_p(E,\mathbb{K})$. As was the case with $\mathcal{L}_1(E,\mathbb{K})$, if $\varnothing \ne F \subseteq E$ is a subset of Lebsegue measure zero, then $0 \ne \chi_F$ is measurable and

$$\nu_p(\chi_F) = \left(\int_E |\chi_F|^p\right)^{1/p} = \int_E \chi_F = m(F\cap E) = 0.$$

Thus $\nu_p(\cdot)$ is not a norm.

□

Once again, we shall appeal to Proposition 6.3 to obtain a normed linear space from a semi-normed linear space.

6.13. Definition. *Let $E \in \mathfrak{M}(\mathbb{R})$ and $1 < p < \infty$. We define the L_p-**space***

$$L_p(E,\mathbb{K}) := \mathcal{L}_p(E,\mathbb{K})/\mathcal{N}_p(E,\mathbb{K}),$$

*where $\mathcal{N}_p(E,\mathbb{K}) = \{f \in \mathcal{L}_p(E,\mathbb{K}) : \nu_p(f) = 0\}$. The L_p-**norm** on $L_p(E,\mathbb{K})$ is the norm defined by*

$$\begin{array}{rccc} \|\cdot\|_p : & L_p(E,\mathbb{K}) & \to & \mathbb{R} \\ & [f] & \mapsto & \nu_p(f) \end{array}.$$

6.14. Remark. Consider $f \in \mathcal{N}_p(E,\mathbb{K})$, so that f is measurable and $\int_E |f|^p = 0$. It follows that $|f|^p = 0$ a.e. on E, and hence that $f = 0$ a.e. on E. Conversely, if f is measurable and $f = 0$ a.e., so that $|f|^p = 0$ a.e. on E as well. But then $f \in \mathcal{N}_p(E,\mathbb{K})$.

In other words, $\mathcal{N}_p(E,\mathbb{K}) = \{f : E \to \mathbb{K} : f = 0 \text{ a. e. on } E\}$ for all $1 < p < \infty$, and $[f] = [g]$ in $L_p(E,\mathbb{K})$ if and only if $f, g \in \mathcal{L}_p(E,\mathbb{K})$ and $f = g$ a.e. on E.

For the sake of completeness, and in keeping with the literature, let us restate the Hölder and Minkowski Inequalities for $L_p(E,\mathbb{K})$.

6.15. Theorem. Hölder's Inequality.
Let $E \in \mathfrak{M}(\mathbb{R})$ and $1 < p < \infty$. Let q denote the Lebesgue conjugate of p.

(a) *If $[f] \in L_p(E,\mathbb{K})$ and $[g] \in L_q(E,\mathbb{K})$, then $[f][g] := [fg] \in L_1(E,\mathbb{K})$ is well-defined and*

$$\|[fg]\|_1 \le \|[f]\|_p \ \|[g]\|_q.$$

(b) *If $0 \ne [f] \in L_p(E,\mathbb{K})$ and f^* is the conjugate function of f, then $[f^*] \in L_q(E,\mathbb{K})$, $\|[f^*]\|_q = 1$, and*

$$\|[f][f^*]\|_1 = \|[f]\|_p.$$

Proof. The only part that does not follow immediately from Theorem 6.10 is the well-definedness of the operation $[f][g] = [fg]$, and this is left as an exercise for the reader.

□

6.16. Theorem. Minkowski's Inequality.
Let $E \in \mathfrak{M}(\mathbb{R})$ be a measurable set and $1 < p < \infty$. If $[f], [g] \in L_p(E,\mathbb{K})$, then $[f+g] \in L_p(E,\mathbb{K})$ and

$$\|[f] + [g]\|_p \le \|[f]\|_p + \|[g]\|_p.$$

We are now in position to prove the completeness of $L_p(E,\mathbb{K})$ for all $1 \le p < \infty$.

6.17. Theorem. *Let $E \in \mathfrak{M}(\mathbb{R})$ and $1 \le p < \infty$. Then $L_p(E, \mathbb{K})$ is a Banach space.*

Proof.

As we have already noted, that $L_p(E, \mathbb{K})$ is a normed linear space follows from Proposition 6.3. There remains to show that it is complete. Recall from the Assignments that a normed linear space is complete if and only if every absolutely summable series is summable. Our proof will appeal to this result.

Suppose that $([f_n])_{n=1}^\infty$ is a sequence in $L_p(E, \mathbb{K})$ and suppose furthermore that $\gamma := \sum_{n=1}^\infty \|[f_n]\|_p < \infty$. Our strategy will be to use the representatives f_n, $n \ge 1$ of the elements $[f_n] \in L_p(E, \mathbb{K})$ to produce a measurable function $h \in \mathcal{L}_p(E, \mathbb{K})$ such that $h(x) = \sum_{n=1}^\infty f_n(x)$ *almost everywhere* on E. Then we shall show that in fact, $[h] = \sum_{n=1}^\infty [f_n]$ in the sense of norm convergence in $L_p(E, \mathbb{K})$.

STEP ONE. First we must show that $\sum_{n=1}^\infty f_n(x)$ converges *almost everywhere* on E. To that end, for each $N \ge 1$, we set $g_N := \sum_{n=1}^N |f_n|$, and observe that $g_N \in \mathcal{L}_p(E, \mathbb{R})$ by Corollary 6.12. Observe also that

$$0 \le g_1 \le g_2 \le g_3 \le \cdots.$$

Set $g_\infty = \sup_{N \ge 1} g_N$, so that $0 \le g_\infty \in \mathcal{L}(E, [0, \infty])$. Then $g_\infty^p = \sup_{N \ge 1} g_N^p$, and so by the Monotone Convergence Theorem 5.16,

$$\begin{aligned}
\int_E g_\infty^p &= \lim_{N\to\infty} \int_E g_N^p \\
&= \lim_{N\to\infty} \int_E (|f_1| + |f_2| + \cdots + |f_N|)^p \\
&= \lim_{N\to\infty} (\nu_p(|f_1| + |f_2| + \cdots + |f_N|))^p \\
&\le \lim_{N\to\infty} (\nu_p(f_1) + \nu_p(f_2) + \cdots + \nu_p(f_N))^p \\
&= \lim_{N\to\infty} (\|[f_1]\|_p + \|[f_2]\|_p + \cdots + \|[f_N]\|_p)^p \\
&\le \gamma^p < \infty.
\end{aligned}$$

From this it follows that $B := \{x \in E : g_\infty(x) = \infty\}$ has measure zero. Let $H := E \setminus B$ and $g := \chi_H \cdot g_\infty$. Then $g \in \mathcal{L}(E, [0, \infty))$, and $g = g_\infty$ a.e. on E. But then

$$\int_E g^p = \int_E g_\infty^p \le \gamma^p,$$

and so $g \in \mathcal{L}_p(E, \mathbb{K})$ – i.e., $[g] \in L_p(E, \mathbb{K})$ – and $\|[g]\|_p \le \gamma$. Next, note that for $x \in H$,

$$\left| \sum_{n=1}^\infty f_n(x) \right| \le \sum_{n=1}^\infty |f_n(x)| = g_\infty(x) = g(x) < \infty,$$

and so $\sum_{n=1}^\infty f_n(x) \in \mathbb{K}$ exists, by the completeness of $\mathbb{K}$.

STEP TWO. For each $N \geq 1$, set $h_N := \chi_H \cdot (\sum_{n=1}^N f_n)$, so that $h_N \in \mathcal{L}_p(E,\mathbb{K}) \subseteq \mathcal{L}(E,\mathbb{K})$, and $[h_N] = \sum_{n=1}^N [f_n]$. Furthermore, for $x \in H$,

$$|h_N(x)| \leq \sum_{n=1}^N |f_n(x)| \leq g(x),$$

while for $x \in B$, $|h_N(x)| = 0 = g(x)$. Thus $|h_N| \leq g$ on E, and as a trivial consequence, $|h_N|^p \leq g^p$ on E. From this we conclude that for each $N \geq 1$,

$$\int_E |h_N|^p \leq \int_E g^p \leq \gamma^p.$$

Define $h(x) := \lim_{N\to\infty} h_N(x) \in \mathbb{K}$, $x \in E$. Note that h is measurable, being the limit of measurable functions. From above, $|h| \leq g$ on E, and thus

$$\int_E |h|^p \leq \int_E g^p \leq \gamma^p < \infty.$$

It follows that $h \in \mathcal{L}_p(E,\mathbb{K})$; i.e. $[h] \in L_p(E,\mathbb{K})$.

STEP THREE. Recall that we are trying to prove that $\sum_{n=1}^\infty [f_n]$ converges in $L_p(E,\mathbb{K})$. We chose a specific set of representatives, namely the f_n's themselves, and we showed that there exists an element $h \in \mathcal{L}_p(E,\mathbb{K})$ such that *almost everywhere* on E, namely on $H \subseteq E$, $h = \sum_{n=1}^\infty f_n$ as a *pointwise limit* of functions. If we can show that

$$\lim_{N\to\infty} \|[h] - [h_N]\|_p = \lim_{N\to\infty} \|[h] - \sum_{n=1}^N [f_n]\|_p = 0,$$

then obviously we are done.

Considering that $|h_M - h_N|^p \leq (|h_M| + |h_N|)^p \leq (g+g)^p$ for all M and N and that $\int_E (2|g|)^p < \infty$ by virtue of $0 \leq g$ being in $\mathcal{L}_p(E,\mathbb{K})$, an application of the Lebesgue

Dominated Convergence Theorem 5.33 shows that

$$\begin{aligned}\|[h]-[h_N]\|_p &= \nu_p(h-h_N)\\ &= \left(\int_E |h-h_N|^p\right)^{1/p}\\ &= \left(\int_E \lim_{M\to\infty} |h_M-h_N|^p\right)^{1/p}\\ &= \left(\lim_{M\to\infty}\int_E |h_M-h_N|^p\right)^{1/p}\\ &= \lim_{M\to\infty}\left(\int_E |h_M-h_N|^p\right)^{1/p}\\ &= \lim_{M\to\infty}\|[h_M]-[h_N]\|_p\\ &= \lim_{M\to\infty}\left\|\sum_{n=N+1}^{M}[f_n]\right\|_p\\ &\le \lim_{M\to\infty}\sum_{n=N+1}^{M}\|[f_n]\|_p\\ &= \sum_{n=N+1}^{\infty}\|[f_n]\|_p.\end{aligned}$$

Since $\sum_{n=1}^{\infty}\|[f_n]\|_p = \gamma < \infty$ by hypothesis, it is clear that

$$\lim_{N\to\infty}\|[h]-[h_N]\|_p \le \lim_{N\to\infty}\sum_{n=N+1}^{\infty}\|[f_n]\|_p = 0.$$

As noted above, this completes the proof.

□

The case of $p=\infty$.

6.18. Definition. *Let $E\in\mathfrak{M}(\mathbb{R})$ and suppose that $f\in\mathcal{L}(E,\mathbb{K})$. We define the* ***essential supremum*** *of f to be*

$$\nu_\infty(f) = \operatorname{ess\,sup}(f) := \inf\{\gamma>0 : m\{x\in E : |f(x)|>\gamma\}=0\}.$$

Set $\mathcal{L}_\infty(E,\mathbb{K}) = \{f\in\mathcal{L}(E,\mathbb{K}) : \nu_\infty(f)<\infty\}$.

6.19. Examples.

(a) Let $E=\mathbb{R}$ and $f=\chi_{\mathbb{Q}}$ be the characteristic function of the rationals. For any $\gamma>0$, $m\{x\in\mathbb{R} : |\chi_{\mathbb{Q}}(x)|>\gamma\}\le m\mathbb{Q}=0$, and so $\nu_\infty(\chi_{\mathbb{Q}})=0$.

Clearly there was nothing special about $\mathbb{Q}$ in this example, other than the fact that this set has Lebesgue measure zero.

(b) Suppose that $a<b\in\mathbb{R}$ and that $f\in\mathcal{C}([a,b],\mathbb{K})$.

We claim that $f\in\mathcal{L}_\infty([a,b],\mathbb{K})$ and that $\nu_\infty(f)=\|f\|_{\sup}$. Indeed, every continuous function is measurable, so $f\in\mathcal{L}([a,b],\mathbb{K})$. Also, $[a,b]$ being a compact set, there exists $x_0\in[a,b]$ such that $|f(x_0)|=\|f\|_{\sup}$. By continuity of f, given $\varepsilon > 0$, there exists $\delta > 0$ such that $x \in [a,b]$ and $|x - x_0| < \delta$ implies $|f(x)| > |f(x_0)| - \varepsilon$. Whether or not x_0 is one of the endpoints of the interval, it follows that there exist $c < d \in (a,b)$ such that $x \in [c,d]$ implies that $|f(x)| > |f(x_0)| - \varepsilon$. But then

$$m\{x \in [a,b] : |f(x)| > |f(x_0)| - \varepsilon\} \geq d - c > 0,$$

and so $\nu_\infty(f) \geq \|f\|_{\text{sup}} - \varepsilon$. Since $\varepsilon > 0$ was arbitrary, $\nu_\infty(f) \geq \|f\|_{\text{sup}}$.

If $\gamma > \|f\|_{\text{sup}}$, then $m\{x \in E : |f(x)| > \gamma\} = m\varnothing = 0$, and thus $\nu_\infty(f) \leq \gamma$. Hence $\nu_\infty(f) \leq \|f\|_{\text{sup}}$. Combining these inequalities,

$$\nu_\infty(f) = \|f\|_{\text{sup}}.$$

In particular, $f \in \mathcal{L}_\infty([a,b],\mathbb{K})$.

6.20. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$. Then $\mathcal{L}_\infty(E,\mathbb{K})$ is a vector space over $\mathbb{K}$, and that $\nu_\infty(\cdot)$ is a seminorm on $\mathcal{L}_\infty(E,\mathbb{K})$.*

Proof. Since $\mathcal{L}_\infty(E,\mathbb{K}) \subseteq \mathcal{L}(E,\mathbb{K})$, and since the latter is a vector space, we need only apply the Subspace Test from linear algebra to conclude that $\mathcal{L}_\infty(E,\mathbb{K})$ is a vector space.

It is clear from the definition that $\nu_\infty(f) \geq 0$ for all $f \in \mathcal{L}_\infty(E,\mathbb{K})$, and clearly the constant function $\zeta(x) = 0$, $x \in E$ lies in $\mathcal{L}_\infty(E,\mathbb{K})$, with $\nu_\infty(\zeta) = 0$. Thus $\mathcal{L}_\infty(E,\mathbb{K}) \neq \varnothing$.

Suppose next that $f, g \in \mathcal{L}_\infty(E,\mathbb{K})$, and that $0 \neq k \in \mathbb{K}$. Then $kf \in \mathcal{L}(E,\mathbb{K})$, and a moment's thought should convince the reader that this implies that

$$\begin{aligned}
\nu_\infty(kf) &= \inf\{\gamma > 0 : m\{x \in E : |kf(x)| > \gamma\} = 0\} \\
&= \inf\{|k|\,\delta > 0 : m\{x \in E : |kf(x)| > |k|\,\delta\} = 0\} \\
&= |k| \ \inf\{\delta > 0 : m\{x \in E : |f(x)| > \delta\} = 0\} \\
&= |k| \ \nu_\infty(f) < \infty.
\end{aligned}$$

Hence $kf \in \mathcal{L}_\infty(E,\mathbb{K})$ for all $0 \neq k \in \mathbb{K}$. Since $0\,\nu_\infty(f) = 0 = \nu_\infty(\zeta) = \nu_\infty(0 \cdot f)$ for all $f \in \mathcal{L}_\infty(E,\mathbb{K})$, we see that $kf \in \mathcal{L}_\infty(E,\mathbb{K})$ and $\nu_\infty(kf) = |k|\nu_\infty(f)$ for all $k \in \mathbb{K}$.

Finally, suppose that $f, g \in \mathcal{L}_\infty(E,\mathbb{K})$ and let $\alpha > \nu_\infty(f)$, $\beta > \nu_\infty(g)$ be arbitrary. Let $E_f := \{x \in E : |f(x)| > \alpha\}$ and $E_g := \{x \in E : |g(x)| > \beta\}$, so that $mE_f = 0 = mE_g$. If $x \in H := E \setminus (E_f \cup E_g)$, then $|(f+g)(x)| \leq |f(x)| + |g(x)| \leq \alpha + \beta$, and so

$$\{x \in E : |(f+g)(x)| > \alpha + \beta\} \subseteq E_f \cup E_g.$$

From this it immediately follows that

$$m\{x \in E : |(f+g)(x)| > \alpha + \beta\} \leq mE_f + mE_g = 0,$$

whence $\nu_\infty(f+g) \leq \alpha + \beta$. Since $\alpha > \nu_\infty(f)$ and $\beta > \nu_\infty(g)$ were arbitrary,

$$\nu_\infty(f+g) \leq \nu_\infty(f) + \nu_\infty(g).$$

This completes the proof.

□

6.21. Definition. *As we did with the previous L_p-spaces, $1 \le p < \infty$, we define the subspace*

$$\mathcal{N}_\infty(E, \mathbb{K}) := \{f \in \mathcal{L}_\infty(E, \mathbb{K}) : \nu_\infty(f) = 0\}$$

of $\mathcal{L}_\infty(E, \mathbb{K})$, and we denote by $[f]$ the equivalence class in

$$L_\infty(E, \mathbb{K}) := \mathcal{L}_\infty(E, \mathbb{K}) / \mathcal{N}_\infty(E, \mathbb{K})$$

of $f \in \mathcal{L}_\infty(E, \mathbb{K})$.

Yet again we may appeal to Proposition 6.3 to conclude the following:

6.22. Theorem. *Let $E \in \mathfrak{M}(\mathbb{R})$. Then $L_\infty(E, \mathbb{K})$ is a normed-linear space, where for $[f] \in L_\infty(E, \mathbb{K})$, we set*

$$\|[f]\|_\infty := \nu_\infty(f).$$

6.23. Let $f \in \mathcal{L}_\infty(E, \mathbb{K})$. If we define $F_n := \{x \in E : |f(x)| > \nu_\infty(f) + \frac{1}{n}\}$, then $mF_n = 0$ for all $n \ge 1$, and $F := \cup_{n=1}^\infty F_n = \{x \in E : |f(x)| > \nu_\infty(f)\}$ satisfies

$$mF \le \sum_{n=1}^\infty mF_n = \sum_{n=1}^\infty 0 = 0.$$

From this it follows that given $[f] \in L_\infty(E, \mathbb{K})$, we may always choose a representative $g \in [f]$ in such a way as to guarantee that $|g(x)| \le \|[f]\|_\infty$ for all $x \in E$. Indeed, given $f \in \mathcal{L}_\infty(E, \mathbb{K})$, then as noted above, the set $F = \{x \in E : |f(x)| > \nu_\infty(f)\}$ has measure zero. As such, the function $g := \chi_{E \setminus F} \cdot f$ is measurable, and differs from f only on F, whence $[g] = [f]$, and clearly $|g(x)| \le \nu_\infty(f) = \nu_\infty(g) = \|[g]\|_\infty$ for all $x \in E$.

Moreover, it is readily seen that $\nu_\infty(f) = 0$ if and only if $f = 0$ almost everywhere on E, and thus

$$\mathcal{N}_\infty(E, \mathbb{K}) = \{f \in \mathcal{L}(E, \mathbb{K}) : f = 0 \text{ a.e. on } E\}.$$

There remains to show that $L_\infty(E, \mathbb{K})$ is complete.

6.24. Theorem. *Let $E \in \mathfrak{M}(\mathbb{R})$. Then $L_\infty(E, \mathbb{K})$ is a Banach space.*

Proof. Let $([f_n])_{n=1}^\infty$ be an absolutely summable sequence in $L_\infty(E, \mathbb{K})$, and set $\gamma := \sum_{n=1}^\infty \|[f_n]\|_\infty < \infty$. By our work in the Assignments, we see that it suffices to prove that $\lim_{N \to \infty} \sum_{n=1}^N [f_n] \in L_\infty(E, \mathbb{K})$ exists.

By the argument of the above paragraph, we may assume without loss of generality that for each $x \in E$ and each $n \ge 1$, $|f_n(x)| \le \nu_\infty(f_n) = \|[f_n]\|_\infty$, and thus

$$\sum_{n=1}^\infty |f_n(x)| \le \sum_{n=1}^\infty \|[f_n]\|_\infty = \gamma.$$

As such, for each $x \in E$, the series $\sum_{n=1}^\infty f_n(x)$ is absolutely summable, and hence summable by the completeness of $\mathbb{K}$. Define a function $f : E \to \mathbb{K}$ via

$$f(x) := \sum_{n=1}^\infty f_n(x), \quad x \in E.$$

For each $N \geq 1$, set $h_N = \sum_{n=1}^{N} f_N \in \mathcal{L}_\infty(E, \mathbb{K})$. Since f is the pointwise limit of the measurable functions h_N, we see that $f \in \mathcal{L}(E, \mathbb{K})$. Moreover, the above estimate shows that $\nu_\infty(f) \leq \gamma < \infty$, and thus $f \in \mathcal{L}_\infty(E, \mathbb{K})$.

Let $\varepsilon > 0$ and choose $N_0 > 0$ such that $\sum_{n=N_0+1}^{\infty} \|[f_n]\|_\infty < \varepsilon$. For all $N > N_0$, we have that

$$\|[f] - [h_N]\|_\infty = \nu_\infty(f - h_N).$$

But for $x \in E$,

$$\begin{aligned}
|f(x) - h_N(x)| &= |\sum_{n=N+1}^{\infty} f_n(x)| \\
&\leq \sum_{n=N+1}^{\infty} |f_n(x)| \\
&\leq \sum_{n=N+1}^{\infty} \|[f_n]\|_\infty \\
&< \varepsilon.
\end{aligned}$$

Thus $\|[f] - [h_N]\|_\infty < \varepsilon$, $N \geq N_0$.

This shows that $[f] = \lim_{N\to\infty}[h_N] = \lim_{N\to\infty} \sum_{n=1}^{N} [f_n]$, as required.

□

6.25. Recall that if $E \in \mathfrak{M}(\mathbb{R})$, $1 < p < \infty$, $f \in \mathcal{L}_p(E, \mathbb{K})$ and $g \in \mathcal{L}_q(E, \mathbb{K})$, where q is the Lebesgue conjugate of p, then Hölder's Inequality (Theorem 6.10) states that $fg \in \mathcal{L}_1(E, \mathbb{K})$ and that

$$\nu_1(fg) \leq \nu_p(f)\,\nu_q(g).$$

Let us obtain a version of this inequality for $p = 1$, which will prove especially useful when we examine Fourier series in later chapters.

6.26. Theorem. Hölder's Inequality. *Let $E \in \mathfrak{M}(\mathbb{R})$ with $mE > 0$.*

(a) *If $f \in \mathcal{L}_1(E, \mathbb{K})$ and $g \in \mathcal{L}_\infty(E, \mathbb{K})$, then $fg \in \mathcal{L}_1(E, \mathbb{K})$ and*

$$\nu_1(fg) \leq \nu_1(f)\ \nu_\infty(g).$$

(b) *There exists a function $f^* \in \mathcal{L}_\infty(E, \mathbb{K})$ such that $\nu_\infty(f^*) = 1$ and*

$$\nu_1(f \cdot f^*) = \int_E f \cdot f^* = \nu_1(f).$$

Proof.

(a) By the comments of Paragraph 6.23, we know that we may find $g_0 \in \mathcal{L}_\infty(E, \mathbb{K})$ such that $g_0 = g$ a.e. on E and $|g_0(x)| \leq \nu_\infty(g) = \nu_\infty(g_0)$ for all $x \in E$. Since $g = g_0$ a.e. on E implies that $|fg| = |fg_0|$ a.e. on E, we find that

$$\int_E |fg| = \int_E |fg_0|.$$

Because of this, we may assume without loss of generality (by replacing g by g_0 if necessary) that $|g(x)| \leq \nu_\infty(g)$ for all $x \in E$.

Note that by Proposition 4.8, fg is measurable. But then

$$\nu_1(fg) = \int_E |fg| \le \int_E |f|\,\nu_\infty(g) = \nu_\infty(g) \int_E |f| = \nu_1(f)\,\nu_\infty(g).$$

(b) Arguing as in Proposition 4.10, we may find a measurable function $u : E \to \mathbb{T}$ such that

$$f = u \cdot |f|.$$

But then with $f^* := \overline{u}$ (i.e. $f^*(x) = \overline{u(x)}$ for all $x \in E$), $\nu_\infty(f^*) = 1$, $|f| = f \cdot f^*$, and so

$$\nu_1(f\,f^*) = \int_E |f \cdot f^*| = \int_E f \cdot f^* = \int_E |f| = \nu_1(f).$$

□

Once again, we shall refer to f^* as the **Lebesgue conjugate function** of f. As before, we obtain an immediate corollary.

6.27. Corollary. *Let $E \in \mathfrak{M}(\mathbb{R})$. If $[f] \in L_1(E,\mathbb{K})$ and $[g] \in L_\infty(E,\mathbb{K})$, then $[f][g] := [fg] \in L_1(E,\mathbb{K})$ is well-defined and*

$$\|[fg]\|_1 \le \|[f]\|_1 \;\|[g]\|_\infty.$$

6.28. Corollary. *Suppose that $a < b \in \mathbb{R}$. Consider $h \in \mathcal{C}([a,b],\mathbb{K})$ and $f \in \mathcal{L}_1([a,b],\mathbb{K})$.*

Then $h \cdot f \in \mathcal{L}_1([a,b],\mathbb{K})$ and

$$\nu_1(h \cdot f) \le \nu_1(f)\,\nu_\infty(h) = \nu_1(f)\,\|h\|_{\sup}.$$

Proof. Clearly h is measurable and $h \in \mathcal{L}_\infty([a,b],\mathbb{K})$ with $\|h\|_{\sup} = \nu_\infty(h)$, by Example 6.19.

The result now follows from Hölder's Inequality, Theorem 6.26 above.

□

6.29. It is interesting to consider the relationship between L_p-spaces for differing values of p. In the section below, we consider the case where the underlying set E has finite measure; the case where the measure of E is infinite is left to the exercises.

Let $E \in \mathfrak{M}(\mathbb{R})$ and suppose that $mE < \infty$. Let $[f] \in L_\infty(E,\mathbb{K})$. Then $f \in \mathcal{L}(E,\mathbb{K})$ and $|f(x)| \le \|[f]\|_\infty$ almost everywhere on E. For $1 \le p < \infty$,

$$\int_E |f|^p \le \int_E \|[f]\|_\infty^p = \|[f]\|_\infty^p \; mE < \infty,$$

proving that $[f] \in L_p(E,\mathbb{K})$, with

$$\|[f]\|_p \le \|[f]\|_\infty (mE)^{1/p}.$$

Thus $L_\infty(E,\mathbb{K}) \subseteq L_p(E,\mathbb{K})$, $1 \le p < \infty$ when $mE < \infty$.

Next, suppose that $1 \le p < r < \infty$, and that $[g] \in L_r(E,\mathbb{K})$. Again, $g \in \mathcal{L}(E,\mathbb{K})$ and

$$\int_E |g|^p = \int_E (|g|^r)^{p/r} \le \int_E \max(1,|g|^r) \le \int_E (1 + |g|^r) = mE + \|[g]\|_r^r < \infty.$$

It follows that $[g] \in L_p(E,\mathbb{K})$; i.e. $L_r(E,\mathbb{K}) \subseteq L_p(E,\mathbb{K})$.

6.30. Next, suppose that $a < b \in \mathbb{R}$. It follows from Example 6.19 that

$$[\mathcal{C}([a,b],\mathbb{K})] := \{[f] : f \in \mathcal{C}([a,b],\mathbb{K})\} \subseteq L_\infty([a,b]).$$

Recall that

$$\mathcal{R}_\infty([a,b],\mathbb{K}) = \{f : [a,b] \to \mathbb{K} : f \text{ is Riemann-integrable and bounded}\}.$$

As we saw in Corollary 5.25, $f \in \mathcal{L}([a,b],\mathbb{K})$ and thus $[f] \in L_\infty([a,b],\mathbb{K})$, by virtue of its being bounded.

Our next major goal is to prove that the space $[\mathcal{C}([a,b],\mathbb{K})]$ is dense in each of the spaces $L_p([a,b),\mathbb{K})$, $1 \le p < \infty$. We shall accomplish this through a series of approximations.

6.31. Before proving our next result, we introduce a bit of notation. Given $E \in \mathfrak{M}(\mathbb{R})$ and $1 \le p \le \infty$, we set

$$\textsc{Simp}_p(E,\mathbb{K}) = \textsc{Simp}(E,\mathbb{K}) \cap \mathcal{L}_p(E,\mathbb{K}).$$

Since $\textsc{Simp}(E,\mathbb{K})$ and $\mathcal{L}_p(E,\mathbb{K})$ are vector spaces over $\mathbb{K}$, so is $\textsc{Simp}_p(E,\mathbb{K})$.

We leave it to the exercises for the reader to prove that if $mE < \infty$ or if $p = \infty$, then $\textsc{Simp}_p(E,\mathbb{K}) = \textsc{Simp}(E,\mathbb{K})$.

6.32. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$ be a Lebesgue measurable set and $1 \le p < \infty$. Then $[\textsc{Simp}_p(E,\mathbb{K})] := \{[\varphi] : \varphi \in \textsc{Simp}_p(E,\mathbb{K})\}$ is dense in $(L_p(E,\mathbb{K}), \|\cdot\|_p)$.*

Proof.

Step One. Suppose that $f \in \mathcal{L}_p(E,[0,\infty))$, and let $\varepsilon > 0$. From our Assignment Questions, we know that we may find functions $\varphi_n \in \textsc{Simp}(E,[0,\infty))$, $n \ge 1$ such that

$$0 \le \varphi_1 \le \varphi_2 \le \cdots \le f,$$

and $f(x) = \lim_{n\to\infty} \varphi_n(x)$, $x \in E$.

Since

$$\int_E |\varphi_n|^p \le \int_E |f|^p < \infty,$$

we see that $\varphi_n \in \textsc{Simp}_p(E,[0,\infty))$, $n \ge 1$.

Moreover,

$$|f - \varphi_n|^p \le |f|^p, \quad n \ge 1,$$

and so by the Lebesgue Dominated Convergence Theorem 5.33,

$$\lim_{n\to\infty} \nu_p(f-\varphi_n)^p = \lim_{n\to\infty} \int_E |f-\varphi_n|^p = 0.$$

From this we clearly deduce that given $\varepsilon > 0$, we can find $\varphi \in \textsc{Simp}_p(E,[0,\infty))$ such that $\nu_p(f-\varphi) < \varepsilon$.

Step Two.

Now let $[g] \in L_p(E,\mathbb{K})$ be arbitrary, and note that $g \in \mathcal{L}_p(E,\mathbb{K})$. Recall that we may then write

$$g = (g_1 - g_2) + i(g_3 - g_4),$$

where $g_k \in \mathcal{L}_p(E,[0,\infty))$, $1 \le k \le 4$. (If g is real-valued, then $g_3 = g_4 = 0$.)

Given $\varepsilon > 0$, by Step One, we can find $\varphi_k \in \text{Simp}_p(E,[0,\infty))$ such that

$$\nu_p(g_k - \varphi_k) < \frac{\varepsilon}{4}, \quad 1 \le k \le 4.$$

Let $\varphi := (\varphi_1 - \varphi_2) + i(\varphi_3 - \varphi_4)$. Then $\varphi \in \text{Simp}_p(E,\mathbb{K})$ (since, as noted above, $\text{Simp}_p(E,\mathbb{K})$ is a vector space over $\mathbb{K}$), and

$$\begin{aligned}\|[f]-[\varphi]\|_p &= \nu_p(f-\varphi)\\ &= \nu_p\left(((f_1-f_2)+i(f_3-f_4)) - ((\varphi_1-\varphi_2)+i(\varphi_3-\varphi_4))\right)\\ &= \nu_p\left((f_1-\varphi_1)-(f_2-\varphi_2)+i(f_3-\varphi_3)-i(f_4-\varphi_4)\right)\\ &\le \sum_{k=1}^4 \nu_p(f_k-\varphi_k)\\ &< \sum_{k=1}^4 \frac{\varepsilon}{4}\\ &= \varepsilon.\end{aligned}$$

In particular, $[\text{Simp}_p(E,\mathbb{K})]$ is dense in $(L_p(E,\mathbb{K}),\|\cdot\|_p)$, $1 \le p < \infty$.

□

6.33. Proposition. *Let $E \in \mathfrak{M}(\mathbb{R})$ be a Lebesgue measurable set. Then*

$$[\text{Simp}(E,\mathbb{K})] := \{[\varphi] : \varphi \in \text{Simp}(E,\mathbb{K})\}$$

is dense in $(L_\infty E,\mathbb{K}),\|\cdot\|_\infty)$.

Proof.

Step One. Suppose first that $f \in \mathcal{L}_\infty(E,[0,\infty))$, and that there exists $M > 0$ such that $f(x) \le M$ for all $x \in E$.

Let $\varepsilon > 0$ and choose $N > 0$ such that $\frac{1}{N} < \varepsilon$. Decompose the interval $[0,M]$ into MN intervals of equal length $\frac{1}{N}$, namely $I_k := [\frac{k-1}{N},\frac{k}{N})$, $1 \le k \le MN-1$, and $I_{MN} := [M-\frac{1}{N},M]$.

Set $H_k := f^{-1}(I_k)$, $1 \le k \le MN$, so that H_k is measurable by the measurability of f, and set

$$\varphi := \sum_{k=1}^{MN} (\frac{k-1}{N})\chi_{H_k}.$$

Clearly $\varphi \in \text{Simp}(E,[0,\infty))$, and

$$|f(x)-\varphi(x)| \le \frac{1}{N} < \varepsilon \qquad \text{for all } x \in E.$$

In particular,

$$\nu_\infty(f-\varphi) < \varepsilon.$$

Step Two. Now suppose that $[g] \in L_\infty(E,\mathbb{K})$ is arbitrary, so that $g \in \mathcal{L}_\infty(E,\mathbb{K})$. As we have seen in paragraph 6.23, we may assume without loss of generality that $|g(x)| \le \nu_\infty(g)$ for all $x \in E$.

Moreover, as was the case in Proposition 6.32, we may write

$$g = (g_1 - g_2) + i(g_3 - g_4),$$

where $g_k \in \mathcal{L}_\infty(E, [0, \infty))$, and $|g_k(x)| \le \nu_\infty(g)$ for all $x \in E$ and $1 \le k \le 4$. Given $\varepsilon > 0$, by Step One, we can find $\varphi_k \in \mathcal{L}_\infty(E, [0, \infty))$ such that $\nu_\infty(g_k - \varphi_k) < \frac{\varepsilon}{4}$, $1 \le k \le 4$. The remainder of the proof is similar to that of the case where $1 \le p < \infty$, and we leave it as an exercise for the reader.

We conclude that $[\text{Simp}(E, \mathbb{R})]$ is dense in $(L_\infty(E, \mathbb{K}), \|\cdot\|_\infty)$.

□

We next consider a simple, but useful, result which we shall have occasion to apply twice. Recall that a **linear manifold** in a normed linear space $\mathfrak{X}$ is a vector subspace which may or may not be closed in the norm-topology.

6.34. Lemma. *Let $(\mathfrak{X}, \|\cdot\|)$ be a Banach space, and suppose that $\mathcal{B} \subseteq \mathfrak{X}$ satisfies $\overline{\text{span}}\,\mathcal{B} = \mathfrak{X}$. Suppose also that $\mathcal{L} \subseteq \mathfrak{X}$ is a linear manifold and that $\mathcal{B} \subseteq \overline{\mathcal{L}}$. Then*

$$\mathfrak{X} = \overline{\mathcal{L}}.$$

Proof. The key is to observe that since $\mathcal{L}$ is a linear manifold, $\overline{\mathcal{L}}$ is a closed subspace of $\mathfrak{X}$. Indeed, since $\mathfrak{X}$ is a vector space, we may apply the Subspace Test from first-year linear algebra. That is, it suffices to prove that $\overline{\mathcal{L}} \ne \varnothing$, and that if $x, y \in \overline{\mathcal{L}}$ and $\kappa \in \mathbb{K}$, then $kx + y \in \overline{\mathcal{L}}$.

To that end, note that $0 \in \mathcal{L}$ implies that $0 \in \overline{\mathcal{L}} \ne \varnothing$. Let $\kappa \in \mathbb{K}, x, y \in \overline{\mathcal{L}}$, and choose sequences $(x_n)_n, (y_n)_n \in \mathcal{L}$ such that $x = \lim_n x_n$ and $y = \lim_n y_n$.

Then $\kappa x_n + y_n \in \mathcal{L}$ for all $n \ge 1$ since $\mathcal{L}$ is a linear manifold, and

$$\begin{aligned}
\lim_n \|(\kappa x + y) - (\kappa x_n + y_n)\| &= \lim_n \|\kappa(x - x_n) + (y - y_n)\| \\
&\le \lim_n \; |\kappa|\,\|x - x_n\| + \|y - y_n\| \\
&= |\kappa| \lim_n \|x - x_n\| + \lim_n \|y - y_n\| \\
&= |\kappa| 0 + 0 \\
&= 0,
\end{aligned}$$

whence $\kappa x + y = \lim_n(\kappa x_n + y_n) \in \overline{\mathcal{L}}$. Thus $\overline{\mathcal{L}}$ is a closed subspace of $\mathfrak{X}$.

Since $\mathcal{B} \subseteq \overline{\mathcal{L}}$ and the latter is a closed subspace of $\mathfrak{X}$, it follows that

$$\mathfrak{X} = \overline{\text{span}}\,\mathcal{B} \subseteq \overline{\mathcal{L}}.$$

But $\overline{\mathcal{L}}$ is obviously contained in $\mathfrak{X}$, and thus the two sets are equal.

□

6.35. Proposition. *Let $a < b \in \mathbb{R}$. If $1 \le p < \infty$, then* $[\text{Step}([a,b],\mathbb{K})]$ *is dense in* $(L_p([a,b],\mathbb{K}), \|\cdot\|_p)$.

Proof. Let $\mathcal{B} := \{[\chi_H] : H \subseteq E \text{ measurable}\}$. Then $\operatorname{span}\mathcal{B} = [\text{Simp}([a,b],\mathbb{K})]$, and by Proposition 6.32, this is dense in $L_p([a,b],\mathbb{K})$. By Lemma 6.34, it suffices to show that every $[\chi_H]$ can be approximated arbitrarily well in the $\|\cdot\|_p$-norm by elements of $[\text{Step}([a,b],\mathbb{K})]$.

Let $H \subseteq [a,b]$ be a measurable set and $\varepsilon > 0$. Recall from Theorem 3.14 that we can find an open set $G \subseteq \mathbb{R}$ such that $H \subseteq G$, and $m(G \setminus H) < \frac{\varepsilon}{2}$. Write G as a *disjoint* union of open intervals $G = \cup_{n=1}^\infty (a_n, b_n)$. (Note that each interval is finite, since $mH \le m[a,b] < \infty$, and $m(G \setminus H) < \infty$), implying that $m(G) = m(H) + m(G \setminus H) < \infty$.)

Thus $m(G) = \sum_{n=1}^\infty (b_n - a_n) < \infty$. Set $G_n = \cup_{k=1}^n (a_k, b_k)$, $n \ge 1$, and choose $N \ge 1$ such that

$$m(G \setminus G_N) = \sum_{n=N+1}^\infty (b_n - a_n) < \frac{\varepsilon}{2}.$$

Set $\psi = \chi_{G_N \cap [a,b]}$ and observe that $\psi \in \text{Step}([a,b],\mathbb{R})$.

Moreover,

$$|\chi_H(x) - \psi(x)| = \begin{cases} 1 = |1-0| & \text{if } x \in H \setminus G_N \\ 1 = |0-1| & \text{if } x \in (G_N \cap [a,b]) \setminus H \\ 0 = |0-0| & \text{if } x \notin (G_N \cup H) \\ 0 = |1-1| & \text{if } x \in (G_N \cap H). \end{cases}$$

It follows that

$$\begin{aligned} \nu_p(\chi_H - \psi) &= \int_E |\chi_H - \psi|^p \\ &= \int_E |\chi_H - \psi| \\ &= m(H \setminus G_N) + m((G_N \cap [a,b]) \setminus H) \\ &\le m(H \setminus G_N) + m(G_N \setminus H) \\ &\le m(G \setminus G_N) + m(G \setminus H) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

It follows that $\|[\chi_H] - [\psi]\|_p < \varepsilon$, thus showing that $[\text{Step}([a,b],\mathbb{R})]$ is dense in $(L_p([a,b],\mathbb{R}), \|\cdot\|_p)$.

□

Observe that Lemma 6.34 greatly simplified the proof; instead of approximating an arbitrary element in $L_p([a,b],\mathbb{K})$, or even an arbitrary element of $[\text{Simp}(E,\mathbb{K})]$ by (equivalence classes of) step functions, we reduced the problem to that of approximating characteristic functions of measurable sets. In the same way, in proving that $[\mathcal{C}([a,b],\mathbb{K})]$ is dense in $(L_p([a,b],\mathbb{K}), \|\cdot\|_p)$, $1 \le p < \infty$, Lemma 6.34 will reduce the problem to that of approximating the characteristic function of an interval by a continuous function in the $\nu_p(\cdot)$-seminorm.

6.36. Theorem. *Let $a < b \in \mathbb{R}$. If $1 \le p < \infty$, then $[\mathcal{C}([a,b],\mathbb{K})]$ is dense in $(L_p([a,b],\mathbb{K}), \|\cdot\|_p)$.*

Proof. Let $\mathcal{B} := \{[\chi_{[r,s]} : a \le r < s \le b\}$. Then $\operatorname{span}\mathcal{B} = [\textsc{Step}([a,b],\mathbb{K})]$, which is dense in $L_p([a,b],\mathbb{K})$ by Proposition 6.35 above.

We may therefore appeal once again to Lemma 6.34, which implies that we need only show that every such $[\chi_{[r,s]}]$ can be approximated arbitrarily well in the $\|\cdot\|_p$-norm by elements of $[\mathcal{C}([a,b],\mathbb{K})]$.

To begin, choose $a \le r < s \le b$, and to dispense with a technicality, choose $M > 0$ such that $r + \frac{2}{M} < s$. For $n \ge M$, define

$$f_n(x) = \begin{cases} 0 & \text{if } a \le x \le r \\ n(x-r) & \text{if } r < x \le r + \frac{1}{n} \\ 1 & \text{if } r + \frac{1}{n} < x \le s - \frac{1}{n} \\ n(s-x) & \text{if } s - \frac{1}{n} < x \le s \\ 0 & \text{if } s < x \le b. \end{cases}$$

(If it not entirely clear from the outset why we picked such a sequence of functions, the reader would be well-advised to graph them. How could you approximate a step function by a continuous function in a simpler way? The choice of only defining f_n for $n \ge M$ is to ensure that $f_n \le \chi_{[r,s]}$ for all such n.)

Observe that f_n is continuous for each $n \ge M$, being piecewise linear. Also, $x \notin [r, r+\frac{1}{n}] \cup [s-\frac{1}{n}, s]$ implies that

$$|f_n(x) - \chi_{[r,s]}(x)| = 0,$$

and for all $x \in [a,b]$, $|f_n(x) - \chi_{[r,s]}(x)| \le 1$. It follows that for all $n \ge M$,

$$\begin{aligned} \nu_p(f_n - \chi_{[r,s]}) &= \left(\int_{[a,b]} |f_n(x) - \chi_{[r,s]}(x)|^p \right)^{\frac{1}{p}} \\ &\le \left(\int_{[r,r+\frac{1}{n}]\cup[s-\frac{1}{n},s]} 1 \right)^{\frac{1}{p}} \\ &= \left(\frac{2}{n} \right)^{\frac{1}{p}}. \end{aligned}$$

Thus

$$0 \le \lim_{n\to\infty} \|[f_n] - [\chi_{[r,s]}]\|_p \le \lim_{n\to\infty} \left(\frac{2}{n}\right)^{\frac{1}{p}} = 0,$$

and hence

$$\lim_{n\to\infty} [f_n] = [\chi_{[r,s]}]$$

in $(L_p([a,b],\mathbb{R}), \|\cdot\|_p)$, completing the proof.

□

6.37. This leads to the following interesting result. Recall first that a topological space is said to be **separable** if it admits a countable dense subset.

Secondly, recall that if (X,d) is a separable metric space, $\delta > 0$ is a positive real number and $\{x_\lambda : \lambda \in \Lambda\} \subseteq X$ satisfies $d(x_\alpha, x_\beta) \ge \delta$ for all $\alpha \ne \beta \in \Lambda$, then Λ is countable. The proof is left to the exercises.

6.38. Corollary. *Let $a < b \in \mathbb{R}$.*

(a) *If $1 \le p < \infty$, then $(L_p([a,b],\mathbb{K}), \|\cdot\|_p)$ is separable.*
(b) *The space $(L_\infty([a,b],\mathbb{K}), \|\cdot\|_\infty)$ is not separable.*

Proof.

(a) First fix $1 \le p < \infty$.

Recall from Section 6.29 that $L_\infty([a,b],\mathbb{K}) \subseteq L_p([a,b],\mathbb{K})$. Moreover, the proof of this assertion showed that if $[f],[g] \in L_\infty([a,b],\mathbb{K})$, then

$$\|[f]-[g]\|_p = \|[f-g]\|_p \le \|[f]-[g]\|_\infty (b-a)^{1/p}.$$

Let $\varepsilon > 0$, and let $[h] \in \mathcal{L}_p([a,b],\mathbb{K})$. We know from Theorem 6.36 that $[\mathcal{C}([a,b],\mathbb{K})]$ is dense in $L_p([a,b],\mathbb{K})$ with respect to the p-norm.

Thus we can find $g \in \mathcal{C}([a,b],\mathbb{K})$ such that

$$\|[h]-[g]\|_p < \frac{\varepsilon}{3}.$$

Of course, by the Weierstraß Approximation Theorem and Example 6.19 above, we know that we can find a polynomial $p(x) = p_0 + p_1 x + \cdots + p_m x^m$ such that

$$\|[g]-[p]\|_\infty = \|g-p\|_{\sup} < \frac{\varepsilon}{3(b-a)^{1/p}}.$$

By Exercise 6.9 below, we can find a polynomial $q(x) = q_0 + q_1 x + \cdots + q_m x^m$ such that $q_k \in \mathbb{Q} + i\mathbb{Q}$ for all $0 \le k \le m$ and

$$\|[p]-[q]\|_\infty = \|p-q\|_{\sup} < \frac{\varepsilon}{3(b-a)^{1/p}}.$$

Finally,

$$\begin{aligned}
\|[h]-[q]\|_p &\le \|[h]-[g]\|_p + \|[g]-[p]\|_p + \|[p]-[q]\|_p \\
&\le \|[h]-[g]\|_p + \|[g]-[p]\|_\infty (b-a)^{1/p} + \|[p]-[q]\|_\infty (b-a)^{1/p} \\
&\le \frac{\varepsilon}{3} + \left[\frac{\varepsilon}{3(b-a)^{1/p}}\right](b-a)^{1/p} + \left[\frac{\varepsilon}{3(b-a)^{1/p}}\right](b-a)^{1/p} \\
&= \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon.
\end{aligned}$$

Thus the family $(\mathbb{Q} + i\mathbb{Q})[x]$ of all polynomials with (complex) rational coefficients, when viewed as continuous functions on $[a,b]$, has the property that $[(\mathbb{Q}+i\mathbb{Q})[x]]$ is dense in $(L_p([a,b],\mathbb{K}), \|\cdot\|_p)$. Since there are at most countably many elements in $(\mathbb{Q}+i\mathbb{Q})[x]$ and therefore in $[(\mathbb{Q}+i\mathbb{Q})[x]]$, we see that $L_p([a,b],\mathbb{K})$ is separable, $1 \le p < \infty$.

(b) Now consider $p = \infty$, and suppose that $a \le r_1 < s_1 \le b$ and $a \le r_2 < s_2 \le b$. Suppose furthermore that either $r_1 \ne r_2$ or that $s_1 \ne s_2$. By Exercise 6.10 below, the symmetric difference

$$[r_1, s_1]\Delta[r_2, s_2]$$

of the intervals $[r_1, s_1]$ and $[r_2, s_2]$ contains an interval, say $[u, v] \subseteq [a, b]$, with $u < v$.

For all $x \in [u, v]$, $|\chi_{[r_1,s_1]}(x) - \chi_{[r_2,s_2]}(x)| = 1$, and so

$$\|[\chi_{[r_1,s_1]}] - [\chi_{[r_2,s_2]}]\|_\infty = \|[\chi_{[r_1,s_1]\Delta[r_2,s_2]}]\|_\infty = 1.$$

Let $\Lambda := \{(r, s) \in \mathbb{R}^2 : a \le r < s \le b\}$. Then Λ is uncountable. For $(r_1, s_2) \ne (r_2, s_2) \in \Lambda$, the above estimate shows that

$$\|\chi_{[r_1,s_1]} - \chi_{[r_2,s_2]}\|_\infty = 1.$$

By the comment preceding this Corollary (see Exercise 6.8), $L_\infty([a, b], \mathbb{K})$ is not separable.

□

Appendix to Section 6.

6.39. If the reader consults almost any other text on measure theory (and the reader will absolutely fall in this author's estimation if they do not), they will almost assuredly observe that we have been *incredibly* pedantic in our approach to these notes. Most texts will use the same notation, namely "f", to refer to both a measurable function, as well as to its equivalence class in "$L_p(E, \mathbb{K})$", where $E \in \mathfrak{M}(\mathbb{R})$ is a measurable set. It is left to the reader to keep track of when they are dealing with a function, and when they are dealing with its image in $L_p(E, \mathbb{K})$. Statements such as "$f(x) = x$ a.e. on $[0,1]$" are used to hint that we are talking about an equivalence class rather than a function. Truth be told (and what's the point of not telling the truth in a set of course notes?), we have always felt ambivalent about this approach. Without having taken any courses in math pedagogy, experience has taught us that people inevitably make the mistake of treating an element of $L_p(E, \mathbb{K})$ as a function – for example, referring to $f(x)$, for some $x \in E$, when this concept is no longer valid (for the equivalence class of f).

For this reason, we have attempted to consistently denote *functions* by simple letters, e.g. f and g, while denoting their *equivalence classes* using the bracket notation $[f]$ and $[g]$. While this is more cumbersome, it has the advantage of being more precise, and our hope is that for the person learning about measure theory for the first time, it will help to keep the two concepts separate. Once sufficient mathematical maturity is acquired (oh, maybe a month from now), the reader should be able to consult other texts with sufficient sophistication to handle any notation thrown at them.

6.40. Which brings us to another point. In order to deal with pointwise limits of increasing functions for the Monotone Convergence Theorem 5.16, we introduced the notion of *extended real numbers*, and then of measurable functions $f : E \to \overline{\mathbb{R}}$ (where again, $E \in \mathfrak{M}(\mathbb{R})$ is a measurable set). In many textbooks dealing with Lebesgue measure, the equivalence classes of functions in "$L_p(E)$" consist of *extended real-valued* measurable functions. (We have restricted our attention to equivalence classes of *real-valued* measurable functions.)

In fact, both approaches lead to essentially the same theory. For any $1 \le p \le \infty$, if $f : E \to \overline{\mathbb{R}}$ is measurable and $\nu_p(f) < \infty$, then the set $B := \{x \in E : |f(x)| = \infty\}$ has measure zero, and so we can always find $g : E \to \mathbb{R}$ such that $g = f$ a.e. on E; for example, we can choose $g = f \cdot \chi_{E \setminus B}$. Thus $[f] = [g]$.

So why not start with the vector space $\mathcal{L}_p(E, \overline{\mathbb{R}})$ and use Proposition 6.3 to define the quotient space $L_p(E, \overline{\mathbb{R}})$? Well, perhaps the best answer we can give is that $\mathcal{L}_p(E, \overline{\mathbb{R}})$ DOES NOT define a vector space over $\mathbb{R}$! Suppose

$$f(x) = \begin{cases} \infty & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases} \qquad \text{and} \qquad g(x) = \begin{cases} -\infty & \text{if } x = 0 \\ 0 & \text{if } x \ne 0 \end{cases}.$$

What should $f + g(0)$ be? We haven't defined $\infty + (-\infty)$, and we won't. Rats.

So then where do these other authors get off defining "$L_p(E)$" in terms of equivalence classes of extended real-valued functions? Well, for one thing, their construction of "$L_p(E)$" as often as not does not pass through Proposition 6.3. Instead, they define two extended real-valued functions f and g to be equivalent if $f = g$ a.e. on E. While the set of extended real-valued functions does not form a vector space, the equivalence classes defined **do**! In fact, up to isometric isomorphism, they form exactly the same L_p-spaces we have defined. We like the seminorm-to-norm-via-quotient-spaces approach we have used, in part because we do not have to wave our hands at any point of our construction. There you go.

6.41. One obvious reason for studying Hölder's Inequality is that it was required to obtain Minkowski's Inequality, which we needed to prove that each $\mathcal{L}_p(E, \mathbb{K})$ was a linear space. The construction of the conjugate function $f^* \in L_q(E, \mathbb{K})$ from $f \in L_p(E, \mathbb{K})$ might at first glance seem rather arcane, and the eager novice might wonder why we would be interested in such a thing.

The answer lies in part in Functional Analysis - which is the study of normed linear spaces and the continuous linear maps between them. Given a Banach space $(\mathfrak{X}, \|\cdot\|)$, one defines the **dual space** of $\mathfrak{X}$ to be

$$\mathfrak{X}^* := \mathcal{B}(\mathfrak{X}, \mathbb{K}) = \{x^* : \mathfrak{X} \to \mathbb{K} : x^* \text{ is linear and continuous}\}.$$

Elements of the dual space are called **continuous linear functionals** on $\mathfrak{X}$.

As we have seen in the Assignments, linear maps between normed linear spaces are continuous if and only if they are bounded. The dual space carries a great deal of information about the space $\mathfrak{X}$ itself, and perhaps the most famous and important theorem from Functional Analysis is the **Hahn-Banach Theorem**. This is not actually a single result, but rather two classes of results, all referred to by that same name. It is usually left up to the reader to recognize which version of the Hahn-Banach Theorem is being invoked in any given application. We invite the reader to consult the (free!) reference [**3**] for more details.

One of the important consequences of the Hahn-Banach Theorem is that the dual space $\mathfrak{X}^*$ of $\mathfrak{X}$ has sufficiently many functionals to *separate points* of $\mathfrak{X}$, meaning that if $x_1 \neq x_2 \in \mathfrak{X}$, then there exists $x^* \in \mathfrak{X}^*$ such that $x^*(x_1) \neq x^*(x_2)$. In fact, one can do better – given $y \in \mathfrak{X}$, one can find a linear functional $y^* \in \mathfrak{X}^*$ such that $\|y^*\|_{\mathfrak{X}^*} = 1$ and $y^*(y) = \|y\|$. (Applying this to $y = x_1 - x_2$ above shows that the corresponding y^* will separate x_1 and x_2.)

What Hölder's Inequality (Theorem 6.15) tells us is that with $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, given $[g] \in L_q(E, \mathbb{K})$, we may define a linear functional

$$\begin{array}{rccc} \Phi_{[g]}: & L_p(E, \mathbb{K}) & \to & \mathbb{K} \\ & [f] & \mapsto & \int_E fg \end{array}$$

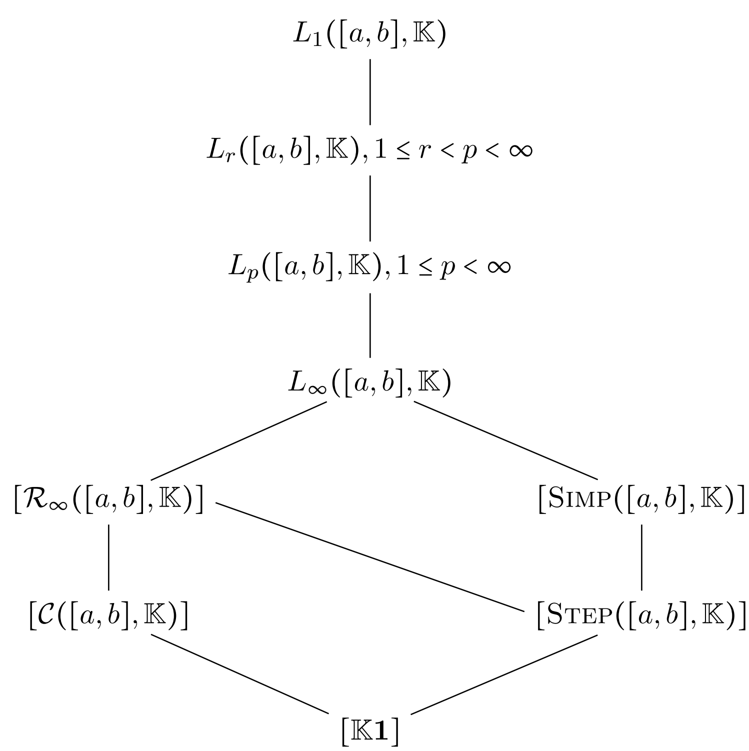
and that this linear map will be bounded with norm at most $\|[g]\|_q$. In the Assignments, we verify that *every* linear functional $\varphi \in L_p(E, \mathbb{K})$ is of the form $\varphi = \Phi_{[g]}$ for some $[g] \in L_q(E, \mathbb{K})$, and that in fact $\|\varphi\| = \|[g]\|_q$. In other words, *we are identifying* $L_q(E, \mathbb{K})$ with the dual space $(L_p(E, \mathbb{K}))^*$ in an isometrically isomorphic manner.

The second half of Theorem 6.15 is just an example verifying the Corollary to the Hahn-Banach Theorem that we mentioned above, namely: the functional $\Phi_{[f^*]} \in (L_p(E,\mathbb{K}))^*$ is the norm-one functional which sends $[f]$ to $\|[f]\|_p$.

For more general measure spaces (Z,μ) – which we have not discussed at all in these notes – we still have that the dual of $L_p(Z,\mu)$ may be isometrically isomorphically identified with $L_q(Z,\mu)$ when $1 < p < \infty$, but problems start to arise with the dual spaces of $L_1(Z,\mu)$ and $L_\infty(Z,\mu)$. Without getting into the details at all (the interested reader may find them online), in the case where the space (Z,μ) is **decomposable**, which includes the case where (Z,μ) is a **σ-finite** measure space, we *do* have that $(L_1(Z,\mu))^* = L_\infty(Z,\mu)$. In particular, this applies to the case of $\ell_1(I)$, where I is any set equipped with counting measure. Thus we can always identify $(\ell_1(I))^*$ with $\ell_\infty(I)$, which is a happy, happy situation. The dual of $L_\infty(Z,\mu)$ tends to be a real can of worms. For example, the dual of $\ell_\infty(\mathbb{N})$ may be identified with the so-called **regular Borel measures** on the **Stone-Čech compactification** $\beta\mathbb{N}$ of the natural numbers. It's big – very, very big.

6.42. The following diagram illustrates the relationship between the L_p-spaces when the underlying measure space is a bounded interval. We use $\mathbf{1}$ to denote constant function $\mathbf{1}(x) = x$ for all $x \in [a,b]$.

Inclusions of L_p-spaces



Exercises for Section 6.

Exercise 6.1. Let $E \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$ be a function. Define

$$\Xi_f(x) = \begin{cases} 1 & \text{if } f(x) \geq 0 \\ -1 & \text{if } f(x) < 0 \end{cases}, \quad x \in E.$$

Prove that if E is measurable and $f \in \mathcal{L}(E, \mathbb{R})$, then $\Xi_f \in \mathcal{L}(E, \mathbb{R})$ as well.

Exercise 6.2.

Let $E \in \mathfrak{M}(\mathbb{R})$ satisfy $mE < \infty$ and $\varphi \in \text{SIMP}(E, \mathbb{K})$. Prove that for all $1 \leq p \leq \infty$, $\varphi \in \mathcal{L}_p(E, \mathbb{K})$.

Exercise 6.3.

Let $E \in \mathfrak{M}(\mathbb{R})$ and $1 < p < \infty$. Suppose that $[f] \in L_p(E, \mathbb{K})$ and $[g] \in L_q(E, \mathbb{K})$. Then

$$[f] \cdot [g] := [fg]$$

is well-defined, and $[fg] \in L_1(E, \mathbb{K})$.

Exercise 6.4. Assignment Question.

Let $(\mathfrak{X}, \|\cdot\|)$ be a normed linear space. Prove that $\mathfrak{X}$ is complete, and hence a Banach space, if and only if every absolutely summable series in $\mathfrak{X}$ is summable.

Here, a series $\sum_{n=1}^{\infty} x_n$ in $\mathfrak{X}$ is said to be **summable** if

$$x := \lim_{N \to \infty} \sum_{n=1}^{N} x_n$$

exists in $\mathfrak{X}$, while the series is said to be **absolutely summable** if

$$\lim_{N \to \infty} \sum_{n=1}^{N} \|x_n\| < \infty.$$

Exercise 6.5.

Paragraph 6.29 and the subsequent paragraphs establish a set of relationships between L_p-spaces for various values of $p \in [1, \infty]$, as well as the spaces of (equivalence classes) of step functions and simple functions, in the case where the underlying measurable domain $E \in \mathfrak{M}(\mathbb{R})$ has finite measure.

Determine what relations hold between these spaces in the case where $mE = \infty$.

Exercise 6.6.

Let $E \in \mathfrak{M}(\mathbb{R})$, and $1 \leq p \leq \infty$. Let $\mathcal{D} \subseteq \mathcal{L}_p(E, \mathbb{K})$. Prove that the following are equivalent:

(a) The set $[\mathcal{D}] := \{[f] : f \in \mathcal{D}\}$ is dense in $L_p(E, \mathbb{K})$.
(b) For each $g \in \mathcal{L}_p(E, \mathbb{K})$ and $\varepsilon > 0$, there exists $f \in \mathcal{D}$ such that

$$\nu_p(f - g) < \varepsilon.$$

Exercise 6.7.

Recall that given $E \in \mathfrak{M}(\mathbb{R})$ and $1 \le p \le \infty$, we defined

$$\mathrm{SIMP}_p(E, \mathbb{K}) = \mathrm{SIMP}(E, \mathbb{K}) \cap \mathcal{L}_p(E, \mathbb{K}).$$

Prove that if $mE < \infty$ or if $p = \infty$, then $\mathrm{SIMP}_p(E, \mathbb{K}) = \mathrm{SIMP}(E, \mathbb{K})$.

Exercise 6.8.

Let (X, d) be a metric space. Suppose that there exists an uncountable set $\{x_\lambda : \lambda \in \Lambda\}$ in X and a positive real number $\delta > 0$ such that $d(x_\alpha, x_\beta) \ge \delta$ for all $\alpha \ne \beta \in \Lambda$. Prove that (X, d) is not separable.

Exercise 6.9.

Let $a < b \in \mathbb{R}$ and let $p(x) = p_0 + p_1 x + \cdots + p_m x^m$ be a polynomial of degree m in $(\mathcal{C}([a, b], \mathbb{R}), \|\cdot\|_{\sup})$. Prove that given any $\varepsilon > 0$, there exists a polynomial $q(x) = q_0 + q_1 x + \cdots + q_m x^m$ in $(\mathcal{C}([a, b], \mathbb{R}), \|\cdot\|_{\sup})$ such that $q_k \in \mathbb{Q}$, $0 \le k \le m$, and

$$\|p - q\|_{\sup} < \varepsilon.$$

Conclude that given a polynomial $r(x) = r_0 + r_1 x + \cdots + r_m x^m$ of degree m in $(\mathcal{C}([a, b], \mathbb{C}), \|\cdot\|_{\sup})$ and $\varepsilon > 0$, we can find a polynomial $s(x) = s_0 + s_1 x + \cdots + s_m x^m$ with $s_k \in \mathbb{Q} + i\mathbb{Q}$, $0 \le k \le m$, such that

$$\|r - s\|_{\sup} < \varepsilon.$$

Exercise 6.10.

Recall that if A and B are sets, we define the **symmetric difference** of A and B to be

$$A \Delta B := (A \cup B) \setminus (B \cap A).$$

Let $a < b$ be real numbers and suppose that $a \le r_1 < s_1 \le b$ and $a \le r_2 < s_2 \le b$. Suppose furthermore that either $r_1 \ne r_2$ or that $s_1 \ne s_2$. Prove that the symmetric difference

$$[r_1, s_1] \,\Delta\, [r_2, s_2]$$

contains a non-degenerate interval $[u, v] \subseteq [a, b]$. (By *non-degenerate*, we simply mean that $u < v$.)

Exercise 6.11.

Let $E \in \mathfrak{M}(\mathbb{R})$. Prove that the map:

$$\begin{array}{rccc} \Omega : & (\mathcal{C}(E, \mathbb{K}), \|\cdot\|_{\sup}) & \to & (L_\infty(E, \mathbb{K}), \|\cdot\|_\infty) \\ & f & \mapsto & [f] \end{array}$$

is a linear isometry, and deduce that $[\mathcal{C}(E, \mathbb{K})] := \{[f] : f \in \mathcal{C}(E, \mathbb{K})\}$ is a closed subspace Banach space of $L_\infty(E, \mathbb{K})$ which is (isometrically) isomorphic to $\mathcal{C}(E, \mathbb{K})$.

In other words, as Banach spaces, we can identify $\mathcal{C}(E, \mathbb{K})$ with its image $[\mathcal{C}(E, \mathbb{K})]$ in $L_\infty(E, \mathbb{K})$.

7. Hilbert spaces

Smoking kills. If you're killed, you've lost a very important part of your life.

Brooke Shields

7.1. We have seen in the previous section that if $E \in \mathfrak{M}(\mathbb{R})$ and $1 \le p \le \infty$, then $(L_p(E, \mathbb{R}), \|\cdot\|_p)$ is a Banach space. The case where $p = 2$ is very special and merits individual attention.

Let us recall the following definition.

7.2. Definition. *An **inner product** on a $\mathbb{K}$-vector space $\mathcal{H}$ is a function $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{K}$ which satisfies:*

(a) $\langle x, x \rangle \ge 0$ *for all* $x \in \mathcal{H}$;
(b) $\langle x, x \rangle = 0$ *if and only if* $x = 0$;
(c) $\langle \kappa x + y, z \rangle = \kappa \langle x, z \rangle + \langle y, z \rangle$ *for all* $x, y, z \in \mathcal{H}$ *and* $\kappa \in \mathbb{K}$; *and*
(d) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ *for all* $x, y \in \mathbb{K}$.

*We say that x and y are **orthogonal** if $\langle x, y \rangle = 0$, and we write $x \perp y$.*

In the case where $\mathbb{K} = \mathbb{R}$, the "complex conjugate" appearing in (d) is obviously superfluous.

7.3. Theorem. The Cauchy-Schwarz Inequality. *Suppose that $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is an inner product space over $\mathbb{K}$. Then*

$$|\langle x, y \rangle| \le \langle x, x \rangle^{1/2} \langle y, y \rangle^{1/2}$$

for all $x, y \in \mathcal{H}$.

Proof. Let $x, y \in \mathcal{H}$. If $\langle x, y \rangle = 0$, then there is nothing to prove.

Suppose therefore that $\langle x, y \rangle \ne 0$. For any $\kappa \in \mathbb{K}$,

$$\begin{aligned} 0 &\le \langle x - \kappa y, x - \kappa y \rangle \\ &= \langle x, x \rangle - \kappa \langle y, x \rangle - \overline{\kappa} \langle x, y \rangle + |\kappa|^2 \langle y, y \rangle. \end{aligned}$$

Setting $\kappa = \dfrac{\langle x, y \rangle}{\langle y, y \rangle}$ yields

$$0 \le \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} + \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle},$$

which is equivalent to

$$|\langle x, y \rangle|^2 \le \langle x, x \rangle \langle y, y \rangle,$$

as required.

□

7.4. Proposition. *Let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be an inner product space. Then the map*

$$\|x\| := \langle x, x \rangle^{1/2}, \quad x \in \mathcal{H}$$

defines a norm on $\mathcal{H}$, called the ***norm induced by the inner product****.*

Proof. We shall prove that it defines a norm, and leave it to the reader to prove that this is what it is called.

Let $x, y \in \mathcal{H}$ and $\kappa \in \mathbb{K}$.

(a) Clearly $\|x\| \ge 0$, as $\langle x, x \rangle \ge 0$.
(b) Note that $\|x\| = 0$ if and only if $\|x\|^2 = \langle x, x \rangle = 0$, which happens if and only if $x = 0$.
(c) $\|\kappa x\|^2 = \langle \kappa x, \kappa x \rangle = |\kappa|^2 \ \langle x, x \rangle = |\kappa|^2 \ \|x\|^2$, and thus

$$\|\kappa x\| = |\kappa| \ \|x\|.$$

(d) From the Cauchy-Schwarz Inequality,

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &\le \|x\|^2 + |\langle x, y \rangle| + |\langle y, x \rangle| + \|y\|^2 \\ &= \|x\|^2 + 2\|x\| \ \|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2.\end{aligned}$$

Thus $\|x+y\| \le \|x\| + \|y\|$.

This completes the proof.

□

It follows that every inner product space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is also a normed linear space $(\mathcal{H}, \|\cdot\|)$, using the norm induced by the inner product. Unless we explicitly mention a different norm for $\mathcal{H}$ (which is highly unlikely), we shall always assume that this is the norm to which we are referring. Also - let us not forget that every normed linear space is also a metric space, using the metric induced by the norm.

As such, every inner product space is a metric space, under the metric induced by the norm induced by the inner product.

Although a Hilbert space is technically an order pair consisting of a vector space $\mathcal{H}$ and an inner product $\langle \cdot, \cdot \rangle$, we shall typically speak informally of the Hilbert space $\mathcal{H}$.

7.5. Definition. *A* ***Hilbert space*** *is a complete inner product space.*

7.6. Examples.

(a) Let $N \ge 1$ be an integer. Consider $x = (x_n)_{n=1}^N$ and $y = (y_n)_{n=1}^N \in \mathcal{H} := \mathbb{C}^N$. The map

$$\langle x, y \rangle := \sum_{n=1}^N x_n \overline{y_n}$$

defines an inner product on $\mathcal{H}$, called the **standard inner product** on $\mathcal{H}$, and $\mathcal{H}$ is complete with respect to the norm induced by this inner product. Thus $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is a Hilbert space.

(b) We can make this somewhat more general as follows. Fix an integer $1 \le N$ and choose strictly positive real numbers $\rho_1, \rho_2, \ldots, \rho_N$. We leave it to the reader that $\mathbb{C}^N$ becomes a Hilbert space when equipped with the inner product

$$\langle x, y \rangle_\rho := \sum_{n=1}^{N} \rho_n x_n \overline{y_n}.$$

(c) $\ell^2 = \{(x_n)_n : x_n \in \mathbb{K}, n \ge 1 \text{ and } \sum_{n=1}^\infty |x_n|^2 < \infty\}$ is a Hilbert space, with the inner product given by

$$\langle (x_n)_n, (y_n)_n \rangle = \sum_n x_n \overline{y_n}.$$

Again, this is referred to as the **standard inner product** on ℓ^2.

7.7. Theorem. *Let $E \in \mathfrak{M}(\mathbb{R})$. The map*

$$\begin{array}{cccc} \langle \cdot, \cdot \rangle & L_2(E, \mathbb{K}) \times L_2(E, \mathbb{K}) & \to & \mathbb{K} \\ & ([f], [g]) & \mapsto & \int_E f \overline{g} \end{array}$$

defines an inner product on $L_2(E)$. Furthermore, the norm induced by this inner product is the L_2-norm $\|\cdot\|_2$ on $L_2(E, \mathbb{K})$. Since $(L_2(E, \mathbb{K}), \|\cdot\|_2)$ is complete, it is a Hilbert space.

Proof. First observe that the map above is well-defined. That is, if $[f_1] = [f]$ and $[g_1] = [g]$, then $f_1 = f$ and $g_1 = g$ a.e. on E, so that $f_1 \overline{g_1} = f \overline{g}$ a.e. on E.

Furthermore, by Hölder's Inequality,

$$\int_E |f \overline{g}| \le \|[f]\|_2 \ \|[g]\|_2 < \infty,$$

and so $[f\overline{g}] = [f_1 \overline{g_1}] \in L_1(E)$. The fact that $f_1 \overline{g_1} = f \overline{g}$ a.e. on E also implies that

$$\int_E f_1 \overline{g_1} = \int_E f \overline{g} \in \mathbb{K},$$

and so the map is well-defined, as claimed.

Let $[f], [g], [h] \in L_2(E)$ and $\kappa \in \mathbb{K}$. Then

(a)

$$\langle [f], [f] \rangle = \int_E |f|^2 \ge 0,$$

and equality occurs if and only if $f = 0$ a.e. on E, i.e. if and only if $[f] = 0$. (This also shows that $\langle [f], [f] \rangle^{\frac{1}{2}} = \|[f]\|_2$.)

(b)

$$\begin{aligned} \langle \kappa [f] + [g], [h] \rangle &= \int_E (\kappa f + g) \overline{h} \\ &= \kappa \int_E f \overline{h} + \int_E g \overline{h} \\ &= \kappa \langle [f], [h] \rangle + \langle [g], [h] \rangle. \end{aligned}$$

(c)

$$\langle [f],[g]\rangle = \int_E f\overline{g} = \overline{\int_E \overline{f} g} = \overline{\langle [g],[f]\rangle}.$$

Thus $\langle \cdot, \cdot \rangle$ is an inner product (this is the **standard inner product on** $L_2(E, \mathbb{K})$), and the norm induced by this inner product is the $\| \cdot \|_2$-norm. The last statement is clear.

□

7.8. Recall that a subset $\mathcal{E}$ of an inner product space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is said to be **orthogonal** if $x \neq y$ in $\mathcal{E}$ implies that $\langle x, y\rangle = 0$. Also, $\mathcal{E}$ is said to be an **orthonormal set** if $\mathcal{E}$ is orthogonal and $x \in \mathcal{E}$ implies that $\|x\| = 1 = \langle x, x\rangle^{\frac{1}{2}}$.

7.9. Definition. *Let $\mathcal{H}$ be a Hilbert space. An* ***orthonormal basis*** *(or* ***Hilbert space basis*** *for $\mathcal{H}$ is a maximal orthonormal set in $\mathcal{H}$. (Here, maximal is with respect to inclusion.)*

We shall abbreviate the phrase "orthonormal basis" to the acronym ONB.

7.10. Remarks.

(a) By Zorn's Lemma, every orthonormal set in $\mathcal{H}$ can be extended to an ONB for $\mathcal{H}$.

(b) If $\mathcal{H}$ is infinite-dimensional, then an ONB for $\mathcal{H}$ is *not* a Hamel basis for $\mathcal{H}$.

7.11. Examples.

(a) Let $N \geq 1$ be an integer and consider $\mathcal{H} = \mathbb{C}^N$, equipped with the standard inner product $\langle \cdot, \cdot \rangle$. For $1 \leq n \leq N$, define $e_n = (\delta_{n,k})_{k=1}^N$, where $\delta_{a,b}$ denotes the **Kronecker delta function**. Then $\{e_n\}_{n=1}^N$ is an ONB for $\mathcal{H}$.

(b) Let $N \geq 1$ be an integer and $\rho_k = k$, $1 \leq k \leq N$. Set $e_n = (\frac{1}{\sqrt{k}}\delta_{n,k})_{k=1}^N$, where $\delta_{a,b}$ denotes the Kronecker delta function. Then $\{e_n\}_{n=1}^N$ is an ONB for $\mathbb{C}^N$, equipped with the inner product from Example 7.6 (b), namely

$$\langle x, y\rangle_\rho := \sum_{n=1}^N \rho_n x_n \overline{y_n}.$$

(c) Generalizing example (a) above to infinite dimensions, let $\mathcal{H} = \ell_2$, equipped with the standard inner product, and for $n \geq 1$, let $e_n = (\delta_{n,k})_{k=1}^\infty$, where $\delta_{a,b}$ denotes the Kronecker delta function.Then $\{e_n\}_{n=1}^\infty$ is an ONB for $\mathcal{H}$. The proof is left as an exercise.

(d) Let $\mathcal{H} = L^2([0, 2\pi], \mathbb{C})$, equipped with the inner product

$$\langle [f],[g]\rangle := \int_{[0,2\pi]} f\overline{g}.$$

For $n \in \mathbb{Z}$, define the continuous function

$$\begin{array}{cccc} \xi_n : & [0,2\pi] & \to & \mathbb{C} \\ & \theta & \mapsto & \frac{1}{\sqrt{2\pi}} e^{in\theta}. \end{array}$$

Then $[\xi_n] \in L^2([0, 2\pi], \mathbb{C})$ for all $n \in \mathbb{Z}$. In the Assignments, we shall see that $\{[\xi_n]\}_{n\in\mathbb{Z}}$ is an ONB for $L^2([0, 2\pi], \mathbb{C})$.

We recall from Linear Algebra:

7.12. Theorem. ***The Gram-Schmidt Orthogonalisation Process***
If $\mathcal{H}$ is a Hilbert space over $\mathbb{K}$ and $\{x_n\}_{n=1}^\infty$ is a linearly independent set in $\mathcal{H}$, then we can find an orthonormal set $\{y_n\}_{n=1}^\infty$ in $\mathcal{H}$ so that $\operatorname{span}\{x_1, x_2, ..., x_N\} = \operatorname{span}\{y_1, y_2, ..., y_N\}$ *for all $N \ge 1$.*
Proof. We leave it to the reader to verify that setting $y_1 = x_1/\|x_1\|$, and recursively defining

$$y_N := \frac{x_N - \sum_{n=1}^{N-1}\langle x_N, y_n\rangle y_n}{\|x_N - \sum_{n=1}^{N-1}\langle x_N, y_n\rangle y_n\|}, \qquad N \ge 1$$

will do.

□

7.13. Theorem. *Let $\mathcal{H}$ be a Hilbert space and suppose that $x_1, x_2, ..., x_n \in \mathcal{H}$.*

(a) [***The Pythagorean Theorem***] *If the vectors are pairwise orthogonal, then*

$$\|\sum_{j=1}^n x_j\|^2 = \sum_{j=1}^n \|x_j\|^2.$$

(b) [***The Parallelogram Law***]

$$\|x_1 + x_2\|^2 + \|x_1 - x_2\|^2 = 2\left(\|x_1\|^2 + \|x_2\|^2\right).$$

Proof. Both of these results follow immediately from the definition of the norm in terms of the inner product.

□

It follows that if $(\mathfrak{X}, \|\cdot\|)$ is a Banach space and the Parallelogram Law does *not* hold in $\mathfrak{X}$, then there does not exist an inner product on $\mathfrak{X}$ such that $\|\cdot\|$ is the norm induced by the inner product. In particular, $\mathfrak{X}$ is not an Hilbert space. Far less obvious, but still true, is the fact that if the norm *does* satisfy the Parallelogram Law, then there exists an inner product such that $\|\cdot\|$ is the norm induced by that inner product. We shall not prove this here.

7.14. Theorem. *Let $\mathcal{H}$ be a Hilbert space, and $K \subseteq \mathcal{H}$ be a closed, non-empty convex subset of $\mathcal{H}$. Given $x \in \mathcal{H}$, there exists a unique point $y \in K$ which is closest to x; that is,*

$$\|x - y\| = \operatorname{dist}(x, K) := \min\{\|x - z\| : z \in K\}.$$

Proof. The proof is left as an Assignment question.

□

7.15. Theorem. *Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace. Let $x \in \mathcal{H}$, and $m \in \mathcal{M}$. The following are equivalent:*

(a) $\|x - m\| = \text{dist}(x, \mathcal{M})$;

(b) *The vector $x - m$ is orthogonal to $\mathcal{M}$, i.e., $\langle x - m, y\rangle = 0$ for all $y \in \mathcal{M}$.*

Proof.

(a) implies (b): Suppose that $\|x-m\| = \text{dist}\,(x, \mathcal{M})$, and suppose to the contrary that there exists $y \in \mathcal{M}$ so that $\kappa := \langle x - m, y\rangle \neq 0$. There is no loss of generality in assuming that $\|y\| = 1$. Consider $z = m + \kappa y \in \mathcal{M}$. Then

$$\begin{aligned} \|x - z\|^2 &= \|x - m - \kappa y\|^2 \\ &= \langle x - m - \kappa y, x - m - \kappa y\rangle \\ &= \|x - m\|^2 - \kappa\langle y, x - m\rangle - \overline{\kappa}\langle x - m, y\rangle + |\kappa|^2\,\|y\|^2 \\ &= \|x - m\|^2 - |\kappa|^2 \\ &< \text{dist}(x, \mathcal{M}), \end{aligned}$$

a contradiction. Hence $x - m \in \mathcal{M}^\perp$.

(b) implies (a): Suppose that $x - m \in \mathcal{M}^\perp$. If $z \in \mathcal{M}$ is arbitrary, then $y := z - m \in \mathcal{M}$, so by the Pythagorean Theorem,

$$\|x - z\|^2 = \|(x - m) - y\|^2 = \|x - m\|^2 + \|y\|^2 \geq \|x - m\|^2,$$

and thus $\text{dist}\,(x, \mathcal{M}) \geq \|x - m\|$. Since the other inequality is obvious, (a) holds.

□

7.16. Remarks.

(a) Given any non-empty subset $\mathcal{S} \subseteq \mathcal{H}$, let

$$\mathcal{S}^\perp := \{y \in \mathcal{H} : \langle x, y\rangle = 0 \text{ for all } x \in \mathcal{S}\}.$$

It is routine to show that $\mathcal{S}^\perp$ is a norm-closed subspace of $\mathcal{H}$. In particular,

$$\left(\mathcal{S}^\perp\right)^\perp \supseteq \overline{\text{span}}\,\mathcal{S},$$

the norm closure of the linear span of $\mathcal{S}$.

(b) Recall from Linear Algebra that if $\mathcal{V}$ is a vector space and $\mathcal{W}$ is a (vector) subspace of $\mathcal{V}$, then there exists a (vector) subspace $\mathcal{X} \subseteq \mathcal{V}$ such that

(i) $\mathcal{W} \cap \mathcal{X} = \{0\}$, and

(ii) $\mathcal{V} = \mathcal{W} + \mathcal{X} := \{w + x : w \in \mathcal{W}, y \in \mathcal{X}\}$.

We say that $\mathcal{W}$ is **algebraically complemented** by $\mathcal{X}$. The existence of such a $\mathcal{X}$ for each $\mathcal{W}$ says that every vector subspace of a vector space is algebraically complemented. We shall write $\mathcal{V} = \mathcal{W} \dot{+} \mathcal{X}$ to denote the fact that $\mathcal{X}$ is an algebraic complement for $\mathcal{W}$ in $\mathcal{V}$.

If $\mathfrak{X}$ is a Banach space and $\mathfrak{Y}$ is a closed subspace of $\mathfrak{X}$, we say that $\mathfrak{Y}$ is **topologically complemented** if there exists a *closed* subspace $\mathfrak{Z}$ of $\mathfrak{X}$ such that $\mathfrak{Z}$ is an algebraic complement to $\mathfrak{Y}$. The issue here is that both $\mathfrak{Y}$ and $\mathfrak{Z}$ must be *closed* subspaces. It can be shown that the closed subspace

c_0 of ℓ_∞ is *not* topologically complemented in ℓ_∞. This result is known as Phillips' Theorem (see the paper of R. Whitley [**7**] for a short but elegant proof). We shall write $\mathfrak{X} = \mathfrak{Y} \oplus \mathfrak{Z}$ if $\mathfrak{Z}$ is a topological complement to $\mathfrak{Y}$ in $\mathfrak{X}$.

Now let $\mathcal{H}$ be a Hilbert space and let $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace of $\mathcal{H}$. We claim that $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$. Indeed, if $z \in \mathcal{M} \cap \mathcal{M}^\perp$, then $\|z\|^2 = \langle z, z\rangle = 0$, so $z = 0$. Also, if $x \in \mathcal{H}$, then we may let $m_1 \in \mathcal{M}$ be the element satisfying

$$\|x - m_1\| = \operatorname{dist}(x, \mathcal{M}).$$

The existence of m_1 is guaranteed by Theorem 7.14. By Theorem 7.15, $m_2 := x - m_1$ lies in $\mathcal{M}^\perp$, and so $x = m_1 + m_2 \in \mathcal{M} + \mathcal{M}^\perp$. Since $\mathcal{M}$ and $\mathcal{M}^\perp$ are closed subspaces of the Hilbert space $\mathcal{H}$ and they are algebraic complements, we are done.

In this case, the situation is even stronger. Given a Banach space $\mathfrak{X}$ and a topologically complemented closed subspace $\mathfrak{Y}$ of $\mathfrak{X}$, there is in general no reason to expect a *unique* topological complement for $\mathfrak{Y}$. For example, if $\mathfrak{X} = \mathbb{R}^2$ (equipped with your favourite norm – say $\|\cdot\|_\infty$), and if $\mathfrak{Y}$ denotes the x-axis in $\mathfrak{X}$, then any line passing through the origin and not equal to the x-axis describes a closed subspace $\mathfrak{Z}$ which is a topological complement to $\mathfrak{Y}$. In our case, however, the space $\mathcal{M}^\perp$ above is unique in that it is an **orthogonal complement**. That is, as well as being a topological complement to $\mathcal{M}$, every vector in $\mathcal{M}^\perp$ is orthogonal to every vector in $\mathcal{M}$.

(c) With $\mathcal{M}$ as in (b), we have $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$, so that if $x \in \mathcal{H}$, then we may write $x = m_1 + m_2$ with $m_1 \in \mathcal{M}$, $m_2 \in \mathcal{M}^\perp$ in a *unique* way. Consider the map:

$$\begin{array}{rccc} P: & \mathcal{H} & \to & \mathcal{M} \oplus \mathcal{M}^\perp \\ & x & \mapsto & m_1, \end{array}$$

relative to the above decomposition of x. It is elementary to verify that P is linear, and that P is **idempotent**, i.e., $P = P^2$. We remark in passing that $m_2 = (I - P)x$, and that $(I - P)^2 = (I - P)$ as well.

In fact, for $x \in \mathcal{H}$, $\|x\|^2 = \|m_1\|^2 + \|m_2\|^2$ by the Pythagorean Theorem, and so $\|Px\| = \|m_1\| \le \|x\|$, from which it follows that $\|P\| \le 1$. If $\mathcal{M} \ne \{0\}$, then choose $m \in \mathcal{M}$ with $\|m\| \ne 0$. Then $Pm = m$, and so $\|P\| \ge 1$. Combining these estimates, $\mathcal{M} \ne 0$ implies $\|P\| = 1$.

We refer to the map P as the **orthogonal projection** of $\mathcal{H}$ onto $\mathcal{M}$. The map $Q := (I - P)$ is the orthogonal projection onto $\mathcal{M}^\perp$, and we leave it to the reader to verify that if $\mathcal{M} \ne \mathcal{H}$, then $\|Q\| = 1$.

(d) Let $\varnothing \ne \mathcal{S} \subseteq \mathcal{H}$. We saw in (a) that $\mathcal{S}^{\perp\perp} \supseteq \overline{\operatorname{span}}\, \mathcal{S}$. In fact, if we let $\mathcal{M} = \overline{\operatorname{span}}\, \mathcal{S}$, then $\mathcal{M}$ is a closed subspace of $\mathcal{H}$, and so by (b),

$$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp.$$

It is routine to check that $\mathcal{S}^\perp = \mathcal{M}^\perp$. Suppose that there exists $0 \ne x \in \mathcal{S}^{\perp\perp}$, $x \notin \mathcal{M}$. Then $x \in \mathcal{H}$, and so we can write $x = m_1 + m_2$ with $m_1 \in \mathcal{M}$, and

$m_2 \in \mathcal{M}^\perp = \mathcal{S}^\perp$ ($m_2 \neq 0$, otherwise $x \in \mathcal{M}$). But then $0 \neq m_2 \in \mathcal{S}^\perp$ and so

$$\begin{aligned}\langle m_2, x\rangle &= \langle m_2, m_1\rangle + \langle m_2, m_2\rangle \\ &= 0 + \|m_2\|^2 \\ &\neq 0.\end{aligned}$$

This contradicts the fact that $x \in \mathcal{S}^{\perp\perp}$. It follows that $\mathcal{S}^{\perp\perp} = \overline{\operatorname{span}}\,\mathcal{S}$.

7.17. Lemma. *Let $\mathcal{H}$ be a Hilbert space over $\mathbb{K}$ and suppose that $\mathcal{M} \subseteq \mathcal{H}$ is a finite-dimensional linear manifold in $\mathcal{H}$. Then $\mathcal{M}$ is norm-closed, and therefore a subspace of $\mathcal{H}$.*

Proof. The proof of this Lemma is left to the Assignments.

□

7.18. Proposition. *Suppose that $\mathcal{M}$ is a finite-dimensional subspace of a Hilbert space $\mathcal{H}$ over $\mathbb{K}$. Suppose that $1 \le N$ is an integer and that $\{e_1, e_2, \ldots, e_N\}$ is an* ONB *for $\mathcal{M}$. If P is the orthogonal projection of $\mathcal{H}$ onto $\mathcal{M}$, then*

$$Px = \sum_{n=1}^{N} \langle x, e_n\rangle e_n, \qquad x \in \mathcal{H}.$$

Proof. Suppose that $\mathcal{M}$ admits an orthonormal basis $\{e_k\}_{k=1}^n$. Let $x \in \mathcal{H}$, and let P denote the orthogonal projection onto $\mathcal{M}$. By (b), Px is the unique element of $\mathcal{M}$ so that $x - Px$ lies in $\mathcal{M}^\perp$. Consider the vector $w = \sum_{k=1}^n \langle x, e_k\rangle e_k$. Then

$$\begin{aligned}\langle x - w, e_j\rangle &= \langle x, e_j\rangle - \sum_{k=1}^{n} \langle \langle x, e_k\rangle e_k, e_j\rangle \\ &= \langle x, e_j\rangle - \sum_{k=1}^{n} \langle x, e_k\rangle \langle e_k, e_j\rangle \\ &= \langle x, e_j\rangle - \langle x, e_j\rangle \, \|e_j\|^2 \\ &= 0.\end{aligned}$$

It follows that $x - w \in \mathcal{M}^\perp$, and thus $w = Px$. That is, $Px = \sum_{k=1}^n \langle x, e_k\rangle e_k$.

□

7.19. Theorem. ***Bessel's Inequality***

If $\{e_n\}_{n=1}^\infty$ is an orthonormal set in a Hilbert space $\mathcal{H}$, then for each $x \in \mathcal{H}$,

$$\sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2 \le \|x\|^2.$$

Proof. For each $N \ge 1$, let P_N denote the orthogonal projection of $\mathcal{H}$ onto span $\{e_n\}_{n=1}^N$. Given $x \in \mathcal{H}$, we have seen that $\|P_N\| \le 1$, and that $P_N x = \sum_{n=1}^N \langle x, e_n\rangle e_n$. Hence

$$\sum_{n=1}^{N} |\langle x, e_n\rangle|^2 = \|P_N x\|^2 \le \|x\|^2$$

for all $N \ge 1$, from which the result follows.

□

Before stating our next result, we recall the following from your previous Analysis course.

7.20. Lemma. *Let (X, d) be a separable metric space. Let $\delta > 0$ be a real number and $Y \subseteq X$ be a set with the property that $y, z \in Y$ with $y \neq z$ implies that $d(y, z) \geq \delta$. Then Y is countable.*

Proof. The proof is left to the exercises.

□

7.21. Theorem. *Let $\mathcal{H}$ be a separable Hilbert space, and suppose that $\mathcal{E} \subseteq \mathcal{H}$ be an orthonormal set. Then $\mathcal{E}$ is countable, say $\mathcal{E} = \{e_n\}_{n=1}^\infty$, and if $x \in \mathcal{H}$, then*

$$\sum_{n=1}^\infty \langle x, e_n \rangle e_n$$

converges in $\mathcal{H}$.

Proof. First note that if $x, y \in \mathcal{E}$ with $x \neq y$, then

$$\|x - y\| = \langle x - y, x - y \rangle^{1/2} = \left(\|x\|^2 + \|y\|^2 \right)^{1/2} = \sqrt{2}.$$

By Lemma 7.20 above, $\mathcal{E}$ must be countable. Thus we may write $\mathcal{E} = \{e_n\}_{n=1}^\infty$.

Let $x \in \mathcal{H}$ and $\varepsilon > 0$. For each $N \geq 1$, set

$$y_N = \sum_{n=1}^N \langle x, e_n \rangle e_n.$$

(The astute reader will note that with $\mathcal{M}_N := \text{span}\{e_1, e_2, \ldots, e_N\}$ and P_N defined as the orthogonal projection of $\mathcal{H}$ onto $\mathcal{M}_N$, we have that $y_N := P_N x$, $N \geq 1$.

It results from Bessel's Inequality that we can find $N_0 > 0$ so that

$$\sum_{k=N_0+1}^\infty |\langle x, e_k \rangle|^2 < \varepsilon^2.$$

If $M \geq N \geq N_0$, then (by the Pythagorean Theorem),

$$\begin{aligned} \|y_M - y_N\|^2 &= \|\sum_{n=N+1}^M \langle x, e_n \rangle e_n\|^2 \\ &= \sum_{n=N+1}^M |\langle x, e_n \rangle|^2 \\ &\leq \sum_{n=N_0+1}^\infty |\langle x, e_n \rangle|^2 \\ &< \varepsilon^2. \end{aligned}$$

This shows that $(y_N)_{N=1}^\infty$ is a Cauchy sequence. Since $\mathcal{H}$ is complete, this Cauchy sequence converges, i.e.

$$\sum_{n=1}^\infty \langle x, e_n\rangle e_n = \lim_{N\to\infty} y_N \in \mathcal{H}.$$

□

7.22. Theorem. *Let $\mathcal{E} = \{e_n\}_{n=1}^\infty$ be an orthonormal set in an infinite-dimensional, separable Hilbert space $\mathcal{H}$. The following are equivalent:*

(a) *The set $\mathcal{E}$ is an* ONB *for $\mathcal{H}$; i.e. $\mathcal{E}$ is a maximal orthonormal set in $\mathcal{H}$.*
(b) $\overline{\operatorname{span}}\,\mathcal{E} = \mathcal{H}$.
(c) *For all $x \in \mathcal{H}$, $x = \sum_{n=1}^\infty \langle x, e_n\rangle e_n$.*
(d) *For all $x \in \mathcal{H}$, $\|x\|^2 = \sum_{n=1}^\infty |\langle x, e_n\rangle|^2$.* [***Parceval's Identity***]

Proof.

(a) implies (b): Let $\mathcal{M} = \overline{\operatorname{span}}\,\mathcal{E}$. If $\mathcal{M} \ne \mathcal{H}$, then $\mathcal{M}^\perp \ne \{0\}$, so we can find $z \in \mathcal{M}^\perp$, $\|z\| = 1$. But then $\mathcal{E} \cup \{z\}$ is an orthonormal set, contradicting the maximality of $\mathcal{E}$.

(b) implies (c): Let $y = \sum_{n=1}^\infty \langle x, e_n\rangle e_n$, which exists by Theorem 7.21.

A routine calculation shows that $\langle y - x, e_n\rangle = 0$ for all $n \ge 1$, so $y - x$ is orthogonal to $\mathcal{M} = \overline{\operatorname{span}}\,\mathcal{E} = \mathcal{H}$. In particular, $y - x$ is orthogonal to $y - x \in \mathcal{H}$, so that $y - x = 0$, i.e. $y = x$.

(c) implies (d): Let $x \in \mathcal{H}$. By hypothesis, $x = \sum_{n=1}^\infty \langle x, e_n\rangle e_n$.

Thus

$$\begin{aligned}\|x\|^2 &= \langle \sum_{n=1}^\infty \langle x, e_n\rangle e_n, \sum_{k=1}^\infty \langle x, e_k\rangle e_k\rangle \\ &= \sum_{n=1}^\infty \sum_{k=1}^\infty \langle x, e_n\rangle\ \overline{\langle x, e_k\rangle}\ |\langle e_n, e_k\rangle| \qquad \text{[Check!]} \\ &= \sum_{n=1}^\infty |\langle x, e_n\rangle|^2.\end{aligned}$$

(d) implies (a): If $e \perp e_n$ for all $n \ge 1$, then by hypothesis,

$$\|e\|^2 = \sum_{n=1}^\infty |\langle e, e_n\rangle|^2 = 0,$$

so that $\mathcal{E}$ is a maximal orthonormal set in $\mathcal{H}$, i.e. $\mathcal{E}$ is an ONB in $\mathcal{H}$.

□

The appropriate notion of isomorphism in the category of Hilbert spaces involve linear maps that preserve the inner product.

7.23. Definition. *Two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ are said to be* ***isomorphic*** *if there exists a linear bijection $U : \mathcal{H}_1 \to \mathcal{H}_2$ so that*

$$\langle Ux, Uy\rangle = \langle x, y\rangle$$

for all $x, y \in \mathcal{H}_1$. We write $\mathcal{H}_1 \simeq \mathcal{H}_2$ to denote this isomorphism.

We also refer to the linear maps implementing the above isomorphism as **unitary operators**. Note that

$$\|Ux\|^2 = \langle Ux, Ux\rangle = \langle x, x\rangle = \|x\|^2$$

for all $x \in \mathcal{H}_1$, so that unitary operators are isometries. Moreover, the inverse map $U^{-1} : \mathcal{H}_2 \to \mathcal{H}_1$ defined by $U^{-1}(Ux) := x$ is also linear, and

$$\langle U^{-1}(Ux)U^{-1}(Uy)\rangle = \langle x, y\rangle = \langle Ux, Uy\rangle,$$

so that U^{-1} is also a unitary operator.

Note that if $\mathcal{L} \subseteq \mathcal{H}_1$ is a closed subspace, then $\mathcal{L}$ is complete, whence $U\mathcal{L}$ is also complete and hence closed in $\mathcal{H}_2$.

7.24. Theorem. *Any two separable, infinite-dimensional Hilbert spaces over* $\mathbb{K}$ *are isomorphic.*

Proof. The proof is left to the Assignments.

□

Appendix to Section 7.

7.25. As mentioned in the Appendix to Chapter 6, given $E \in \mathfrak{M}(\mathbb{R})$ and $1 < p < \infty$, one may identify the dual space of $L_p(E,\mathbb{R})$ with $L_q(E,\mathbb{R})$ via a linear, isometric isomorphism. Since the Lebesgue conjugate of $p = 2$ is $q = 2$, this gives us an identification of the dual space of $L_2(E,\mathbb{R})$ with itself, via a linear, isometric isomorphism.

Once again, the case where $p = 2$ is special. In this case, we obtain a second identification of the dual of a Hilbert $\mathcal{H}$ space with itself. Given $y \in \mathcal{H}$, we define the linear function $\Phi_y \in \mathcal{H}^*$ by setting

$$\Phi_y(x) = \langle x, y\rangle, \qquad x \in \mathcal{H}.$$

As we shall see in the Assignments, the map

$$\begin{array}{rccc} \varrho : & \mathcal{H} & \to & \mathcal{H}^* \\ & y & \mapsto & \Phi_y \end{array}$$

is a *conjugate-linear* isometric isomorphism. Here, **conjugate-linear** means that $\varrho(y+kz) = \varrho(y) + \overline{k}\varrho(z)$ for all $y, z \in \mathcal{H}$ and $k \in \mathbb{K}$. Of course, when $\mathbb{K} = \mathbb{R}$, ϱ is linear.

There are reasons for preferring this identification of $\mathcal{H}^*$ with $\mathcal{H}$ to the identification that we obtained in the Appendix to Chapter 6. Given Banach spaces $\mathfrak{X}$ and $\mathfrak{Y}$ and an operator $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$, we may define an operator $T^* \in \mathcal{B}(\mathfrak{Y}^*, \mathfrak{X}^*)$ via $T^*y^*(x) = y^*(Tx)$ for all $x \in \mathfrak{X}$, $y^* \in \mathfrak{Y}^*$. The identification of $\mathcal{H}$ with $\mathcal{H}^*$ mentioned in the paragraph above provides us with an **involution** on $\mathcal{B}(\mathcal{H})$, that is, a map $* : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ which satisfies

- $(T^*)^* = T$ for all $T \in \mathcal{B}(\mathcal{H})$;
- $(T_1 + kT_2)^* = T_1^* + \overline{k}T_2^*$ for all $T_1, T_2 \in \mathcal{B}(\mathcal{H})$, $k \in \mathbb{K}$;
- $(T_1T_2)^* = T_2^*T_1^*$ for all $T_1, T_2 \in \mathcal{B}(\mathcal{H})$.

Moreover, one may check that $\|T^*\| = \|T\|$ for all $T \in \mathcal{B}(\mathcal{H})$, and finally that

$$\|T^*T\| = \|T\|^2 \qquad \text{for all } T \in \mathcal{B}(\mathcal{H}).$$

Thus $\mathcal{B}(\mathcal{H})$ becomes what is called an **involutive Banach algebra** using this operation $*$, and the last equality is referred to as the **C^*-equation**. Involutive Banach algebras whose involution satisfies the C^*-equation are called **C^*-algebras**. The C^*-equation, as anodyne as it might at first appear, has untold consequences for the structure and representation theory of the algebra. People have spent their lives studying C^*-algebras. If you haven't been moved to pity yet, your heart is made of stone.

7.26. Theorem. ***The Riesz Representation Theorem.*** *Let $\{0\} \neq \mathcal{H}$ be a Hilbert space over $\mathbb{K}$, and let $\varphi \in \mathcal{H}^*$. Then there exists a unique vector $y \in \mathcal{H}$ so that*

$$\varphi(x) = \langle x, y\rangle \qquad \textit{for all } x \in \mathcal{H}.$$

Moreover, $\|\varphi\| = \|y\|$.

Proof. Given a fixed $y \in \mathcal{H}$, let us denote by β_y the map $\beta_y(x) = \langle x, y\rangle$. Our goal is to show that $\mathcal{H}^* = \{\beta_y : y \in \mathcal{H}\}$. First note that if $y \in \mathcal{H}$, then $\beta_y(kx_1 + x_2) =$

$\langle kx_1 + x_2, y\rangle = k\langle x_1, y\rangle + \langle x_2, y\rangle = k\beta_y(x_1) + \beta_y(x_2)$, and so β_y is linear. Furthermore, for each $x \in \mathcal{H}$, $|\beta_y(x)| = |\langle x, y\rangle| \le \|x\|\|y\|$ by the Cauchy-Schwarz Inequality. Thus $\|\beta_y\| \le \|y\|$, and hence β_y is continuous - i.e. $\beta_y \in \mathcal{H}^*$.

It is not hard to verify that the map

$$\begin{array}{rccc}\Theta: & \mathcal{H} & \to & \mathcal{H}^* \\ & y & \mapsto & \beta_y\end{array}$$

is conjugate-linear (if $\mathbb{K} = \mathbb{C}$), otherwise it is linear (if $\mathbb{K} = \mathbb{R}$). From the first paragraph, it is also contractive. But $[\Theta(y)](y) = \beta_y(y) = \langle y, y\rangle = \|y\|^2$, so that $\|\Theta(y)\| \ge \|y\|$ for all $y \in \mathcal{H}$, and Θ is isometric as well. It immediately follows that Θ is injective, and there remains only to prove that Θ is surjective.

Let $\varphi \in \mathcal{H}^*$. If $\varphi = 0$, then $\varphi = \Theta(0)$. Otherwise, let $\mathcal{M} = \ker\varphi$, so that $\operatorname{codim}\mathcal{M} = 1 = \dim \mathcal{M}^\perp$, since $\mathcal{H}/\mathcal{M} \simeq \mathbb{K} \simeq \mathcal{M}^\perp$. Choose $e \in \mathcal{M}^\perp$ with $\|e\| = 1$.

Let P denote the orthogonal projection of $\mathcal{H}$ onto $\mathcal{M}$, constructed as in Remark 7.16. Then, as $I - P$ is the orthogonal projection onto $\mathcal{M}^\perp$, and as $\{e\}$ is an orthonormal basis for $\mathcal{M}^\perp$, by Proposition 7.18, for all $x \in \mathcal{H}$, we have

$$x = Px + (I - P)x = Px + \langle x, e\rangle e.$$

Thus for all $x \in \mathcal{H}$,

$$\varphi(x) = \varphi(Px) + \langle x, e\rangle\varphi(e) = \langle x, \overline{\varphi(e)}e\rangle = \beta_y(x),$$

where $y := \overline{\varphi(e)}e$. Hence $\varphi = \beta_y$, and Θ is onto.

□

7.27. Remark. The fact that the map Θ defined in the proof the Riesz Representation Theorem above induces an isometric, conjugate-linear automorphism of $\mathcal{H}$ is worth remembering.

Exercises for Section 7.

Exercise 7.1. Let $\mathcal{H} = \ell^2$, equipped with standard inner product. For $n \geq 1$, let $e_n = (\delta_{n,k})_{k=1}^{\infty}$, where $\delta_{a,b}$ is the Kronecker delta function. Prove that $\{e_n\}_{n=1}^{\infty}$ is an ONB for $\mathcal{H}$.

8. Fourier analysis - an introduction

The word "genius" isn't applicable in football. A genius is a guy like Norman Einstein.

Joe Theismann

8.1. One of the main achievements of linear algebra is that – in particular when dealing with finite-dimensional vector spaces – one is able to reduce a great many questions about abstract linear maps to very concrete and computation-friendly questions about matrices. As we know, to each linear map T from an n-dimensional vector space $\mathcal{V}$ over a field $\mathbb{F}$ to a m-dimensional vector space $\mathcal{W}$ over $\mathbb{F}$ we may associate an $m \times n$ matrix $[T] \in \mathbb{M}_{m,n}(\mathbb{F})$ as follows: we let $\mathcal{B}_\mathcal{V} := \{v_1, v_2, \dots, v_n\}$ and $\mathcal{B}_\mathcal{W} = \{w_1, w_2, \dots, w_m\}$ be bases for $\mathcal{V}$ and $\mathcal{W}$ respectively. Given $x \in \mathcal{V}$, we may express x as a linear combination of the elements of $\mathcal{B}_\mathcal{V}$ in a unique way, say

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n,$$

where $\alpha_i \in \mathbb{F}$, $1 \le i \le n$. We write $[x]_{\mathcal{B}_\mathcal{V}}$ to denote the corresponding n-tuple $(\alpha_k)_{k=1}^n$. Recall that the map

$$\begin{array}{rccc} \Gamma_\mathcal{V}: & \mathcal{V} & \to & \mathbb{F}^n \\ & x & \mapsto & [x]_{\mathcal{B}_\mathcal{V}}, \end{array}$$

is an isomorphism of $\mathcal{V}$ and $\mathbb{F}^n$.

Similarly, every $y \in \mathcal{W}$ corresponds to a unique m-tuple $[y]_{\mathcal{B}_\mathcal{W}} = (\beta_j)_{j=1}^m \in \mathbb{F}^m$, so that $y = \sum_{j=1}^m \beta_j w_j$, and the map $\Gamma_\mathcal{W} : \mathcal{W} \to \mathbb{F}^m$ defined by $\Gamma_\mathcal{W}(y) = [y]_{\mathcal{B}_\mathcal{W}}$ is an isomorphism of vector spaces.

The matrix that we then associate to T is $[T] := [t_{ij}]$, where, for $1 \le j \le n$, $(t_{i,j})_{i=1}^m = [Tv_j]_{\mathcal{B}_\mathcal{W}}$. Obviously the matrix for T depends upon our choice of bases $\mathcal{B}_\mathcal{V}$ for $\mathcal{V}$ and $\mathcal{B}_\mathcal{W}$ for $\mathcal{W}$.

While T acts upon the elements of $\mathcal{V}$, the matrix $[T]$ acts upon the *coordinates* of vectors $x \in \mathcal{V}$, and returns the *coordinates* of the vector Tx; that is,

$$[T][x]_{\mathcal{B}_\mathcal{V}} = [Tx]_{\mathcal{B}_\mathcal{W}},$$

where $[x]_{\mathcal{B}_\mathcal{V}}$ (resp. $[Tx]_{\mathcal{B}_\mathcal{W}}$) represents the coordinates of x (resp. Tx) with respect to the basis $\mathcal{B}_\mathcal{V}$ for $\mathcal{V}$ (resp. $\mathcal{B}_\mathcal{W}$ for $\mathcal{W}$.)

In practice, finding the coordinates of a vector in a vector space $\mathcal{V}$ with respect to a basis for a given vector space often reduces to solving a system of linear equations. The situation is greatly simplified, however, if the vector space $\mathcal{H}$ is a finite-dimensional Hilbert space over the field $\mathbb{K}$, and the basis $\mathcal{E} := \{e_1, e_2, \dots, e_n\}$ in question is an *orthonormal* basis.

In this case, the coordinates of $x \in \mathcal{H}$ are deduced from the formula

$$x = \langle x, e_1 \rangle e_1 + \langle x, e_2 \rangle e_2 + \cdots + \langle x, e_n \rangle e_n.$$

Thus, for $T : \mathcal{H} \to \mathcal{H}$ linear, the matrix of T corresponding to the basis $\mathcal{E}$ (for both the domain and the codomain) is simply $[T] = [t_{ij}]$, where $t_{ij} = \langle Te_j, e_i \rangle$, $1 \le i, j \le n$.

8.2. When the vector spaces $\mathcal{V}$ and $\mathcal{W}$ above are infinite-dimensional, it is still possible to associate to each linear map $T : \mathcal{V} \to \mathcal{W}$ a (generalized) matrix as above; the problem now lies in the fact that the bases involved are obviously infinite, and there is no need for them to be countable. For example, it can be shown that if $(\mathfrak{X}, \|\cdot\|)$ is an infinite-dimensional Banach space, then any (Hamel) basis for $\mathfrak{X}$ must be uncountable. This greatly mitigates the usefulness of the (generalized) matrix representation of linear maps from $\mathfrak{X}$ to $\mathfrak{X}$.

Analysis, however, distinguishes itself from Algebra in its ubiquitous recourse to approximation. We may ask whether or not we can find a suitable subset $\mathcal{B}$ of $\mathfrak{X}$, preferably countable, such that every $x \in \mathfrak{X}$ can be *approximated* in norm by finite linear combinations of elements of $\mathcal{B}$.

In the case of separable Hilbert spaces, we have already seen that this is the case. When $\mathcal{H}$ is an infinite-dimensional separable Hilbert space over $\mathbb{K}$, we have seen that $\mathcal{H}$ admits a countable, maximal orthonormal set $\mathcal{E} = \{e_n\}_{n=1}^\infty$ (which we have dubbed an "*orthonormal basis*" for $\mathcal{H}$ – or a "*Hilbert space basis*" for $\mathcal{H}$), such that $x \in \mathcal{H}$ implies that

$$x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n := \lim_{N\to\infty} \sum_{n=1}^{N} \langle x, e_n \rangle e_n.$$

While $\mathcal{E}$ is certainly linearly independent, it is never a Hamel basis (see Exercise 1).

8.3. In particular, we saw in Theorem 7.7 that if $E \in \mathfrak{M}(\mathbb{R})$, then $L_2(E, \mathbb{C})$ is a Hilbert space over $\mathbb{C}$, when equipped with the inner product

$$\langle [f], [g] \rangle = \int_E f\overline{g},$$

and that $L_2([-\pi, \pi], \mathbb{C})$ is separable (see Corollary 6.38).

Let us now investigate a different, but closely related example. Consider the space

$$\mathcal{L}_2(\mathbb{T}, \mathbb{C}) := \{f : \mathbb{R} \to \mathbb{C} : f \text{ is measurable}, 2\pi\text{-periodic, and } \int_{[-\pi,\pi)} |f|^2 < \infty\}.$$

We leave it as an exercise for the reader to prove that $\mathcal{L}_2(\mathbb{T}, \mathbb{C})$ is a vector space and that the function

$$\begin{array}{rccc} \nu_2 : & \mathcal{L}_2(\mathbb{T}, \mathbb{C}) & \to & \mathbb{R} \\ & f & \mapsto & \left(\frac{1}{2\pi} \int_{[-\pi,\pi)} |f|^2 \right)^{1/2} \end{array}$$

defines a seminorm on $\mathcal{L}_2(\mathbb{T}, \mathbb{C})$. Set $\mathcal{N}(\mathbb{T}, \mathbb{C}) := \{f \in \mathcal{L}_2(\mathbb{T}, \mathbb{C}) : \nu_2(f) = 0\}$. Arguing as in Section 6.4, we see that $[f] = [g]$ in $L_2(\mathbb{T}, \mathbb{C}) := \mathcal{L}_2(\mathbb{T}, \mathbb{C})/\mathcal{N}(\mathbb{T}, \mathbb{C})$ if and only if $f = g$ a.e. on $\mathbb{R}$, or equivalently $f = g$ a.e. on $[-\pi, \pi)$ (given that f, g are 2π-periodic on $\mathbb{R}$). By Proposition 6.3, we obtain a norm on $L_2(\mathbb{T}, \mathbb{C})$ by setting $\|[f]\|_2 := \nu_2(f)$, $[f] \in L_2(\mathbb{T}, \mathbb{C})$.

Furthermore, the function

$$\begin{array}{rccc} \langle \cdot, \cdot \rangle_{\mathbb{T}} : & L_2(\mathbb{T}, \mathbb{C}) \times L_2(\mathbb{T}, \mathbb{C}) & \to & \mathbb{C} \\ & ([f], [g]) & \mapsto & \frac{1}{2\pi} \int_{[-\pi,\pi)} f\overline{g} \end{array}$$

defines an inner product on $L_2(\mathbb{T},\mathbb{C})$, and $\|\cdot\|_2$ is precisely the norm induced by this inner product. *A fortiori*, $L_2(\mathbb{T},\mathbb{C})$ is complete with respect to this norm, and is therefore a Hilbert space. Finally, we leave it to the reader to verify that with $\xi_n(\theta) = e^{in\theta}$, $\theta \in \mathbb{R}$, $n \in \mathbb{Z}$, the set $\{[\xi_n]\}_{n\in\mathbb{Z}}$ forms an ONB for $L_2(\mathbb{T},\mathbb{C})$.

Given $n \in \mathbb{Z}$, we shall refer to the complex number $\alpha_n^{[f]} := \langle [f], [\xi_n] \rangle_{\mathbb{T}}$ as the n^{th}-**Fourier coefficient** of $[f]$ relative to the ONB $([\xi_n])_{n\in\mathbb{Z}}$.

As we shall see in the Assignments, the map

$$\begin{array}{rcll} U: & L_2(\mathbb{T},\mathbb{C}) & \mapsto & \ell_2(\mathbb{Z},\mathbb{C}) \\ & [f] & \mapsto & (\alpha_n^{[f]})_{n\in\mathbb{Z}} \end{array}$$

defines a unitary operator from the Hilbert space $L_2(\mathbb{T},\mathbb{C})$ to $\ell_2(\mathbb{Z},\mathbb{C})$. In particular, therefore, it is injective. Thus, if $[f],[g] \in L_2(\mathbb{T},\mathbb{C})$ and $\alpha_n^{[f]} = \alpha_n^{[g]}$ for all $n \in \mathbb{Z}$, then $f = g$ a.e. on $\mathbb{R}$. In other words, an element $[f] \in L_2(\mathbb{T},\mathbb{C})$ *is entirely determined by its Fourier coefficients.* Moreover, given any sequence $(\beta_n)_{n\in\mathbb{Z}} \in \ell_2(\mathbb{Z},\mathbb{C})$, there exists $[f] \in L_2(\mathbb{T},\mathbb{C})$ such that $\alpha_n^{[f]} = \beta_n$, $n \in \mathbb{Z}$.

Let $[f] \in L_2(\mathbb{T},\mathbb{C})$. For each $1 \le N \in \mathbb{N}$, set

$$\Delta_N([f]) = \sum_{n=-N}^{N} \alpha_n^{[f]} [\xi_n].$$

We shall say that $\Delta_N([f])$ is the N^{th} partial sum of the **Fourier series** of $[f]$. It follows from Theorem 7.22 that

$$[f] = \lim_{N\to\infty} \Delta_N([f]),$$

where the convergence is relative to the $\|\cdot\|_2$-norm defined above.

This is an entirely satisfactory state of affairs, and demonstrates clearly just how well-behaved Hilbert spaces are. Our next goal is to try to extend this theory to the L_1-setting. Here, we will quickly discover that things are far more complicated, and for that reason perhaps also that much more interesting!

8.4. In trying to extend this theory beyond the Hilbert space setting, we shall once again require some notations and definitions. We define:

- $\mathrm{Trig}(\mathbb{T},\mathbb{C}) := \mathrm{span}\{\xi_n : n \in \mathbb{Z}\} = \{\sum_{n=-N}^{N} \alpha_n \xi_n : \alpha_n \in \mathbb{C}, 1 \le N \in \mathbb{N}\}$;
- $\mathcal{C}(\mathbb{T},\mathbb{C}) := \{f : \mathbb{R} \to \mathbb{C} : f \text{ is continuous and } 2\pi\text{-periodic}\}$;
- $\mathrm{SIMP}(\mathbb{T},\mathbb{C}) := \{f : \mathbb{R} \to \mathbb{C} : f|_{[-\pi,\pi)} \text{ is a simple function and } f \text{ is } 2\pi\text{-periodic}\}$;
- $\mathrm{STEP}(\mathbb{T},\mathbb{C}) := \{f : \mathbb{R} \to \mathbb{C} : f|_{[-\pi,\pi)} \text{ is a step function and } f \text{ is } 2\pi\text{-periodic}\}$;
- for $1 \le p < \infty$,

  $$\mathcal{L}_p(\mathbb{T},\mathbb{C}) := \{f : \mathbb{R} \to \mathbb{C} : f \text{ is measurable}, 2\pi\text{-periodic, and } \int_{[-\pi,\pi)} |f|^p < \infty\};$$
- and for $p = \infty$,

  $$\mathcal{L}_\infty(\mathbb{T},\mathbb{C}) = \{f : \mathbb{R} \to \mathbb{C} : f \text{ is measurable}, 2\pi\text{-periodic, and essentially bounded}\}.$$

Observe that

$$\mathrm{Trig}(\mathbb{T},\mathbb{C}) \subseteq \mathcal{C}(\mathbb{T},\mathbb{C}) \subseteq \mathcal{L}_p(\mathbb{T},\mathbb{C}), \quad 1 \le p \le \infty.$$

As was the case with $p = 2$ above, for each $1 \le p < \infty$, the set $\mathcal{L}_p(\mathbb{T}, \mathbb{C})$ forms a vector space over $\mathbb{C}$, and the map

$$\begin{array}{rccc} \nu_p : & \mathcal{L}_p(\mathbb{T}, \mathbb{C}) & \to & \mathbb{R} \\ & f & \mapsto & \left(\frac{1}{2\pi} \int_{[-\pi,\pi)} |f|^p \right)^{1/p} \end{array}$$

defines a seminorm on $\mathcal{L}_p(\mathbb{T}, \mathbb{C})$.

Moreover, if $p = \infty$, we repeat the arguments of Chapter 6. That is, for $f \in \mathcal{L}_\infty(\mathbb{T}, \mathbb{C})$, we set

$$\nu_\infty(f) := \text{ess sup}(f) := \inf\{\delta > 0 : m\{\theta \in [-\pi, \pi) : |f(\theta)| > \delta\} = 0\},$$

and verify that this defines a seminorm on $\mathcal{L}_\infty(\mathbb{T}, \mathbb{C})$.

Appealing to Proposition 6.3, for each $1 \le p \le \infty$, we obtain a norm $\|\cdot\|_p$ on $L_p(\mathbb{T}, \mathbb{C}) := \mathcal{L}_p(\mathbb{T}, \mathbb{C}) / \mathcal{N}_p(\mathbb{T}, \mathbb{C})$, where $\mathcal{N}_p(\mathbb{T}, \mathbb{C}) := \{f \in \mathcal{L}_p(\mathbb{T}, \mathbb{C}) : \nu_p(f) = 0\}$. Perhaps unsurprisingly, we find that $[f] = [g] \in L_p(\mathbb{T}, \mathbb{C})$ if and only if $f = g$ a.e. on $\mathbb{R}$ (equivalently $f = g$ a.e. on $[-\pi, \pi)$, because of 2π-periodicity). The details are left to the reader (see the Exercises below).

It is also a simple but important exercise (which we again leave to the reader) to verify that for $f \in \mathcal{C}(\mathbb{T}, \mathbb{C})$,

$$\|[f]\|_\infty = \|f\|_{\sup} := \sup\{|f(\theta)| : -\pi \le \theta < \pi\}.$$

Note that the supremum on the right hand side of this equation exists as a finite number since $f \in \mathcal{C}(\mathbb{T}, \mathbb{C})$ implies that f is continuous on $\mathbb{R}$, and hence f is bounded on $[-\pi, \pi] \supseteq [-\pi, \pi)$.

Given any function $f : [-\pi, \pi) \to \mathbb{C}$, denote by $\check{f} : \mathbb{R} \to \mathbb{C}$ its 2π-periodic extension, that is: $\check{f}(\theta) = f(\theta)$, $\theta \in [-\pi, \pi)$, and $\check{f}(\theta + 2\pi) = \check{f}(\theta)$, $\theta \in \mathbb{R}$. Clearly $\check{f}$ always exists and is uniquely defined by f.

8.5. Theorem. *Let $1 \le p \le \infty$. The map*

$$\begin{array}{rccc} \Xi_p : & L_p([-\pi, \pi), \mathbb{C}) & \to & L_p(\mathbb{T}, \mathbb{C}) \\ & [f] & \mapsto & [\check{f}] \end{array}$$

is an isometric isomorphism.

Proof. See the Exercises.

□

It follows that all of our favourite results from the previous Chapters hold in $L_p(\mathbb{T}, \mathbb{C})$; for example, Hölder's Inequality holds in $L_p(\mathbb{T}, \mathbb{C})$ and says that if $1 \le p < \infty$ and q is the Lebesgue conjugate of p, then for $[f] \in L_p(\mathbb{T}, \mathbb{C})$ and $[g] \in L_q(\mathbb{T}, \mathbb{C})$ we have that $[fg] \in L_1(\mathbb{T}, \mathbb{C})$ and

$$\|[fg]\|_1 \le \|[f]\|_p \, \|[g]\|_q.$$

A second result which carries over is that $[\text{Simp}(\mathbb{T}, \mathbb{C})]$ is dense in $L_p(\mathbb{T}, \mathbb{C})$ (in the $\|\cdot\|_p$-norm) for all $1 \le p \le \infty$. Again, the explicit proof of this is left to the exercises. Note that the fact that $[\text{Step}(\mathbb{T}, \mathbb{C})]$ and $[\mathcal{C}(\mathbb{T}, \mathbb{C})]$ are dense in $L_p(\mathbb{T}, \mathbb{C})$ for all $1 \le p < \infty$ is not an immediate consequence of Theorem 8.5, since, for example, a function f which lies in $\mathcal{C}(\mathbb{T},\mathbb{C})$ must satisfy $f(-\pi) = \lim_{\theta\to\pi} f(\theta)$. Despite this, $[\mathrm{STEP}(\mathbb{T},\mathbb{C})]$ and $[\mathcal{C}(\mathbb{T},\mathbb{C})]$ *are* dense in $L_p(\mathbb{T},\mathbb{C})$ for all such p, and this is also left to the exercises. These facts will prove useful below.

Of course, this raises the question of why we even bother with $L_p(\mathbb{T},\mathbb{C})$, given that it is isomorphic to $L_p([-\pi,\pi),\mathbb{C})$. This would be a good time for the reader to consult the Appendix, even if this is not something the reader typically does.

8.6. Definition. *For $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ and $n \in \mathbb{Z}$, we refer to*

$$\widehat{f}(n) := \frac{1}{2\pi}\int_{[-\pi,\pi)} f\overline{\xi_n}$$

*as the n^{th}-**Fourier coefficient** of f. We also refer to*

$$\sum_{n\in\mathbb{Z}} \widehat{f}(n)\,\xi_n$$

*as the **Fourier series** of f in $\mathcal{L}_1(\mathbb{T},\mathbb{C})$.*

8.7. Remark. Observe that if $f,g \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ and $f = g$ a.e. on $[-\pi,\pi)$, then

$$\widehat{f}(n) = \widehat{g}(n) \qquad \text{for all } n \in \mathbb{Z}.$$

In other words, if we set n^{th}-**Fourier coefficient** of $[f] \in L_1(\mathbb{T},\mathbb{C})$ to be

$$\alpha_n^{[f]} := \widehat{f}(n), \quad n \in \mathbb{Z},$$

then this is well-defined. We then define

$$\sum_{n\in\mathbb{Z}} \alpha_n^{[f]}\,[\xi_n]$$

to be the **Fourier series** of $[f]$.

The perspicacious reader (hopefully you) will have observed that at no point have we said anything about convergence of the above series. Indeed, at this stage, the series notation is *strictly formal*, and is meant only as a shorthand to represent the sequence of partial sums $\left(\sum_{n=-N}^{N} \alpha_n^{[f]}[\xi_n]\right)_{N=0}^{\infty}$. The remainder of the course is devoted to examining in what sense the series (i.e. the sequence of partial sums) above converges.

We also point out that we may extend the notion of a Fourier coefficient to non-integer powers of $e^{i\theta}$. That is, for $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ and $r \in \mathbb{R}$, we define

$$\widehat{f}(r) = \frac{1}{2\pi}\int_{[-\pi,\pi)} f\overline{\xi_r},$$

where $\xi_r(\theta) = e^{ir\theta}$ for all $\theta \in \mathbb{R}$.

8.8. In the case $p = 2$, we have more than once seen that $(\alpha_n^{[f]})_{n\in\mathbb{Z}} \in \ell_2(\mathbb{Z}, \mathbb{C})$. While nothing so nice holds for $[f] \in L_1(\mathbb{T}, \mathbb{C})$, the situation is not altogether hopeless.

First note that $|\xi_r(\theta)| = 1$ for all $\theta \in \mathbb{R}$ and all $r \in \mathbb{R}$. As such, for $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$, we have

$$\begin{aligned} |\widehat{f}(r)| &= \left| \frac{1}{2\pi} \int_{[-\pi,\pi)} f\overline{\xi_r} \right| \\ &\le \frac{1}{2\pi} \int_{[-\pi,\pi)} |f\overline{\xi_r}| \\ &= \frac{1}{2\pi} \int_{[-\pi,\pi)} |f| \\ &= \nu_1(f) \\ &= \|[f]\|_1. \end{aligned}$$

As before, if $f, g \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$ and $f = g$ a.e., then $\widehat{f}(r) = \widehat{g}(r)$ for all $r \in \mathbb{R}$, and so we may define $\alpha_r^{[f]} := \widehat{f}(r)$, $r \in \mathbb{R}$. It trivially follows that $\sup_{r\in\mathbb{R}} |\alpha_r^{[f]}| \le \|[f]\|_1$ for all $[f] \in L_1(\mathbb{T}, \mathbb{C})$, and more specifically $(\alpha_n^{[f]})_{n\in\mathbb{Z}} \in \ell_\infty(\mathbb{Z}, \mathbb{C})$. In fact, we can do better.

8.9. Theorem. The Riemann-Lebesgue Lemma.
Let $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$. Then

$$\lim_{r\to\infty} \widehat{f}(r) = 0 = \lim_{r\to-\infty} \widehat{f}(r).$$

In particular,

$$(\alpha_n^{[f]})_{n\in\mathbb{Z}} \in c_0(\mathbb{Z}, \mathbb{C}).$$

Proof. The key to the proof of this result is to notice that it is really quite simple to prove when $f|_{[-\pi,\pi)}$ is the characteristic function of an interval. But Lebesgue integration is linear, and the span of these (2π-periodic extensions of) characteristic functions of intervals is $\textsc{Step}(\mathbb{T}, \mathbb{C})$. So the result will hold for this class as well. Given this, one simply appeals to the density of $[\textsc{Step}(\mathbb{T}, \mathbb{C})]$ in $L_1(\mathbb{T}, \mathbb{C})$ to obtain the full result. As a rule, life is not great, but some parts of it really are.

- Suppose first that f_0 is the characteristic function of an interval, say

  $$f_0 = \chi_{[s,t]},$$

  where $-\pi \le s < t < \pi$. Let $f := \check{f}_0$ denote the 2π-periodic extension of f_0 to $\mathbb{R}$, so that $f \in \textsc{Step}(\mathbb{T}, \mathbb{C})$. Then, keeping in mind that every bounded, Riemann-integrable function over a bounded interval is Lebesgue integral, and that the Lebesgue and Riemann integrals coincide (see Theorem 5.24),

$$\begin{aligned}\widehat{f}(r) &= \frac{1}{2\pi}\int_{[-\pi,\pi)} \chi_{[s,t]}\overline{\xi_r}\\ &= \frac{1}{2\pi}\int_{[s,t]} e^{-ir\theta}\\ &= \frac{1}{2\pi}\int_s^t e^{-ir\theta}d\theta\\ &= \frac{1}{2\pi}\left.\frac{e^{-ir\theta}}{-ir}\right]_{\theta=s}^{\theta=t}\\ &= \frac{1}{2\pi}\left[\frac{e^{-irt}-e^{-irs}}{-ir}\right].\end{aligned}$$

From this it easily follows that

$$|\widehat{f}(r)| \le \frac{2}{2\pi|r|} = \frac{1}{\pi|r|},$$

whence

$$\lim_{r\to\infty}\widehat{f}(r) = 0 = \lim_{r\to-\infty}\widehat{f}(r).$$

- Next, suppose that $f \in \text{STEP}(\mathbb{T},\mathbb{C})$. Let $f_0 := f|_{[-\pi,\pi)}$, and write $f_0 = \sum_{k=1}^M \beta_k\chi_{H_k}$ as a disjoint representation, where each $H_k = [s_k, t_k]$ is a subinterval of $[-\pi,\pi)$.

  The result is now a simple consequence of the argument above, combined with the linearity of the Lebesgue integral, and is left to the reader.
- Finally, let $[f] \in L_1(\mathbb{T},\mathbb{C})$, $\varepsilon > 0$ be arbitrary, and choose $g \in \text{STEP}(\mathbb{T},\mathbb{C})$ such that $\|[f]-[g]\|_1 < \varepsilon/2$.

  Then

$$\begin{aligned}\widehat{f}(r) &= \frac{1}{2\pi}\int_{[-\pi,\pi)} f\overline{\xi_r}\\ &= \frac{1}{2\pi}\int_{[-\pi,\pi)}(f-g)\overline{\xi_r} + \frac{1}{2\pi}\int_{[-\pi,\pi)} g\overline{\xi_r}\\ &= \widehat{f-g}(r) + \widehat{g}(r).\end{aligned}$$

But as we have seen, $|\widehat{f-g}(r)| \le \nu_1(f-g) = \|[f-g]\|_1 = \|[f]-[g]\|_1 < \varepsilon/2$ for all $r \in \mathbb{R}$. Since $g \in \text{STEP}(\mathbb{T},\mathbb{C})$, from the previous case, we see that we may choose $N > 0$ such that $|r| \ge N$ implies that $|\widehat{g}(r)| < \varepsilon/2$, and thus $|r| \ge N$ implies that

$$|\widehat{f}(r)| \le |\widehat{f-g}(r)| + |\widehat{g}(r)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves that $\lim_{r\to\infty}\widehat{f}(r) = 0 = \lim_{r\to-\infty}\widehat{f}(r)$, completing the proof of the first statement of the Theorem. The second statement is an easy consequence of the first, and is left to the reader.

□

8.10. We began by recalling that $[f] \in L_2(\mathbb{T}, \mathbb{C})$ if and only if $(\alpha_n^{[f]})_{n\in\mathbb{Z}} \in \ell_2(\mathbb{Z}, \mathbb{C})$. We then defined the Fourier coefficients of $[f] \in L_1(\mathbb{T}, \mathbb{C})$ in exactly the same way as for elements of $L_2(\mathbb{T}, \mathbb{C})$, and we have succeeded in showing that

$$(\alpha_n^{[f]})_{n\in\mathbb{Z}} \in c_0(\mathbb{Z}, \mathbb{C}).$$

So far, however, we have not shown that this is an "if and only if" statement, and one reason for this is that it is not. We shall see by Chapter 12 that the map

$$\begin{array}{rccc} \Lambda : & (L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1) & \to & (c_0(\mathbb{Z}, \mathbb{C}), \|\cdot\|_\infty) \\ & [f] & \mapsto & (\alpha_n^{[f]})_{n\in\mathbb{Z}} \end{array}$$

is a continuous, injective linear map, but that it is not surjective. Indeed, linearity is a simple result which is left to the exercises, while continuity of Λ is an easy consequence of the estimate of Section 8.8. In analogy to the situation for $L_2(\mathbb{T}, \mathbb{C})$, it is tempting to ask whether or not the range of Λ is $\ell_1(\mathbb{Z}, \mathbb{C})$. In Chapter 12, we shall discover that this is overly optimistic.

We are left with a number of questions. Let $[f] \in L_1(\mathbb{T}, \mathbb{C})$.

(a) Does the Fourier series $\sum_{n\in\mathbb{Z}} \alpha_n^{[f]} [\xi_n]$ converge, and if so, in which sense? Pointwise (a.e.)? Uniformly? In the L_1-norm?
(b) If the Fourier series *does* converge in some sense, does it converge back to f?
(c) Is $[f]$ completely determined by its Fourier series? That is, if $[f], [g] \in L_1(\mathbb{T}, \mathbb{C})$ and $\alpha_n^{[f]} = \alpha_n^{[g]}$ for all $n \in \mathbb{Z}$, is $[f] = [g]$? (We know that this was true for $[f], [g] \in L_2(\mathbb{T}, \mathbb{C})$.)

These are some of the questions we will consider in the following sections.

Appendix to Section 8.

8.11. So where does the notation $\mathcal{L}_1(\mathbb{T}, \mathbb{C})$ come from, given that we are dealing with 2π-periodic functions on $\mathbb{R}$? The issue lies in the fact that we are *really* interested in studying functions on $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$, but that we have not yet defined what we mean by a *measure* on that set. We are therefore identifying $[-\pi, \pi)$ with $\mathbb{T}$ via the bijective function $\psi(\theta) = e^{i\theta}$. Thus, an alternative approach to this would be to say that a subset $E \subseteq \mathbb{T}$ is **measurable** if and only if $\psi^{-1}(E) \subseteq [-\pi, \pi)$ is Lebesgue measurable. In order to "normalise" the measure of $\mathbb{T}$ (i.e. to make its measure equal to 1), we simply divide Lebesgue measure on $[-\pi, \pi)$ by 2π.

This still doesn't quite explain why we are interested in 2π-periodic functions on $\mathbb{R}$, rather than just functions on $[-\pi, \pi)$, though. Here is the "*kicker*". The unit circle $\mathbb{T} \subseteq \mathbb{C}$ has a very special property, namely, that it is a group. Given $\theta_0 \in \mathbb{T}$, we can "rotate" a function $f : \mathbb{T} \to \mathbb{C}$ in the sense that we set $g(\theta) = f(\theta \cdot \theta_0)$. Observe that rotation along $\mathbb{T}$ corresponds to translation (modulo 2π) of the interval $[-pi, \pi)$. The key is the irritating "modulo 2π" problem. If we don't use modular arithmetic, and if a function g is only defined on $[-\pi, \pi)$, we can not "translate" it, since the new function need no longer have as its domain: $[-\pi, \pi)$. We get around this by extending the domain of g to $\mathbb{R}$ and making g 2π-periodic. Then we may translate g by any real number $\tau_s^\circ(g)(\theta) := g(\theta - s)$, which has the effect that if we set $f(e^{i\theta}) = g(\theta)$, then $g(\theta - s) = f(e^{i\theta} \cdot e^{-is})$. That is, translation of g under addition corresponds to rotation of f under multiplication.

The last thing that we need to know is that such translations of functions will play a crucial role in our study of Fourier series of elements of $L_1(\mathbb{T}, \mathbb{C})$. Aside from being a Banach space, $L_1(\mathbb{T}, \mathbb{C})$ can be made into an algebra under **convolution**. While our analysis will not take us as far as that particular result, we will still need to delve into the theory of convolutions of *continuous* functions with functions in $\mathcal{L}_1(\mathbb{T}, \mathbb{C})$. This will provide us with a way of understanding how and why various series associated to the Fourier series of an element $[f] \in L_1(\mathbb{T}, \mathbb{C})$ converge or diverge. Since convolutions are defined as averages under translation by the group action, and since $\mathbb{T}$ is a group under multiplication and $\mathbb{R}$ is a group under addition, our identification of $(\mathbb{T}, \cdot)$ with $([-\pi, \pi), +)$ (using modular arithmetic) is not an unreasonable way of doing things.

We have been speaking in vague generalities. The next four sections will hopefully add meaning to the above statements.

Exercises for Section 8.

Exercise 8.1.

Let $\mathcal{H}$ be an infinite-dimensional Hilbert space over $\mathbb{K}$, and let $\mathcal{E}$ be an ONB for $\mathcal{H}$.

(a) Prove that $\mathcal{E}$ is linearly independent.
(b) Prove that $\mathcal{E}$ is not a Hamel basis for $\mathcal{H}$.

Exercise 8.2.

Let $1 \le p \le \infty$. Prove that $[f] = [g]$ in $L_p(\mathbb{T}, \mathbb{C})$ if and only if $f = g$ a.e. on $\mathbb{R}$.

Exercise 8.3.

(a) Prove that the map

$$\begin{array}{rccc} \Phi: & L_2([-\pi,\pi],\mathbb{C}) & \to & L_2([-\pi,\pi),\mathbb{C}) \\ & [f] & \mapsto & [f|_{[-\pi,\pi)}] \end{array}$$

is an isomorphism of Hilbert spaces.

(b) Recall from Example[**5.11**] that if we set

$$\xi_n(\theta) = \frac{1}{\sqrt{2\pi}} e^{in\theta}, \ \theta \in [-\pi,\pi], \ n \in \mathbb{Z},$$

then $([\xi_n])_{n\in\mathbb{Z}}$ is an ONB for $L_2([-\pi,\pi],\mathbb{C})$. Let $\psi_n := \xi_n|_{[-\pi,\pi)}$, $n \in \mathbb{Z}$, and prove that $([\psi_n])_{n\in\mathbb{Z}}$ is an ONB for $L_2([-\pi,\pi),\mathbb{C})$.

Exercise 8.4.

Let $1 \le p \le \infty$. Prove that the map

$$\begin{array}{rccc} \Xi_p: & L_p([-\pi,\pi),\mathbb{C}) & \to & L_p(\mathbb{T},\mathbb{C}) \\ & [f] & \mapsto & [\check{f}] \end{array}$$

is a well-defined, isometric isomorphism of Banach spaces.

Exercise 8.5.

Let $1 \le p < \infty$.

(a) Prove that $[\mathrm{SIMP}(\mathbb{T},\mathbb{C})]$ is dense in $(L_p(\mathbb{T},\mathbb{C}), \|\cdot\|_p)$.
(b) Prove that $[\mathrm{STEP}(\mathbb{T},\mathbb{C})]$ is dense in $(L_p(\mathbb{T},\mathbb{C}), \|\cdot\|_p)$.
(c) Prove that $[\mathcal{C}(\mathbb{T},\mathbb{C})]$ is dense in $(L_p(\mathbb{T},\mathbb{C}), \|\cdot\|_p)$.
(d) Prove that $[\mathrm{SIMP}(\mathbb{T},\mathbb{C})]$ is dense in $(L_\infty(\mathbb{T},\mathbb{C}), \|\cdot\|_\infty)$.

Exercise 8.6.

Let $f \in \mathcal{C}(\mathbb{T},\mathbb{C})$. Prove that

$$\|[f]\|_\infty = \|f\|_{\sup} := \sup\{|f(\theta)| : \theta \in [-\pi,\pi)\}.$$

Exercise 8.7. Prove that the map

$$\begin{array}{rccc} \Lambda : & (L_1(\mathbb{T},\mathbb{C}), \|\cdot\|_1) & \to & (c_0(\mathbb{Z},\mathbb{C}), \|\cdot\|_\infty) \\ & [f] & \mapsto & (\alpha_n^{[f]})_{n\in\mathbb{Z}} \end{array}$$

is linear and continuous.

Exercise 8.8.

Let $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$. Prove (or disprove) that the function

$$\begin{array}{rccc} \widehat{f} : & \mathbb{R} & \to & \mathbb{C} \\ & r & \to & \frac{1}{2\pi}\int_{[-\pi,\pi)} f\overline{\xi_r} \end{array}$$

is continuous on $\mathbb{R}$.

9. Convolution

Reporter: "Where do you go from here, Mike?"

"I don't know, man, I mean, I might just fade into Bolivian."

Mike Tyson

9.1. Recall that an **algebra** $\mathcal{B}$ is a vector space over a field $\mathbb{F}$ which also happens to be a ring. A **Banach algebra** $\mathcal{A}$ is a Banach space over $\mathbb{K}$ which is simultaneously an algebra, and for which multiplication is jointly continuous by virtue of its satisfying the inequality

$$\|ab\| \le \|a\| \ \|b\|$$

for all $a, b \in \mathcal{A}$.

For example, $(\mathcal{C}(X, \mathbb{K}), \|\cdot\|_{\sup})$ is a Banach algebra for each locally compact, Hausdorff topological space X, as is $\mathbb{M}_n(\mathbb{K}) \simeq \mathcal{B}(\mathbb{K}^n)$ (for each $n \ge 1$), when equipped with the operator norm.

So far, we have seen that $L_1(\mathbb{T}, \mathbb{C})$ is a Banach space, but we have not investigated any multiplicative structure on it. One can equip $L_1(\mathbb{T}, \mathbb{C})$ with an operation $*$ under which it becomes a Banach algebra. Indeed, given $f, g \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$, we set

$$g \diamond f(\theta) := \frac{1}{2\pi} \int_{[-\pi,\pi)} g(s) f(\theta - s) dm(s),$$

and we refer to this as the **convolution** of g and f. One shows that $[f_1] = [f_2]$ and $[g_1] = [g_2]$ in $L_1(\mathbb{T}, \mathbb{C})$ yields that $g_1 \diamond f_1 = g_2 \diamond f_2$ almost everywhere, allowing one to define

$$[g] * [f] := [g \diamond f]$$

for all $[f], [g] \in L_1(\mathbb{T}, \mathbb{C})$.

It is not entirely clear, *a priori*, that $g \diamond f(\theta) \in \mathbb{C}$ for any $\theta \in \mathbb{R}$, and it is even less clear that $g \diamond f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$. Nevertheless, it is true. The proof of this, however, requires a bit more measure theory than we have developed so far. The key ingredient we are missing is **Fubini's Theorem**, which is stated in the Appendix to this Chapter.

What is easier to prove, however, and what we *shall* prove is that we can turn $\mathcal{L}_1(\mathbb{T}, \mathbb{C})$ (and consequently $L_1(\mathbb{T}, \mathbb{C})$) into a **left module** over $\mathcal{C}(\mathbb{T}, \mathbb{C})$ using convolution. That is, given $g \in \mathcal{C}(\mathbb{T}, \mathbb{C})$ and $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$, we shall set

$$g \diamond f(\theta) := \frac{1}{2\pi} \int_{[-\pi,\pi)} g(s) f(\theta - s) dm(s),$$

and we shall prove that $g \diamond f \in \mathcal{C}(\mathbb{T}, \mathbb{C}) \subseteq \mathcal{L}_1(\mathbb{T}, \mathbb{C})$. Assuming this for the moment, if $f_1 \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$ and $f_1 = f$ a.e. on $\mathbb{R}$, then clearly $g \diamond f(\theta) = g \diamond f_1(\theta)$ for all $\theta \in \mathbb{R}$, whence $g \diamond f = g \diamond f_1$, and this allows us to define

$$g * [f] = [g \diamond f], \quad [f] \in L_1(\mathbb{T}, \mathbb{C}).$$

The advantage to convolving with *continuous* functions only is that we shall be able to reformulate the convolution as a $L_1(\mathbb{T},\mathbb{C})$ *Riemann integral* (in the sense of Chapter 1), and this will allow us gather more information about the continuity properties of this operation, and ultimately about convergence properties of Fourier series.

We invite the reader to consult the Appendix for more information on modules, if necessary.

We begin with a Lemma.

9.2. Lemma. *Let $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ and $s \in \mathbb{R}$ be fixed.*

(a)
$$\int_{[-\pi,\pi)} f = \int_{[-\pi,\pi)} \tau_s^\circ(f),$$
where $\tau_s^\circ(f)(\theta) = f(\theta - s)$.

(b) *If $h(\theta) := f(-\theta)$, $\theta \in \mathbb{R}$, then*
$$\int_{[-\pi,\pi)} h = \int_{[-\pi,\pi)} f.$$

(c) *Define $\varphi_{f,\theta} : \mathbb{R} \to \mathbb{C}$ by $\varphi_{f,\theta}(s) = f(\theta - s)$. Then $\varphi_{f,\theta} \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ and*
$$\nu_1(\varphi_{f,\theta}) = \nu_1(f).$$
That is,
$$\frac{1}{2\pi} \int_{[-\pi,\pi)} |f(\theta - s)| dm(s) = \frac{1}{2\pi} \int_{[-\pi,\pi)} |f(t)| dm(t).$$

Proof. This is an Assignment question.

□

9.3. Definition. *Let $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ and $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$. We define the **convolution of** f **by** g to be the function*
$$\begin{array}{rccc} g \diamond f : & \mathbb{R} & \to & \mathbb{C} \\ & \theta & \mapsto & \frac{1}{2\pi} \int_{[-\pi,\pi)} g(s) f(\theta - s) dm(s). \end{array}$$

The alert reader (hopefully you) will have observed that there is a problem with this definition. For one thing - how do we know that $g \diamond f(\theta)$ exists as a complex number for each $\theta \in \mathbb{R}$? Let us resolve this issue immediately.

Fix $\theta \in \mathbb{R}$. Then
$$\begin{aligned} |g \diamond f(\theta)| &= \frac{1}{2\pi} \left| \int_{[-\pi,\pi)} g(s) f(\theta - s) dm(s) \right| \\ &\le \frac{1}{2\pi} \int_{[-\pi,\pi)} |g(s)| |\varphi_{f,\theta}(s)| dm(s) \\ &\le \|g\|_{\sup} \, \nu_1(\varphi_{f,\theta}) \\ &= \|g\|_{\sup} \, \nu_1(f) < \infty. \end{aligned}$$

Thus $g \diamond f$ is indeed a complex-valued function.

The following extremely useful computation will be used repeated below.

9.4. Lemma. *Let $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$, $g \in \mathcal{L}_\infty(\mathbb{T},\mathbb{C})$. If $\theta \in \mathbb{R}$, then*

$$\int_{[-\pi,\pi)} g(s)f(\theta - s)dm(s) = \int_{[-\pi,\pi)} g(\theta - t)f(t)dm(t).$$

In particular, this holds if $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ and $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$.

Proof. This is an Assignment question.

□

9.5. Remark. In light of Lemma 9.4, for $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ and $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$, we shall define the **convolution of** g **by** f to be

$$f \diamond g(\theta) = \frac{1}{2\pi}\int_{[-\pi,\pi)} g(\theta - t)f(t)dm(t).$$

In so doing, we have guaranteed that $f \diamond g(\theta) = g \diamond f(\theta)$ for all $\theta \in \mathbb{R}$, and so henceforth we shall simply refer to this function as **the convolution of** f **and** g.

9.6. Proposition. *Let $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$ and $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$. Then $g \diamond f \in \mathcal{C}(\mathbb{T},\mathbb{C})$.*

Proof. Note that since g is continuous on $\mathbb{R}$ and 2π-periodic, it is in fact uniformly continuous on $\mathbb{R}$. (See the exercises below.)

Let $\varepsilon > 0$ and choose $\delta > 0$ such that $|x - y| < \delta$ implies that $|g(x) - g(y)| < \varepsilon$.

Let $\theta_0, \theta \in \mathbb{R}$ and suppose that $|\theta - \theta_0| < \delta$. Then, by noting that

$$|g(\theta - s) - g(\theta_0 - s)| < \varepsilon \text{ for all } s \in \mathbb{R},$$

and by applying Lemma 9.4 above, we get:

$$\begin{aligned}
|g \diamond f(\theta) - g \diamond f(\theta_0)| &= \frac{1}{2\pi}\left|\int_{[-\pi,\pi)} g(s)\left(f(\theta - s) - f(\theta_0 - s)\right) dm(s)\right| \\
&= \frac{1}{2\pi}\left|\int_{[-\pi,\pi)} \left(g(\theta - s) - g(\theta_0 - s)\right) f(s)\, dm(s)\right| \\
&\leq \frac{1}{2\pi}\int_{[-\pi,\pi)} |g(\theta - s) - g(\theta_0 - s)||f(s)|\, dm(s) \\
&\leq \frac{1}{2\pi}\int_{[-\pi,\pi)} \varepsilon\, |f(s)|\, dm(s) \\
&= \nu_1(f)\,\varepsilon.
\end{aligned}$$

From this it clearly follows that $g \diamond f$ is (uniformly) continuous.

The fact that $g \diamond f$ is 2π-periodic is clear from the 2π-periodicity of both g and f.

□

9.7. Remark. Suppose that $g \in \mathcal{C}(\mathbb{T}, \mathbb{C})$, $f_1, f_2 \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$ and that $[f_1] = [f_2] \in L_1(\mathbb{T}, \mathbb{C})$; i.e. that $f_1 = f_2$ a.e. on $\mathbb{R}$. Then, since $\varphi_{g,\theta} f_1 = \varphi_{g,\theta} f_2$ a.e. on $\mathbb{R}$ (where $\varphi_{g,\theta}$ is defined as in Lemma 9.2(c) above), we find that for each $\theta \in \mathbb{R}$,

$$\begin{aligned} g \diamond f_1(\theta) &= f_1 \diamond g(\theta) \\ &= \frac{1}{2\pi} \int_{[-\pi,\pi)} g(\theta - t) f_1(t) dm(t) \\ &= \frac{1}{2\pi} \int_{[-\pi,\pi)} g(\theta - t) f_2(t) dm(t) \\ &= f_2 \diamond g(\theta) \\ &= g \diamond f_2(\theta). \end{aligned}$$

This allows us to extend our notion of convolution as follows:

9.8. Definition. *Given $g \in \mathcal{C}(\mathbb{T})$ and $[f] \in L_1(\mathbb{T})$, we define the **convolution** of g and $[f]$ to be*

$$g * [f] := [g \diamond f],$$

where $g \diamond f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$ is the convolution of Definition 9.3.

*We also define the **convolution operator** with **kernel** g to be the map:*

$$\begin{array}{rccc} C_g : & L_1(\mathbb{T}, \mathbb{C}) & \to & L_1(\mathbb{T}, \mathbb{C}) \\ & [f] & \mapsto & g * [f]. \end{array}$$

Observe that if $[f_1]$ and $[f_2] \in L_1(\mathbb{T}, \mathbb{C})$, and if $\kappa \in \mathbb{C}$, then

$$\begin{aligned} C_g(\kappa [f_1] + [f_2]) &= g * [\kappa f_1 + f_2] \\ &= \frac{1}{2\pi} \int_{[-\pi,\pi)} g(s)(\kappa f_1(\theta - s) + f_2(\theta - s)) dm(s) \\ &= \kappa \frac{1}{2\pi} \int_{[-\pi,\pi)} g(s) f_1(\theta - s) dm(s) + \frac{1}{2\pi} \int_{[-\pi,\pi)} g(s) f_2(\theta - s) dm(s) \\ &= \kappa g * [f_1] + g * [f_2] \\ &= \kappa C_g([f_1]) + C_g([f_2]), \end{aligned}$$

proving that C_g is a linear map on $L_1(\mathbb{T}, \mathbb{C})$. Given that C_g *is* a linear map and $(L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1)$ is a Banach space, it is of interest to determine whether or not C_g is bounded, and if so, then what is its norm? As we shall see - the answer to this question will be intimately related to the question of convergence of Fourier series of elements of $L_1(\mathbb{T}, \mathbb{C})$.

A direct computation of the norm of C_g as defined above is rather difficult, using the tools currently at our disposition. Our strategy, therefore, is to reformulate C_g as a vector-valued Riemann integral on $L_1(\mathbb{T}, \mathbb{C})$, as developed in Chapter 1. In fact, we shall be able to extend this notion of convolution beyond the Banach space $L_1(\mathbb{T}, \mathbb{C})$. For this, we shall need the concept of a *homogeneous Banach space*, which we now define.

9.9. Homogeneous Banach spaces. Let $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$, and let $s \in \mathbb{R}$. Consider the function

$$\begin{array}{rccc} \tau_s^\circ(f): & \mathbb{R} & \to & \mathbb{C} \\ & \theta & \mapsto & f(\theta - s). \end{array}$$

One should think of τ_s° as *translating* f *by* s. The superscript $\circ$ above the τ_s is to indicate that we are acting on *functions*. When acting on elements of $L_1(\mathbb{T}, \mathbb{C})$, we shall drop this superscript.

As we shall see in the Assignments, the fact that $\mathfrak{M}(\mathbb{R})$ is invariant under translation, that Lebesgue measure is translation-invariant, and that the set of 2π-periodic functions is again invariant under translation implies that

$$\tau_s^\circ(f) \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$$

as well. Furthermore, if $[f] = [g] \in L_1(\mathbb{T}, \mathbb{C})$, then $[\tau_s^\circ(f)] = [\tau_s^\circ(g)]$, as is easily verified. Thus we define the operation of **translation by** s on $L_1(\mathbb{T}, \mathbb{C})$ via

$$\tau_s([f]) := [\tau_s^\circ(f)].$$

9.10. Definition. *A* ***homogeneous Banach space over*** $\mathbb{T}$ *is a linear manifold* $\mathfrak{B}$ *in* $L_1(\mathbb{T})$ *equipped with a norm* $\|\cdot\|_{\mathfrak{B}}$ *with respect to which* $(\mathfrak{B}, \|\cdot\|_{\mathfrak{B}})$ *is a Banach space, and satisfying*

(a) $\|[f]\|_1 \le \|[f]\|_{\mathfrak{B}}$ *for all* $[f] \in \mathfrak{B}$.
(b) $[\mathrm{Trig}(\mathbb{T})] \subseteq \mathfrak{B}$;
(c) $\mathfrak{B}$ *is invariant under translation; that is, for all* $[f] \in \mathfrak{B}$ *and* $s \in \mathbb{R}$,

$$\tau_s[f] = [\tau_s^\circ(f)] \in \mathfrak{B};$$

(d) *for all* $[f] \in \mathfrak{B}$ *and* $s \in \mathbb{R}$, $\|\tau_s[f]\|_{\mathfrak{B}} = \|[f]\|_{\mathfrak{B}}$ *; and*
(e) *for each* $[f] \in \mathfrak{B}$*, the map*

$$\begin{array}{rccc} \Psi_{[f]}: & \mathbb{R} & \to & \mathfrak{B} \\ & s & \mapsto & \tau_s([f]) \end{array}$$

is continuous.

The idea that a linear manifold $\mathfrak{M}$ of a Banach space $\mathfrak{X}$ might not be closed in the ambient norm, but that $(\mathfrak{M}, \|\cdot\|_{\mathfrak{M}})$ might be complete in its own norm and hence a Banach space might seem a bit strange at first. Come to think of it, however, we have already seen this multiple times! Note that each of the spaces $L_p(\mathbb{T}, \mathbb{C})$ is dense in $L_1(\mathbb{T}, \mathbb{C})$, $1 \le p < \infty$, and that each is complete using the corresponding $\|\cdot\|_p$ norm.

9.11. Example. Recall that $[\mathcal{C}(\mathbb{T}, \mathbb{C})] \subseteq L_\infty(\mathbb{T}, \mathbb{C})$ is a subset of $L_1(\mathbb{T}, \mathbb{C})$ and that it is clearly a linear manifold. Furthermore, for $f \in \mathcal{C}(\mathbb{T}, \mathbb{C})$, it follows from Example 6.19 (b) (see also Exercise 6.11) that

$$\|[f]\|_\infty = \|f\|_{\sup} := \sup\{|f(\theta)| : \theta \in [-\pi, \pi)\},$$

and that $([\mathcal{C}(\mathbb{T}, \mathbb{C})], \|\cdot\|_\infty)$ is a Banach space. We claim that it is in fact a *homogeneous* Banach space over $\mathbb{T}$.

(a) Let $f \in \mathcal{C}(\mathbb{T},\mathbb{C})$. Then

$$\|[f]\|_1 = \frac{1}{2\pi}\int_{[-\pi,\pi)} |f| \le \frac{1}{2\pi}\int_{[-\pi,\pi)} \|f\|_{\sup} = \|f\|_{\sup} = \|[f]\|_\infty.$$

(b) Since, for each $n \in \mathbb{Z}$, $\xi_n \in \mathcal{C}(\mathbb{T},\mathbb{C})$, and since the latter is a linear manifold, we have that $[\mathrm{Trig}(\mathbb{T},\mathbb{C})] \subseteq [\mathcal{C}(\mathbb{T},\mathbb{C})]$.

(c) If $f \in \mathcal{C}(\mathbb{T},\mathbb{C})$ and $s \in \mathbb{R}$, then clearly $\tau_s^\circ(f) \in \mathcal{C}(\mathbb{T},\mathbb{C})$, so that $\tau_s[f] \in [\mathcal{C}(\mathbb{T},\mathbb{C})]$. It follows that $[\mathcal{C}(\mathbb{T},\mathbb{C})]$ is translation-invariant.

(d) Given $f \in \mathcal{C}(\mathbb{T},\mathbb{C})$ and $s \in \mathbb{R}$,

$$\begin{aligned}\|\tau_s[f]\|_\infty &= \|[\tau_s^\circ(f)]\|_\infty \\ &= \|\tau_s^\circ(f)\|_{\sup} \\ &= \sup\{|f(\theta - s)| : \theta \in \mathbb{R}\} \\ &= \sup\{|f(\theta)| : \theta \in \mathbb{R}\} \\ &= \|f\|_{\sup} \\ &= \|[f]\|_\infty.\end{aligned}$$

(e) Let $[f] \in [\mathcal{C}(\mathbb{T},\mathbb{C})]$, and assume without loss of generality that $f \in \mathcal{C}(\mathbb{T},\mathbb{C})$. Then $f : \mathbb{R} \to \mathbb{C}$ is continuous and 2π-periodic. From this we easily conclude that f is uniformly continuous on $\mathbb{R}$. Let $\varepsilon > 0$, and choose $\delta > 0$ such that $x, y \in \mathbb{R}$ and $|x - y| < \delta$ implies that $|f(x) - f(y)| < \frac{\varepsilon}{2}$.

Let $s_0 \in \mathbb{R}$ be fixed. If $|s - s_0| < \delta$, then

$$\begin{aligned}\|\Psi_{[f]}(s) - \Psi_{[f]}(s_0)\|_\infty &= \|\tau_s[f] - \tau_{s_0}[f]\|_\infty \\ &= \|\tau_s^\circ(f) - \tau_{s_0}^\circ(f)\|_{\sup} \\ &= \sup_{\theta\in[-\pi,\pi)} |f(\theta - s) - f(\theta - s_0)| \\ &\le \frac{\varepsilon}{2} \\ &< \varepsilon,\end{aligned}$$

since $|(\theta - s) - (\theta - s_0)| = |s - s_0| < \delta$ for all $\theta \in [-\pi,\pi)$.

Thus $\Psi_{[f]}$ is continuous at s_0. But $s_0 \in \mathbb{R}$ was arbitrarily chosen, so $\Psi_{[f]}$ is continuous on $\mathbb{R}$. Since $[f] \in [\mathcal{C}(\mathbb{T},\mathbb{C})]$ was also arbitrary, this completes the proof of the fact that $([\mathcal{C}(\mathbb{T},\mathbb{C}), \|\cdot\|_\infty)$ is a homogeneous Banach space over $\mathbb{T}$.

9.12. Example. Let $1 \le p < \infty$. We claim that $(L_p(\mathbb{T},\mathbb{C}), \|\cdot\|_p)$ is a homogeneous Banach space over $\mathbb{T}$.

(a) Let $f \in \mathcal{L}_p(\mathbb{T},\mathbb{C})$ and let q denote the Lebesgue conjugate of p; i.e. $\frac{1}{p} + \frac{1}{q} = 1$. Recall from Proposition 4.10 that there exists a measurable function $u : \mathbb{R} \to \mathbb{T}$ such that $f = u \cdot |f|$. Note that $u \in \mathcal{L}_q(\mathbb{T},\mathbb{C})$; indeed, the fact that f is 2π-periodic implies that u is, and u has already been seen to be measurable. Moreover,

$$\|[u]\|_q = \left(\frac{1}{2\pi}\int_{[-\pi,\pi)} |u|^q\right)^{1/q} = \left(\frac{1}{2\pi}\int_{[-\pi,\pi)} 1\right)^{1/q} = 1.$$

Of course, $\|[\overline{u}]\|_q = \|[u]\|_q = 1$ as well. By Hölder's Inequality,

$$\|[f]\|_1 = \frac{1}{2\pi}\int_{[-\pi,\pi)} |f\cdot\overline{u}| \le \|[f]\|_p \ \|[\overline{u}]\|_q \le \|[f]\|_p.$$

(b) Observe that $[\mathrm{Trig}(\mathbb{T},\mathbb{C})] \subseteq [\mathcal{C}(\mathbb{T},\mathbb{C})] \subseteq L_p(\mathbb{T},\mathbb{C}) \subseteq L_1(\mathbb{T},\mathbb{C})$.

(c) and (d) Let $[f] \in L_p(\mathbb{T},\mathbb{C})$ and $s \in \mathbb{R}$. As noted in Section 9.9, $\tau_s[f] \in L_1(\mathbb{T},\mathbb{C})$, and in particular $\tau_s^\circ(f)$ is measurable.

Moreover, using Lemma 9.2:

$$\begin{aligned}\|\tau_s[f]\|_p &= \left(\frac{1}{2\pi}\int_{[-\pi,\pi)} |f(\theta - s)|^p dm(s)\right)^{1/p}\\ &= \left(\frac{1}{2\pi}\int_{[-\pi,\pi)} |f(\theta)|^p dm(\theta)\right)^{1/p}\\ &= \|[f]\|_p < \infty.\end{aligned}$$

In particular, $\tau_s[f] \in L_p(\mathbb{T},\mathbb{C})$.

(e) Let $[f] \in L_p(\mathbb{T},\mathbb{C})$ and $s_0 \in \mathbb{R}$. Fix $\varepsilon > 0$. Since $[\mathcal{C}(\mathbb{T},\mathbb{C})]$ is dense in $(L_p(\mathbb{T},\mathbb{C}), \|\cdot\|_p)$, we can find $h \in \mathcal{C}(\mathbb{T},\mathbb{C})$ such that

$$\|[f] - [h]\|_p < \frac{\varepsilon}{3}.$$

By Example 9.11, the map $s \mapsto \tau_s[h]$ is continuous with respect to the $\|\cdot\|_\infty$-norm on $[\mathcal{C}(\mathbb{T},\mathbb{C})]$. Thus we may choose $\delta > 0$ such that $|s - s_0| < \delta$ implies that

$$\|\tau_s[h] - \tau_{s_0}[h]\|_\infty < \frac{\varepsilon}{3}.$$

By the triangle inequality,

$$\|\tau_s[f] - \tau_{s_0}[f]\|_p \le \|\tau_s[f] - \tau_s[h]\|_p + \|\tau_s[h] - \tau_{s_0}[h]\|_p + \|\tau_{s_0}[h] - \tau_{s_0}[f]\|_p.$$

Let us estimate each of the terms on the right-hand side of this inequality.

Now (since translation is isometric on $L_p(\mathbb{T},\mathbb{C})$ from above)

$$\|\tau_s[f] - \tau_s[h]\|_p = \|\tau_s[f-h]\|_p = \|[f-h]\|_p = \|[f] - [h]\|_p < \frac{\varepsilon}{3},$$

and similarly,

$$\|\tau_{s_0}[f] - \tau_{s_0}[h]\|_p < \frac{\varepsilon}{3}.$$

Moreover,

$$\begin{aligned}\|\tau_s[h] - \tau_{s_0}[h]\|_p &= \left(\frac{1}{2\pi}\int_{[-\pi,\pi)} |\tau_s^\circ(h) - \tau_{s_0}^\circ(h)|^p\right)^{1/p}\\ &\le \left(\frac{1}{2\pi}\int_{[-\pi,\pi)} \|\tau_s[h] - \tau_{s_0}[h]\|_\infty^p\right)^{1/p}\\ &\le \left(\frac{1}{2\pi}\int_{[-\pi,\pi)} (\varepsilon/3)^p\right)^{1/p}\\ &= \frac{\varepsilon}{3}.\end{aligned}$$

Substituting these estimates into the inequality above shows that for $|s - s_0| < \delta$,

$$\|\tau_s[f] - \tau_{s_0}[f]\|_p < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

and thus translation is continuous on $(L_p(\mathbb{T},\mathbb{C}), \|\cdot\|_p)$.

Hence $(L_p(\mathbb{T},\mathbb{C}), \|\cdot\|_p)$ is a homogeneous Banach space over $\mathbb{T}$ when $1 \le p < \infty$.

9.13. Example. When $p = \infty$, the situation is rather different. We shall now prove that $(L_\infty(\mathbb{T},\mathbb{C}), \|\cdot\|_\infty)$ is *not* a homogeneous Banach space over $\mathbb{T}$.

We leave it to the exercises for the reader to show that

(a) $\|[f]\|_1 \le \|[f]\|_\infty$ for all $[f] \in L_\infty(\mathbb{T},\mathbb{C})$.
(b) $[\mathrm{Trig}(\mathbb{T},\mathbb{C})] \subseteq L_\infty(\mathbb{T},\mathbb{C})$;
(c) for all $[f] \in L_\infty(\mathbb{T},\mathbb{C})$ and $s \in \mathbb{R}$,

$$\tau_s[f] \in L_\infty(\mathbb{T},\mathbb{C}); \text{ and}$$

(d) for all $[f] \in L_\infty(\mathbb{T},\mathbb{C})$ and $s \in \mathbb{R}$,

$$\|\tau_s[f]\|_\infty = \|[f]\|_\infty.$$

The failure of $L_\infty(\mathbb{T},\mathbb{C})$ to be a homogeneous Banach space over $\mathbb{T}$ comes down to the fact that translation is not continuous in $L_\infty(\mathbb{T},\mathbb{C})$.

To see this, consider the function $f_0 := \chi_{[0,\pi)} \in \mathcal{L}_\infty([-\pi,\pi))$, and let $f = \check{f}_0 \in \mathcal{L}_\infty(\mathbb{T},\mathbb{C})$ be its 2π-periodic extension, as defined in Section 8.4.

If $-\pi < s < 0$, then $\tau_s^\circ(f)(\theta) - \tau_0^\circ(f)(\theta) = 1 - 0 = 1$ for all $\theta \in (s, 0)$, and thus

$$\|\tau_s[f] - \tau_0[f]\|_\infty \ge 1.$$

(It is in fact equal to 1.) In particular,

$$\lim_{s\to 0} \|\tau_s[f] - \tau_0[f]\|_\infty = 1 \ne 0 = \|\tau_0[f] - \tau_0[f]\|_\infty,$$

and so the map $s \mapsto \tau_s[f]$ is not continuous at 0.

9.14. Let $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$ and $[f] \in L_1(\mathbb{T},\mathbb{C})$. We have defined (see Definition 9.8) the convolution of g and $[f]$ to be $g * [f] := [g \diamond f]$, where

$$g \diamond f(\theta) = \frac{1}{2\pi} \int_{[-\pi,\pi)} g(s) f(\theta - s)\ dm(s).$$

That is, we define $g * [f]$ by first defining $g \diamond f$ pointwise, using Lebesgue integration.

Now consider that Example 9.12 shows that $(L_1(\mathbb{T},\mathbb{C}), \|\cdot\|_1)$ is a homogeneous Banach space over $\mathbb{T}$. As such, the function

$$\begin{array}{rccc} \beta: & \mathbb{R} & \to & L_1(\mathbb{T},\mathbb{C}) \\ & s & \mapsto & g(s)\tau_s[f] \end{array}$$

is continuous. By Theorem 1.14,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \beta(s)\ ds = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)\tau_s[f]\ ds$$

exists in $L_1(\mathbb{T},\mathbb{C})$, and is obtained as an $\|\cdot\|_1$-limit of Riemann sums $(\beta, P_N, P_N^*) \in L_1(\mathbb{T},\mathbb{C})$ using partitions P_N of $[-2\pi, 2\pi]$ with corresponding choices P_N^* of test values for P_N.

If we fix $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$, then we obtain a map

$$\begin{array}{rccc} \Gamma_g: & L_1(\mathbb{T},\mathbb{C}) & \to & L_1(\mathbb{T},\mathbb{C}) \\ & [f] & \mapsto & \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)\tau_s[f]\ ds. \end{array}$$

We leave it to the reader to verify that Γ_g is linear.

The reader may have noticed that there is a striking resemblance between the operators C_g and Γ_g. After all, $\tau_s[f] = [\tau_s^\circ(f)]$, where $\tau_s^\circ(f)(\theta) = f(\theta - s)$, $\theta \in \mathbb{R}$.

Our next goal is to show that in fact, $\Gamma_g = C_g$, so that for each $[f] \in L_1(\mathbb{T},\mathbb{C})$,

$$C_g[f] = g * [f] = [g \diamond f] = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)\tau_s[f]\ ds = \Gamma_g[f].$$

This is not an obvious nor a trivial result: the two constructions are entirely different. Quite frankly, some of us – and you know who you are – don't deserve this.

9.15. In Example 9.11, we showed that $([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty)$ is a homogeneous Banach space over $\mathbb{T}$. It follows that given $f \in \mathcal{C}(\mathbb{T},\mathbb{C})$, the map $s \mapsto \tau_s[f]$ is continuous, or equivalently, the map $s \mapsto \tau_s^\circ(f)$ is continuous from $(\mathbb{R}, |\cdot|)$ to $(\mathcal{C}(\mathbb{T},\mathbb{C}), \|\cdot\|_{\sup})$.

Before presenting our general result, we shall need a crucial Lemma.

9.16. Lemma. *Let $f, g \in (\mathcal{C}(\mathbb{T}), \|\cdot\|_{\sup})$. Let*

$$\Gamma_g^\circ(f) := \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)\tau_s^\circ(f)\ ds,$$

taken as a Banach space Riemann integral in $(\mathcal{C}(\mathbb{T}), \|\cdot\|_{\sup})$ in the sense of Chapter 1. Then

$$\Gamma_g^\circ(f)(\theta) = g \diamond f(\theta) = \frac{1}{2\pi}\int_{[-\pi,\pi)} g(s) f(\theta - s)\ dm(s) \quad \text{for all } \theta \in \mathbb{R}.$$

Proof. Perhaps the most difficult part of this proof is to first ensure that we understand the difference between $\Gamma_g^\circ(f)$ and $g \diamond f$. On the one hand, $(\mathcal{C}(\mathbb{T}), \|\cdot\|_{\sup})$ is a Banach space and the map $\beta : \mathbb{R} \to \mathcal{C}(\mathbb{T},\mathbb{C})$ defined by $\beta(s) := g(s)\tau_s^\circ(f) \in \mathcal{C}(\mathbb{T},\mathbb{C})$ is continuous. By Theorem 1.14, $\Gamma_g^\circ(f) \in \mathcal{C}(\mathbb{T},\mathbb{C})$ exists as a $\|\cdot\|_{\sup}$-limit of Riemann sums $S(\beta, P_N, P_N^*) \in (\mathcal{C}(\mathbb{T},\mathbb{C}), \|\cdot\|_{\sup})$. In fact, as we saw there, we may suppose without loss of generality that for each $N \geq 1$, $P_N \in \mathcal{P}([-\pi,\pi])$ is a regular partition of $[-\pi,\pi]$ into 2^N subintervals of equal length $\frac{2\pi}{2^N}$, and $P_N^* = P_N \setminus \{-\pi\}$, so that P_N^* is a set of test values for P_N.

Meanwhile, $g \diamond f$ is the convolution of g and f defined pointwise through Lebesgue integration, as in Definition 9.3.

Let us temporarily fix $\theta_0 \in \mathbb{R}$ and define a function $\gamma (= \gamma_{\theta_0}) : \mathbb{R} \to \mathbb{K}$ via:

$$\gamma(s) = g(s) f(\theta_0 - s), \quad s \in \mathbb{R}.$$

In the present case where g and f are both continuous, the map γ is easily seen to also be continuous, and thus both bounded and Riemann integrable on $[-\pi,\pi)$. By Theorem 5.24 therefore,

$$\begin{aligned} g \diamond f(\theta_0) &:= \frac{1}{2\pi} \int_{[-\pi,\pi)} g(s) f(\theta_0 - s)\ dm(s) \\ &= \frac{1}{2\pi} \int_{[-\pi,\pi)} \gamma(s)\ dm(s) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \gamma(s)\ ds \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(s) f(\theta_0 - s)\ ds, \end{aligned}$$

where the last two integrals are Riemann integrals.

Since $(\mathbb{C}, |\cdot|)$ is a Banach space, the argument of Theorem 1.14 also shows that with P_N and P_N^* defined as above (alternatively, from first-year Calculus),

$$\begin{aligned} g \diamond f(\theta_0) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \gamma(s)\ ds \\ &= \lim_{N\to\infty} S(\gamma, P_N, P_N^*). \end{aligned}$$

Finally,

$$\begin{aligned} \|\Gamma_g^\circ(f) - S(\beta, P_N, P_N^*)\|_{\sup} &\geq |\Gamma_g^\circ(f)(\theta_0) - S(\beta, P_N, P_N^*)(\theta_0)| \\ &= |\Gamma_g^\circ(f)(\theta_0) - \sum_{n=1}^{2^N} (\beta(p_n))(\theta_0)(p_n - p_{n-1})| \\ &= |\Gamma_g^\circ(f)(\theta_0) - \sum_{n=1}^{2^N} (g(p_n)) f(\theta_0 - p_n)(p_n - p_{n-1})| \end{aligned}$$

$$= |\Gamma_g^\circ(f)(\theta_0) - \sum_{n=1}^{2^N} \gamma(p_n)(p_n - p_{n-1})|$$
$$= |\Gamma_g^\circ(f)(\theta_0) - S(\gamma, P_N, P_N^*)|.$$

But $\lim_{N\to\infty} \|\Gamma_g^\circ(f) - S(\beta, P_N, P_N^*)\|_{\sup} = 0$, so

$$0 = \lim_{N\to\infty} |\Gamma_g^\circ(f)(\theta_0) - S(\gamma, P_N, P_N^*)|.$$

That is,

$$\Gamma_g^\circ(f)(\theta_0) = \lim_{N\to\infty} S(\gamma, P_N, P_N^*) = g \diamond f(\theta_0).$$

Since $\theta_0 \in \mathbb{R}$ was arbitrary, $\Gamma_g^\circ(f) = g \diamond f$.

□

9.17. Theorem. *Let $g \in \mathcal{C}(\mathbb{T}, \mathbb{C})$ and $[f] \in L_1(\mathbb{T}, \mathbb{C})$. Let*

$$\Gamma_g[f] := \frac{1}{2\pi} \int_{-\pi}^{\pi} g(s)\tau_s[f] \; ds,$$

where the integral is a Banach space Riemann integral in $(L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1)$ in the sense of Chapter 1. Then

$$\Gamma_g[f] = g * [f] = [g \diamond f] = C_g[f].$$

Proof. Since $[\mathcal{C}(\mathbb{T}, \mathbb{C})]$ is $\|\cdot\|_1$-dense in $L_1(\mathbb{T}, \mathbb{C})$, we can find a sequence $(f_m)_{m=1}^\infty$ in $\mathcal{C}(\mathbb{T}, \mathbb{C})$ such that

$$\lim_{n\to\infty} \|[f_m] - [f]\|_1 = 0.$$

For each $m \geq 1$, define

$$\Gamma_g[f_m] := \frac{1}{2\pi} \int_{-\pi}^{\pi} g(s)\tau_s[f_m] \; ds$$

as a Riemann integral in $(L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1)$ in the sense of Chapter 1.

Because $f_m \in \mathcal{C}(\mathbb{T}, \mathbb{C})$ for each $m \geq 1$, we have that the map $s \mapsto g(s)\tau_s^\circ(f)$ is continuous, and thus the integral

$$\Gamma_g^\circ(f_m) := \frac{1}{2\pi} \int_{-\pi}^{\pi} g(s)\tau_s^\circ(f_m) \; ds,$$

converges in $(\mathcal{C}(\mathbb{T}, \mathbb{C}), \|\cdot\|_{\sup})$ by Lemma 9.16. Equivalently, the Banach space Riemann integral

$$[\Gamma_g^\circ(f_m)] := \frac{1}{2\pi} \int_{-\pi}^{\pi} g(s)\tau_s[f_m] \; ds$$

converges in $([\mathcal{C}(\mathbb{T}, \mathbb{C})], \|\cdot\|_\infty)$. But $\|[h]\|_1 \leq \|[h]\|_\infty$ for all $h \in \mathcal{C}(\mathbb{T}, \mathbb{C})$, and thus

$$[\Gamma_g^\circ(f_m)] = \Gamma_g[f_m],$$

as the integral also converges (to the same element) in $(L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1)$. (More details may be found in the Appendix – Remark 9.39.)

Step One.

First we shall show that $\Gamma_g[f_m] = g * [f_m]$ for all $m \geq 1$.

By Lemma 9.16, $\Gamma_g^\circ(f_m) = g \diamond f_m$ for all $m \geq 1$. Thus

$$\Gamma_g[f_m] = [\Gamma_g^\circ(f_m)] = [g \diamond f_m] = g * [f_m], \quad m \geq 1.$$

Step Two.

Next we show that $g * [f] = \lim_{m\to\infty} g * [f_m]$ in $(L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1)$.

Now, for all $m \geq 1$ and all $\theta \in \mathbb{R}$,

$$|g \diamond (f - f_m)(\theta)| \leq \nu_\infty(g)\ \nu_1(f - f_m) = \|g\|_{\sup}\ \|[f] - [f_m]\|_1,$$

as we saw in the paragraph following Definition 9.3. Thus for all $m \geq 1$,

$$\begin{aligned}
\|g * [f] - g * [f_m]\|_1 &= \|g * [f - f_m]\|_1 \\
&= \frac{1}{2\pi} \int_{[-\pi,\pi)} |g \diamond (f - f_m)(\theta)| dm(\theta) \\
&\leq \frac{1}{2\pi} \int_{[-\pi,\pi)} \|g\|_{\sup}\ \|[f] - [f_m]\|_1 dm(\theta) \\
&= \|g\|_{\sup}\ \|[f] - [f_m]\|_1,
\end{aligned}$$

from which it easily follows that

$$g * [f] = \lim_{m\to\infty} g * [f_m] \quad \text{in } (L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1).$$

Step Three.

We now show that $\Gamma_g[f] = \lim_{m\to\infty} \Gamma_g[f_m]$ in $(L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1)$.
Indeed,

$$\begin{aligned}
\|\Gamma_g[f] - \Gamma_g[f_m]\|_1 &= \left\| \frac{1}{2\pi} \int_{-\pi}^{\pi} g(s) \tau_s([f - f_m])\ ds \right\|_1 \\
&\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(s)|\ \|\tau_s([f - f_m])\|_1\ ds \\
&\leq \|g\|_{\sup} \frac{1}{2\pi} \int_{-\pi}^{\pi} \|[f] - [f_m]\|_1 ds \\
&= \|g\|_{\sup}\ \|[f] - [f_m]\|_1.
\end{aligned}$$

As before, it easily follows that

$$\Gamma_g[f] = \lim_{m\to\infty} \Gamma_g[f_m] \quad \text{in } (L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1).$$

Step Four.

Finally (!) we see that in $(L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1)$ we have that

$$\Gamma_g[f] = \lim_{m\to\infty} \Gamma_g[f_m] = \lim_{m\to\infty} g * [f_m] = g * [f] = C_g[f].$$

□

Viewing Γ_g as a map from $L_1(\mathbb{T},\mathbb{C})$ into itself, we have shown that $\Gamma_g = C_g$; in other words, the two notions of "convolution" agree. But in fact, when $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$, we may define

$$\begin{array}{rccc} \Gamma_g^{\mathfrak{B}}: & \mathfrak{B} & \to & \mathfrak{B} \\ & [f] & \mapsto & \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)\tau_s[f]\, ds \end{array}$$

as a map on *any* homogeneous Banach space $\mathfrak{B}$ over $\mathbb{T}$. Also, $C_g[f] = [g \diamond f] \in \mathcal{B}$, Let us now show that it always agrees with convolution.

9.18. Theorem. *Let $(\mathfrak{B}, \|\cdot\|_{\mathfrak{B}})$ be a homogeneous Banach space over $\mathbb{T}$, $[f] \in \mathfrak{B}$, and let $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$. Then*

$$\Gamma_g^{\mathfrak{B}}[f] \coloneqq \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)\tau_s[f]ds$$

converges in $\mathfrak{B}$.

(a) *Furthermore*

$$\Gamma_g^{\mathfrak{B}}[f] = g * [f] = [g \diamond f] = C_g[f],$$

where – as always – for all $\theta \in \mathbb{R}$,

$$g \diamond f(\theta) = \frac{1}{2\pi}\int_{[-\pi,\pi)} g(s) f(\theta - s)\, dm(s).$$

(b) *Moreover,*

$$\|g * [f]\|_{\mathfrak{B}} \le \nu_1(g)\, \|[f]\|_{\mathfrak{B}}.$$

Proof.

(a) Since $(\mathfrak{B}, \|\cdot\|_{\mathfrak{B}})$ is a Banach space, and since – for fixed $[f] \in \mathfrak{B}$ – the function $\beta : \mathbb{R} \to \mathfrak{B}$ defined by $\beta(s) \coloneqq g(s)\tau_s[f]$ is continuous, we see from Theorem 1.14 that

$$\Gamma_g^{\mathfrak{B}}[f] \coloneqq \frac{1}{2\pi}\int_{-\pi}^{\pi} \beta(s)\, ds = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)\tau_s[f]\, ds$$

exists in $\mathfrak{B}$ (i.e. the Riemann sums corresponding to the integral converge in $\mathfrak{B}$).

As always, using the arguments of Theorem 1.14, we see that if we set P_N to be a regular partition of $[-\pi,\pi]$ into 2^N subintervals of equal length, and if $P_N^* = P_N \setminus \{\pi\}$ is a corresponding set of test values for P_N, then

$$\lim_{N\to\infty} \|\Gamma_g^{\mathfrak{B}}[f] - S(\beta, P_N, P_N^*)\|_{\mathfrak{B}} = 0.$$

But $\|[h]\|_1 \le \|[h]\|_{\mathfrak{B}}$ for all $[h] \in \mathfrak{B}$, as $\mathfrak{B}$ is a homogeneous Banach space over $\mathbb{T}$. Thus

$$\lim_{N\to\infty} \|\Gamma_g^{\mathfrak{B}}[f] - S(\beta, P_N, P_N^*)\|_1 = 0.$$

In other words,

$$\Gamma_g^{\mathfrak{B}}[f] = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(s)\tau_s[f]ds \quad \text{in } (L_1(\mathbb{T},\mathbb{C}), \|\cdot\|_1).$$

Phrased another way, $\Gamma_g^{\mathfrak{B}}[f] = \Gamma_g^{L_1(\mathbb{T},\mathbb{C})}[f]$. By Theorem 9.17,

$$\Gamma_g^{\mathfrak{B}}[f] = \Gamma_g^{L_1(\mathbb{T},\mathbb{C})}[f] = g * [f] = C_g[f].$$

(b) Recall that – in any homogeneous Banach space over $\mathbb{T}$ – we defined the continuous map $\Phi_{[f]} : \mathbb{R} \to \mathfrak{B}$ via $\Psi_{[f]}(s) = \tau_s[f]$, and we have that $\|\Psi_{[f]}(s)\|_{\mathfrak{B}} = \|[f]\|_{\mathfrak{B}}$ for all $s \in \mathbb{R}$.

Next, consider that

$$\begin{aligned}\|g * [f]\|_{\mathfrak{B}} &= \frac{1}{2\pi}\left\|\int_{-\pi}^{\pi} g(s)\tau_s[f]\,ds\right\|_{\mathfrak{B}}\\ &= \frac{1}{2\pi}\left\|\int_{-\pi}^{\pi} g(s)\Psi_{[f]}(s)\,ds\right\|_{\mathfrak{B}}\\ &\le \frac{1}{2\pi}\int_{-\pi}^{\pi} |g(s)|\;\|\Psi_{[f]}(s)\|_{\mathfrak{B}}\,ds\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |g(s)|\,\|[f]\|_{\mathfrak{B}}\,ds\\ &= \|[f]\|_{\mathfrak{B}}\,\nu_1(g).\end{aligned}$$

□

9.19. Remark. The conclusion of Theorem 9.18 (a) is actually somewhat stronger than it might at first appear. We showed in Proposition 9.6 that if $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$ and $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$, then $g \diamond f \in \mathcal{C}(\mathbb{T},\mathbb{C})$, and thus $g * [f] := [g \diamond f] \in [\mathcal{C}(\mathbb{T},\mathbb{C})]$. So why should it lie in $\mathfrak{B}$? After all, there is no reason why $\mathfrak{B}$ should contain all continuous functions (despite the fact that it contains all trigonometric polynomials). What we have just shown is that $g * [f] \in \mathfrak{B}$ even if $\mathfrak{B}$ doesn't contain $[\mathcal{C}(\mathbb{T},\mathbb{C})]$. In other words, convolution (at least by a continuous function) places us in the smaller space $\mathfrak{B}$ (when we start in the smaller space).

It can in fact be shown that if $[g] \in L_1(\mathbb{T},\mathbb{C})$, $[f] \in \mathfrak{B} \subseteq L_1(\mathbb{T},\mathbb{C})$, and if $[g] * [f]$ is defined as in Section 9.1, then $[g] * [f] \in \mathfrak{B}$, but that is beyond the scope of the course.

9.20. Theorem. *Let $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$, and let*

$$C_g : ([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty) \to ([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty)$$

be the convolution operator corresponding to g, so that $C_g[h] = g * [h]$. Then $\|C_g\| = \nu_1(g) = \|[g]\|_1$.

Proof. By Theorem 9.18 (b), we see that for all $[f] \in ([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty)$,

$$\|C_g[f]\|_\infty = \|g * [f]\|_\infty \le \nu_1(g)\;\|[f]\|_\infty,$$

whence $\|C_g\| \le \nu_1(g)$. There remains to show that $\|C_g\| \ge \nu_1(g)$.

Let $f \in \mathcal{C}(\mathbb{T},\mathbb{C})$, with $\|[f]\|_\infty \le 1$. Then $g * [f] \in [\mathcal{C}(\mathbb{T},\mathbb{C})]$, so

$$\|C_g[f]\|_\infty = \|g * [f]\|_\infty = \|g \diamond f\|_{\sup} \ge |g \diamond f(0)|.$$

Next, using the fact that the Lebesgue and Riemann integrals of bounded, continuous functions are equal, we find that

$$\begin{aligned} g \diamond f(0) &= \frac{1}{2\pi} \int_{[-\pi,\pi)} g(s) f(0-s) dm(s) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(s) f(-s) ds. \end{aligned}$$

Case One. Suppose that g is invertible; i.e. $g(s) \neq 0$ for all $s \in \mathbb{R}$. Choose

$$f(s) = \frac{\overline{g(-s)}}{|g(-s)|}, \quad s \in \mathbb{R},$$

so that f is continuous and $\|[f]\|_\infty = \|f\|_{\sup} = 1$.

Furthermore,

$$\begin{aligned} \|C_g\| &\geq \|C_g[f]\|_\infty \\ &\geq |C_g(f)(0)| \\ &= \frac{1}{2\pi} \left| \int_{-\pi}^{\pi} g(s) f(-s) ds \right| \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(s)| ds \\ &= \|[g]\|_1. \end{aligned}$$

Case Two. We shall make use of the following result, which was once a bonus question for the Assignments, but has now been relegated to the Appendix (see Theorem 9.40): let $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$ and $\varepsilon > 0$. Then there exists $h \in \mathcal{C}(\mathbb{T},\mathbb{C})$ such that $h(s) \neq 0$ for all $s \in \mathbb{R}$ and $\|h - g\|_{\sup} < \varepsilon$.

From this it immediately follows that

$$|\|[g]\|_1 - \|[h]\|_1| \leq \|[g] - [h]\|_1 \leq \|[g] - [h]\|_\infty < \varepsilon.$$

But then

$$\|C_g - C_h\| = \|C_{g-h}\| \leq \|[g-h]\|_1 < \varepsilon.$$

By Case One, $\|C_h\| = \|[h]\|_1$, whence

$$\|C_g\| \geq \|C_h\| - \|C_{g-h}\| \geq \|[h]\|_1 - \varepsilon > \|[g]\|_1 - 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary,

$$\|C_g\| \geq \|[g]\|_1,$$

completing the proof.

□

We now establish a similar result for convolution by a continuous function g acting on $L_1(\mathbb{T},\mathbb{C})$.

9.21. Theorem. *Let $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$, and let $C_g : L_1(\mathbb{T},\mathbb{C}) \to L_1(\mathbb{T},\mathbb{C})$ be the convolution operator corresponding to g, so that $C_g[f] = g * [f]$. Then $\|C_g\| = \nu_1(g) = \|[g]\|_1$.*

Proof. By Theorem 9.18 (b), we see that for all $[f] \in L_1(\mathbb{T},\mathbb{C})$,

$$\|C_g[f]\|_1 = \|g * [f]\|_1 \leq \nu_1(g)\ \|[f]\|_1,$$

whence $\|C_g\| \leq \nu_1(g)$.

There remains to show that $\|C_g\| \geq \nu_1(g)$, or equivalently, that $\|C_g\| \geq \|[g]\|_1$. To that end, we consider the functions $f_n = n\pi\chi_{[-1/n,1/n]}$, $n \geq 1$.

Clearly $[f_n] \in L_1(\mathbb{T},\mathbb{C})$ and $\|[f_n]\|_1 = 1$ for all n.

Moreover, for all $\theta \in [-\pi,\pi)$, and hence for all $\theta \in \mathbb{R}$, we have (using Lemma 9.2)

$$\begin{aligned}(g \diamond f_n)(\theta) &= \frac{1}{2\pi}\int_{[-\pi,\pi)} g(s) f_n(\theta - s)dm(s)\\ &= \frac{1}{2\pi}\int_{[-\pi,\pi)} g(\theta - t) f_n(t)dm(t)\\ &= \frac{1}{2\pi}(n\pi)\int_{[-1/n,\,1/n]} g(\theta - t)dm(t)\\ &= \frac{n}{2}\int_{[-1/n,\,1/n]} g(\theta - t)dm(t).\end{aligned}$$

Note also that

$$g(\theta) = \frac{n}{2}\int_{[-1/n,1/n]} g(\theta)\ dm(t),$$

since $g(\theta)$ acts as a constant in this integral.

Thus

$$\begin{aligned}|g \diamond f_n(\theta) - g(\theta)| &= \left|\frac{n}{2}\int_{[-1/n,\,1/n]} g(\theta - t) - g(\theta)\ dm(t)\right|\\ &\leq \frac{n}{2}\int_{[-1/n,\,1/n]} |g(\theta - t) - g(\theta)|dm(t)\end{aligned}$$

Let $\varepsilon > 0$ and choose $1 \leq N \in \mathbb{N}$ such that $|x-y| < \frac{1}{N}$ implies that $|g(x)-g(y)| < \varepsilon$. This is possible because g is uniformly continuous on $\mathbb{R}$ (by virtue of being continuous on $[-\pi,\pi)$ and 2π-periodic).

For $n > N$, $t \in [-\frac{1}{n}, \frac{1}{n}]$ implies that $|(\theta - t) - \theta| \leq \frac{1}{n} < \delta$, and so $|g(\theta - t) - g(\theta)| < \varepsilon$. But then for all $n \geq N$ and for all $\theta \in \mathbb{R}$,

$$|g \diamond f_n(\theta) - g(\theta)| \leq \frac{n}{2}\int_{[-1/n,\,1/n]} \varepsilon\ dm(t) = \varepsilon.$$

Hence

$$\begin{aligned}\nu_1(g \diamond f_n - g) &= \frac{1}{2\pi}\int_{[-\pi,\pi)} |g \diamond f_n(\theta) - g(\theta)|\ dm(\theta)\\ &\leq \frac{1}{2\pi}\int_{[-\pi,\pi)} \varepsilon\ dm(\theta)\\ &= \varepsilon.\end{aligned}$$

From this it follows that in $L_1(\mathbb{T}, \mathbb{C})$,

$$\lim_{n\to\infty} C_g([f_n]) = [g],$$

and so

$$\lim_{n\to\infty} \|C_g([f_n])\|_1 = \|[g]\|_1.$$

But $\|[f_n]\|_1 = 1$ for all $n \geq 1$, and thus

$$\|C_g\| \geq \sup_{n\geq 1} \|C_g([f_n])\|_1 = \|[g]\|_1,$$

completing the proof.

□

In the next couple of chapters, we shall explore the connection between convolution operators and convergence of Fourier series.

Appendix to Section 9.

9.22. As mentioned at the start of the Chapter, an **algebra** $\mathbb{A}$ over a field $\mathbb{F}$ is a vector space over $\mathbb{F}$ which is also a ring. In other words, as well as being a vector space, the algebra $\mathbb{A}$ must also admit a multiplication, denoted by $\cdot$, which satisfies the following properties:

(i) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in \mathbb{A}$;
(ii) $a \cdot (b + c) = a \cdot b + a \cdot c$ for all $a, b, c \in \mathbb{A}$; and
(iii) $(a + b) \cdot c = a \cdot c + b \cdot c$ for all $a, b, c \in \mathbb{A}$.

It often transpires that we simply write ab instead of $a \cdot b$ for the product in an algebra. We shall do so in the examples below.

If there is an element $e \in \mathbb{A}$ such that $e \cdot a = a = a \cdot e$ for all $a \in \mathbb{A}$, then we say that $\mathbb{A}$ is a **unital algebra**, and that e is the **multiplicative identity** for $\mathbb{A}$. There is a simple construction which allows us to embed any algebra $\mathbb{A}$ into a unital algebra $\mathbb{B}$, however in practice this construction is not always natural. In fact, it is precisely in the case that interests us, namely in the case where $\mathbb{A} = L_1(\mathbb{T}, \mathbb{C})$ equipped with convolution as the "multiplication" that the construction leads to an identity which operates as a *point mass* at zero - and there is no natural way of thinking of this as a nice "function" on $\mathbb{R}$.

We are not deterred by the lack of a multiplicative identity in our algebras. Below we shall discuss what is often a suitable and perfectly acceptable alternative.

9.23. Example. Since the notion of an algebra is an algebraic one (a most reasonable state of affairs), it seems only fair to begin with an example which will warm the cockles of every algebraist's heart:

Let $\mathbb{F}$ be a field. We define

$$\mathbb{F}[x] := \{p_0 + p_1 x + p_2 x^2 + \cdots + p_n x^n : p_j \in \mathbb{F}, 1 \le j \le n, n \ge 1\}.$$

Then $\mathbb{F}[x]$ is the set of all polynomials with coefficients in $\mathbb{F}$ in an indeterminate x; using the usual operations governing addition, scalar multiplication, and multiplication of polynomials, $\mathbb{F}[x]$ becomes an algebra over $\mathbb{F}$.

9.24. Example. Let $n \ge 1$ be an integer. Then $\mathbb{A} := \mathbb{M}_n(\mathbb{C})$ is a unital algebra over $\mathbb{C}$, using the usual notions of matrix multiplication, scalar multiplication, and addition. The identity under multiplication is the *identity matrix* $I_n := \text{diag}(1, 1, 1, \ldots, 1)$.

9.25. Example. Let $\mathfrak{X}$ be a Banach space, and set $\mathbb{A} = \mathcal{B}(\mathfrak{X})$. Then $\mathbb{A}$ is a unital algebra, where the multiplication in $\mathcal{B}(\mathfrak{X})$ refers to the composition of linear maps. That is, $(AB)x = A(Bx)$ for all $x \in \mathfrak{X}$, $A, B \in \mathbb{A}$.

The *multiplicative identity* in this case is again the identity map $Ix = x$, $x \in \mathfrak{X}$.

9.26. Example. Consider $\mathbb{A} := \ell_\infty(\mathbb{N}, \mathbb{C})$. Clearly $\mathbb{A}$ is a vector space. We may define a product on $\mathbb{A}$ by setting

$$(x_n)_n \cdot (y_n)_n := (x_n y_n)_n.$$

Using this product, $\mathbb{A}$ is a unital algebra. The space $\mathbb{B} := c_0(\mathbb{N}, \mathbb{C})$ is a **subalgebra** of $\mathbb{A}$. That is, it is a subset of $\mathbb{A}$ and is an algebra using the operations inherited from $\mathbb{A}$.

9.27. Closely related to the notion of an algebra is the notion of a **module**. In its most general incarnation, a (left) module is an abelian group which admits a left-action by a ring. We are only interested in a special case of this phenomenon, and so we shall only define the concept of a module in the setting of interest to us. For more information about modules in general, we refer the reader to [**1**] (just one amongst a cornucopia of books that deal with the subject).

9.28. Definition. *Let $\mathcal{M}$ be a vector space over $\mathbb{K}$, and let $\mathcal{A}$ be an algebra over $\mathbb{K}$. We say that $\mathcal{M}$ is a* ***left-$\mathcal{A}$ module*** *if there exists an operation $\circ : \mathcal{A} \times \mathcal{M} \to \mathcal{M}$ which satisfies:*

(i) $(a+b) \circ m = a \circ m + b \circ m$ *for all* $a, b \in \mathcal{A}$, $m \in \mathcal{M}$;
(ii) $a \circ (m+n) = a \circ m + a \circ n$ *for all* $a \in \mathcal{A}$, $m, n \in \mathcal{M}$; *and*
(iii) $(ab) \circ m = a \circ (b \circ m)$ *for all* $a, b \in \mathcal{A}$ *and* $m \in \mathcal{M}$.

If $\mathcal{A}$ is unital, we also ask that $1 \circ m = m$ for all $m \in \mathcal{M}$.

As the reader might imagine, if $\mathcal{B}$ is an algebra, there exists also the concept of $\mathcal{M}$ being a ***right–$\mathcal{B}$ module****. We leave the definition to the reader's vivid imagination.*

Finally, $\mathcal{M}$ is said to be a $\mathcal{A}-\mathcal{B}$ ***bimodule*** *if it is a left-$\mathcal{A}$ module and a right-$\mathcal{B}$ module. If $\mathcal{A} = \mathcal{B}$, we refer to $\mathcal{M}$ simply as an $\mathcal{A}$ bimodule.*

9.29. Example. Let $\mathcal{A} = \mathbb{M}_3(\mathbb{C})$ and $\mathcal{B} = \mathbb{M}_7(\mathbb{C})$. Then $\mathcal{M} := \mathbb{M}_{3\times 7}(\mathbb{C})$ becomes an $\mathcal{A} - \mathcal{B}$ bimodule, using usual matrix multiplication on the left by elements of $\mathcal{A}$, and usual matrix multiplication on the right by elements of $\mathcal{B}$.

9.30. Example.

(a) Let $\mathcal{A} = \ell_\infty(\mathbb{N}, \mathbb{C})$ and $\mathcal{M} = c_0(\mathbb{N}, \mathbb{C})$. Then $\mathcal{M}$ is a $\mathcal{A}$ bimodule, where we define $(a_n)_n \cdot (m_n)_n := (a_n m_n)_n = (m_n)_n \cdot (a_n)_n$ for all $a = (a_n)_n \in \mathcal{A}$ and $(m_n)_n \in \mathcal{M}$.
(b) We can also set $\mathcal{M} = \ell_\infty(\mathbb{N}, \mathbb{C})$ and $\mathcal{A} = c_0(\mathbb{N}, \mathbb{C})$. Using the same operations as above, $\ell_\infty(\mathbb{N}, \mathbb{C})$ becomes a $c_0(\mathbb{N}, \mathbb{C})$ bimodule.

9.31. Our approach to convolution has been to show that $L_1(\mathbb{T}, \mathbb{C})$ is a bimodule over $\mathcal{C}(\mathbb{T}, \mathbb{C})$, using convolution $*$ as our "multiplication" operation. It is relatively straightforward to prove that conditions (i) and (ii) of Definition 9.34 hold. What is left obvious (and what is left as an assignment exercise) is that condition (iii) holds as well.

For this, it is worth observing that given $g, h \in \mathcal{C}(\mathbb{T}, \mathbb{C})$ and $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$, $h \diamond f \in \mathcal{C}(\mathbb{T}, \mathbb{C})$, and thus $g \diamond (h \diamond f)$ may be realised as a *Riemann integral*, instead of a Lebesgue integral. This is crucial in letting us circumvent the use of Fubini's Theorem, which is the *raison-d'être* of our approach.

9.32. Definition. *We say that a Banach space* $(\mathcal{A}, \|\cdot\|)$ *is a* ***Banach algebra*** *if it is also an algebra, and if*

$$\|ab\| \le \|a\| \ \|b\| \qquad \text{for all } a, b \in \mathcal{A}.$$

The norm condition for the product of a and b above ensures that multiplication is jointly continuous; i.e. the map

$$\begin{array}{rccc} \mu: & \mathcal{A} \times \mathcal{A} & \to & \mathcal{A} \\ & (a,b) & \mapsto & ab \end{array}$$

is continuous.

Some of the above examples of algebras are actually examples of Banach algebras.

9.33. Example. If $\mathfrak{X}$ is a Banach space, then $(\mathcal{B}(\mathfrak{X}), \|\cdot\|)$ is a Banach algebra, as

$$\begin{aligned} \|AB\| &= \sup_{x \in \mathfrak{X}, \|x\| \le 1} \|ABx\| \\ &\le \sup_{x \in \mathfrak{X}, \|x\| \le 1} \|A\| \ \|Bx\| \\ &\le \sup_{x \in \mathfrak{X}, \|x\| \le 1} \|A\| \ \|B\| \\ &= \|A\| \ \|B\|. \end{aligned}$$

9.34. We leave it as an exercise for the reader to prove that $(\mathcal{C}(K,\mathbb{C}), \|\cdot\|_{\sup})$ is a unital Banach algebra, whenever K is a compact, Hausdorff space. Here, functions are added and multiplied pointwise, and $(\kappa f)(x) = \kappa(f(x))$ for all $\kappa \in \mathbb{C}$, $f \in \mathcal{C}(K,\mathbb{C})$.

9.35. As mentioned above, one can extend the notion of convolution to obtain a product operation on $L_1(\mathbb{T},\mathbb{C})$, under which the latter becomes a Banach algebra. Alas, this algebra is non-unital: that is, it does not admit an identity element under this operation.

In the study of Banach algebras, and more specifically of C^*-algebras of operators on a Hilbert space, one often comes across the situation where the algebra is non-unital, but does admit the "next best thing" to a unit, namely a **bounded approximate unit**.

9.36. Definition. *Let $(\mathcal{A}, \|\cdot\|)$ be a Banach algebra. An* ***approximate unit*** *for $\mathcal{A}$ is a net $(e_\lambda)_{\lambda\in\Lambda}$ in $\mathcal{A}$ such that*

$$\lim_\lambda \|e_\lambda a - a\| + \|a e_\lambda - a\| = 0$$

for all $a \in \mathbb{A}$.

We say that $(e_\lambda)_{\lambda\in\Lambda}$ is a ***bounded approximate unit*** *if there exists $M > 0$ such that $\|e_\lambda\| \le M$ for all $\lambda \in \Lambda$.*

Typically, when the algebra in question is separable (i.e. $\mathcal{A}$ admits a countable dense set), the approximate unit may be chosen to be a sequence, rather than a net.

We shall come across examples of these in Chapter 11, hidden under the guise of **summability kernels**.

9.37. Example. Let

$$\mathcal{A} = c_0(\mathbb{N}, \mathbb{C}) = \{(w_n)_{n=1}^\infty : w_n \in \mathbb{C} \text{ for all } n \ge 1 \text{ and } \lim_{n\to\infty} w_n = 0\}.$$

Then $\mathcal{A}$ is a Banach space using pointwise operations, and we can also give it a multiplication operation in a similar way: that is, we set $(w_n)_{n=1}^\infty \cdot (z_n)_{n=1}^\infty = (w_n z_n)_{n=1}^\infty$. It is routine to verify that $(\mathcal{A}, \|\cdot\|_\infty)$ is a Banach algebra.

It is clearly non-unital. However, if we set $e_n = (1, 1, \ldots, 1, 0, 0, 0, \ldots)$ for each $n \ge 1$ (where there are n terms which are equal to 1 and all remaining terms are equal to 0), then $(e_n)_{n=1}^\infty$ is easily seen to be a bounded approximate identity for $\mathcal{A}$.

9.38. There are also discrete versions of convolution; let $g \in \ell_\infty(\mathbb{Z}, \mathbb{C})$ and $f \in \ell_1(\mathbb{Z}, \mathbb{C})$. We may define the **discrete convolution** of g and f as follows:

$$g * f(n) = \sum_{m=-\infty}^{\infty} g(m) f(n-m) = \sum_{m=-\infty}^{\infty} g(n-m) f(m).$$

For those who are interested, this operation turns $(\ell_1(\mathbb{Z}, \mathbb{C}), \|\cdot\|_1)$ into a commutative Banach algebra, using discrete convolution as the multiplication operation. Is it unital?

9.39. Remark. There is a minor subtlety in the proof of Theorem 9.17 that goes along the following lines:

Fix $1 \le m$ an integer. We defined $\Gamma_g^\circ(f_m)$ as a Riemann integral in the Banach space $(\mathcal{C}(\mathbb{T}, \mathbb{C}), \|\cdot\|_{\sup})$. Thus, for an appropriate sequence $(P_N)_N$ of partitions of $[-\pi, \pi]$, we have that

$$\Gamma_g^\circ(f_m) = \lim_{N\to\infty} S(\beta_m, P_N, P_N^*),$$

where $\beta_m(s) = g(s)\tau_s^\circ(f_m)$, $s \in [-\pi, \pi)$.

But given any Riemann sum $S(\beta_m, Q, Q^*)$ of the form

$$\sum_{k=1}^M \beta_m(q_k^*)(q_k - q_{k-1}) = \sum_{k=1}^M g(q_k^*)\tau_{q_k^*}^\circ(f_m)(q_k - q_{k-1})$$

in $\mathcal{C}(\mathbb{T},\mathbb{C})$, its image in $[\mathcal{C}(\mathbb{T},\mathbb{C})]$ is

$$[S(\beta_m, Q, Q^*)] = \sum_{k=1}^{M} g(q_k^*)\tau_{q_k^*}[f_m](q_k - q_{k-1}).$$

Since the map $h \to [h]$ from $(\mathcal{C}(\mathbb{T},\mathbb{C}), \|\cdot\|_{\sup})$ to $([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty)$ is a bijective linear isometry, it follows that the image $[\Gamma_g^\circ(f_m)]$ of $\Gamma_g^\circ(f_m)$ is

$$[\Gamma_g^\circ(f_m)] = \lim_{N\to\infty} [S(\beta_m, P_N, P_N^*)],$$

and that this convergence is with respect to the $\|\cdot\|_\infty$ norm. On the other hand, looking how each $[S(\beta_m, P_N, P_N^*)]$ is defined, we see that the latter limit is precisely how we defined $\Gamma_g([f_m]) \in ([\mathcal{C}(\mathbb{T},\mathbb{C}), \|\cdot\|_\infty)$, and thus

$$[\Gamma_g^\circ(f_m)] = \Gamma_g[f_m], \quad m \ge 1.$$

Next, each $[S(\beta_m, P_N, P_N^*)] \in [\mathcal{C}(\mathbb{T},\mathbb{C})] \subseteq L_1(\mathbb{T},\mathbb{C})$, and $\Gamma_g[f_m] \in [\mathcal{C}(\mathbb{T},\mathbb{C})] \subseteq L_1(\mathbb{T},\mathbb{C})$ as well. As remarked in the proof of Theorem 9.17, $\|[h]\|_1 \le \|[h]\|_\infty$ for all $[h] \in [\mathcal{C}(\mathbb{T},\mathbb{C})]$, and thus

$$0 \le \lim_{N\to\infty} \|[\Gamma_g^\circ(f_m)] - [S(\beta_m, P_N, P_N^*)]\|_1 \le \lim_{N\to\infty} \|[\Gamma_g^\circ(f_m)] - [S(\beta_m, P_N, P_N^*)]\|_\infty = 0,$$

proving that

$$\Gamma_g([f_m]) = [\Gamma_g^\circ(f_m)] = \lim_{N\to\infty} [S(\beta_m, P_N, P_N^*)],$$

with the convergence taking place in $(L_1(\mathbb{T},\mathbb{C}), \|\cdot\|_1)$.

The following result was required for Case Two of Theorem 9.20.

9.40. Theorem. *If $f \in \mathcal{C}(\mathbb{T},\mathbb{C})$ and $\varepsilon > 0$, then there exists $g \in \mathcal{C}(\mathbb{T},\mathbb{C})$ so that $g(x) \ne 0$ for all $x \in [-\pi,\pi]$ and $\|f - g\|_{\sup} < \varepsilon$.*

Remark: Just in case you've forgotten - here's a note to remind you that this is very much a *complex* phenomenon. If we replace complex functions by real-valued functions, the corresponding assertion is false.

A Banach algebra with the property that the invertible elements are dense is said to have **topological stable rank one**. As you might imagine, there is a notion of higher topological stable ranks. When X is a compact topological space, the topological stable rank of $\mathcal{C}(X,\mathbb{K})$ is supposed to be a measure of the dimension of X.

Proof. (I) Let $\varepsilon > 0$. By the Stone-Weierstrass Theorem, we can find a polynomial p_0 so that

$$\sup_{x\in[a,b]} |p_0(x) - f(x)| < \varepsilon/100.$$

By adding a constant function $\kappa\mathbf{1}$ to p_0 with $|\kappa| \le \frac{51}{100}\varepsilon$ if necessary, we can assume that for $p = p_0 + \kappa\mathbf{1}$,

$$\min(|p(a)|, |p(b)|) \ge \varepsilon/2.$$

Observe that

$$|p(a) - p(b)| = |p_0(a) - p_0(b)| \le |p_0(a) - f(a)| + |f(a) - f(b)| + |f(b) - p_0(b)| \le \frac{1}{50}\varepsilon,$$

and that

$$\sup_{x \in [a,b]} |f(x) - p(x)| \le \frac{52}{100}\varepsilon.$$

(II)

Let $p(x) = \alpha_0(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$ for the appropriate choices of α_i, $0 \le i \le n$. If we impose the norm $\|b\|_\infty := \sup_{x\in[a,b]} |b(x)|$ for $b \in \mathbb{C}[x] \subseteq \mathcal{C}([a,b],\mathbb{K})$, then it is clear that the map

$$\begin{array}{rccc} \Phi: & (\mathbb{C}^n, \|\cdot\|_\infty) & \to & (\mathbb{C}[x], \|\cdot\|_\infty) \\ & (\beta_1, \beta_2, \dots, \beta_n) & \mapsto & (x-\beta_1)(x-\beta_2)\cdots(x-\beta_n) \end{array}$$

is continuous. Thus, given $\eta > 0$ we can find $\beta_1, \beta_2, \dots, \beta_n \in \mathbb{C} \setminus \mathbb{R}$ so that $\beta_k - \alpha_k$ is sufficiently small, $1 \le k \le n$, to guarantee that with $q(x) = \alpha_0(x-\beta_1)(x-\beta_2)\cdots(x-\beta_n)$, we have $\|p - q\|_\infty < \eta$. Since $\beta_k \notin \mathbb{R}$ for all $1 \le k \le n$, it follows that q has no real roots - i.e. $q(x) \ne 0$ for all $x \in [a,b]$.

Observe that $\|p - q\|_\infty < \eta$ implies that

$$|q(a) - q(b)| \le |q(a) - p(a)| + |p(a) - p(b)| + |p(b) - q(b)| \le \eta + \frac{1}{50}\varepsilon + \eta.$$

In particular, if we choose $\eta = \frac{1}{100}\varepsilon$, then

$$|q(a) - q(b)| \le \frac{1}{25}\varepsilon.$$

(III)

Now we are in the situation where

- $q : [a,b] \to \mathbb{C}$ is a polynomial and $q(x) \ne 0$ for all $x \in [a,b]$,
- $|q(a)| \ge |p(a)| - \eta \ge \frac{1}{2}\varepsilon - \frac{1}{100}\varepsilon = \frac{49}{100}\varepsilon$, and similarly $|q(b)| \ge \frac{49}{100}\varepsilon$, and
- $$|q(a) - q(b)| \le \frac{1}{25}\varepsilon.$$

Since q is continuous at b, we can find $\delta > 0$ so that $b - \delta \le x \le b$ implies that

$$|q(x) - q(b)| < \frac{1}{25}\varepsilon.$$

Let $r : [a,b] \to \mathbb{C}$ be the continuous function on $[a,b]$ defined by

- $r(x) = 0$ if $x \in [a, b-\delta]$,
- $r(x) = \dfrac{x - (b-\delta)}{\delta}\,(q(a) - q(b))\,, \quad x \in [b-\delta, b]$.

Observe that $|r(x)| \le |q(a) - q(b)| \le \frac{1}{25}\varepsilon$ for all $x \in [b-\delta, b]$, and hence for all $x \in [a,b]$. Let $g(x) = q(x) + r(x)$, $x \in [a,b]$. Clearly g is continuous since each of q and r are continuous, and for $x \in [b-\delta, b]$ we have

$$|g(x)| = |q(x) + r(x)| \ge |q(x)| - |r(x)| \ge \left(|q(b)| - \frac{1}{25}\varepsilon\right) - \frac{1}{25}\varepsilon \ge \frac{41}{100}\varepsilon.$$

In particular, $g(x) \ne 0$ for all $x \in [b-\delta, b]$. Of course, $g(x) = q(x) \ne 0$ for all $x \in [a, b-\delta]$, and thus $g(x) \ne 0$ for all $x \in [a,b]$.

Also,

$$g(a) = q(a) = q(b) + r(b) = g(b).$$

Finally, as we have seen above, $|r(x)| \le |q(a) - q(b)| \le \frac{1}{25}\varepsilon$ for all $x \in [a,b]$, so

$$\begin{aligned}\sup_{x\in[a,b]} |f(x) - g(x)| &\le \sup_{x\in[a,b]} |f(x) - p(x)| + \sup_{x\in[a,b]} |p(x) - q(x)| + \sup_{x\in[a,b]} |q(x) - g(x)| \\ &\le \frac{52}{100}\varepsilon + \frac{1}{100}\varepsilon + \frac{1}{25}\varepsilon \\ &< \varepsilon.\end{aligned}$$

This completes the proof.

□

9.41. Finally, we end this appendix with a statement of Fubini's Theorem, as promised.

9.42. Theorem. ***Fubini's Theorem***

Let (X, μ_X) and (Y, μ_Y) be σ−finite measure spaces, and suppose that $X \times Y$ is given the product measure $\mu_{X\times Y} := \mu_X \times \mu_Y$. If $f : X \times Y \to \mathbb{K}$ lies in $\mathcal{L}_1(X \times Y, \mathbb{K})$, then

$$\int_X \left(\int_Y f(x,y) d\mu_Y(y)\right) d\mu_X(x) = \int_Y \left(\int_X f(x,y) d\mu_X(x)\right) d\mu_Y(y),$$

and these coincide with

$$\int_{X\times Y} f(x,y) d\mu_{X\times Y}(x,y).$$

Exercises for Section 9.

Exercise 9.1. Let $h : \mathbb{R} \to \mathbb{C}$ be a 2π-periodic and continuous function. Prove that h is uniformly continuous on $\mathbb{R}$.

Exercise 9.2. Prove that $L_\infty(\mathbb{T}, \mathbb{C}) \subseteq L_1(\mathbb{T}, \mathbb{C})$ and that

(a) $\|[f]\|_1 \le \|[f]\|_\infty$ for all $[f] \in L_\infty(\mathbb{T}, \mathbb{C})$;
(b) $[\mathrm{Trig}(\mathbb{T}, \mathbb{C})] \subseteq L_\infty(\mathbb{T}, \mathbb{C})$;
(c) for all $[f] \in L_\infty(\mathbb{T}, \mathbb{C})$ and $s \in \mathbb{R}$,
$$\tau_s[f] \in L_\infty(\mathbb{T}, \mathbb{C}); \text{ and}$$
(d) for all $[f] \in L_\infty(\mathbb{T}, \mathbb{C})$ and $s \in \mathbb{R}$,
$$\|\tau_s[f]\|_\infty = \|[f]\|_\infty.$$

Exercise 9.3. Consider the function $\chi_{[0,\pi)} \in \mathcal{L}_\infty([-\pi, \pi))$, and let $f = \widehat{\chi_{[0,\pi)}} \in \mathcal{L}_\infty(\mathbb{T}, \mathbb{C})$ be its 2π-periodic extension, as defined in Paragraph 8.4.

If $-\pi < s < 0$, prove that
$$\|\tau_s[f] - \tau_0[f]\|_\infty = 1.$$

Exercise 9.4. Let $(\mathfrak{X}, \|\cdot\|)$ be a Banach space and $g : \mathbb{R} \to \mathfrak{X}$ be a 2π-periodic, continuous function.

Prove that for all $s \in \mathbb{R}$,
$$\int_{-\pi}^{\pi} g(\theta) d\theta = \int_{-\pi}^{\pi} g(\theta - s) ds.$$

Exercise 9.5. Let $g \in \mathcal{C}(\mathbb{T}, \mathbb{C})$. Prove that the convolution operator $C_g : L_1(\mathbb{T}, \mathbb{C}) \to L_1(\mathbb{T}, \mathbb{C})$ defined by $C_g[f] = g * [f]$ is linear.

Exercise 9.6. What does Fubini's Theorem say when $X = Y = \mathbb{N}$ and $\mu_X = \mu_Y$ denote counting measure on $\mathbb{N}$? That is, $\mu_X(E)$ denotes the cardinality of $E \subseteq \mathbb{N}$? (It should seem like a very familiar result!)

Exercise 9.7. (Assignment Exercise)

Let $g, h \in \mathcal{C}(\mathbb{T}, \mathbb{C})$ and $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$. Prove that for all $\theta \in \mathbb{R}$,
$$(g \diamond (h \diamond f))(\theta) = ((g \diamond h) \diamond f)(\theta).$$
Conclude that $L_1(\mathbb{T}, \mathbb{C})$ is a bimodule over $\mathcal{C}(\mathbb{T}, \mathbb{C})$.

Exercise 9.8. Let $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$ and $g \in \mathcal{L}_\infty(\mathbb{T}, \mathbb{C})$. Define
$$\begin{array}{rccc} g \diamond f : & \mathbb{R} & \to & \mathbb{C} \\ & \theta & \mapsto & \frac{1}{2\pi} \int_{[-\pi,\pi)} g(s) f(\theta - s) dm(s). \end{array}$$

Prove that $|g \diamond f(\theta)| \le \nu_\infty(g) \nu_1(f)$ for all $\theta \in \mathbb{R}$, and that $g \diamond f$ is continuous on $\mathbb{R}$.

10. The Dirichlet kernel

I just want to thank everyone who made this day necessary.

Yogi Berra

10.1. We began our discussion of Fourier series by noting that if $[f] \in L_2(\mathbb{T}, \mathbb{C})$, then the sequence $(\Delta_N([f]))_{N=1}^{\infty}$ of partial sums of the Fourier series of $[f]$ converges in the $\|\cdot\|_2$-norm to $[f]$ (see Paragraph 8.3). Our goal was to see to what extent we could extend these results to elements $[f] \in L_1(\mathbb{T}, \mathbb{C})$. Somewhere along the way, we seem to have been distracted by the concept of convolution. Let us show that all roads lead to the Dirichlet kernel, which we now define.

10.2. Definition. *For each $n \in \mathbb{Z}$, recall that $\xi_n \in \mathcal{C}(\mathbb{T}, \mathbb{C})$ is the function $\xi_n(\theta) = e^{in\theta}$. For $N \geq 1$, we define the* ***Dirichlet kernel of order*** *N to be*

$$D_N = \sum_{n=-N}^{N} \xi_n.$$

We mention in passing that this use of the word "*kernel*" has nothing to do with the null space of any linear map. It is just another example of the overuse of certain terminology in mathematics.

10.3. Let $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$. For each $N \geq 1$, we shall define

$$\Delta_N^{\circ}(f) = \sum_{n=-N}^{N} \alpha_n^{[f]} \xi_n = \sum_{n=-N}^{N} \widehat{f}(n) \xi_n.$$

It is clear that $\Delta_N^{\circ}(f) \in \mathcal{C}(\mathbb{T}, \mathbb{C})$, being a finite linear combination of $\{\xi_n\}_{n=-N}^{N}$.

If $f = g$ a.e. on $\mathbb{R}$, then $\alpha_n^{[f]} = \alpha_n^{[g]}$ for all $n \in \mathbb{Z}$, and thus $\Delta_N^{\circ}(f) = \Delta_N^{\circ}(g)$ for all $N \geq 1$. We may therefore define

$$\Delta_N([f]) = [\Delta_N^{\circ}(f)], \qquad N \geq 1.$$

Thus $\Delta_N([f])$ is the N^{th}-partial sum of the Fourier series of $[f]$. Note that when $[f] \in L_2(\mathbb{T}, \mathbb{C})$, this definition coincides with our previous definition.

Now, for $N \geq 1$, $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$ and $\theta \in \mathbb{R}$,

$$\begin{aligned}
\Delta_N^\circ(f)(\theta) &= \sum_{n=-N}^{N} \alpha_n^{[f]} \xi_n \\
&= \sum_{n=-N}^{N} \left(\frac{1}{2\pi} \int_{[-\pi,\pi)} f(s) \overline{\xi_n}(s) \; dm(s) \right) \xi_n(\theta) \\
&= \sum_{n=-N}^{N} \frac{1}{2\pi} \int_{[-\pi,\pi)} f(s) e^{in(\theta - s)} \; dm(s) \\
&= \sum_{n=-N}^{N} \frac{1}{2\pi} \int_{[-\pi,\pi)} f(\theta - s) e^{ins} \; dm(s) \qquad \text{by our Assignments} \\
&= \frac{1}{2\pi} \int_{[-\pi,\pi)} \sum_{n=-N}^{N} f(\theta - s) e^{ins} \; dm(s) \\
&= \frac{1}{2\pi} \int_{[-\pi,\pi)} D_N(s) f(\theta - s) \; dm(s) \\
&= (D_N \diamond f)(\theta).
\end{aligned}$$

In other words, $\Delta_N^\circ(f) = D_N \diamond f$, or

$$\Delta_N([f]) = D_N * [f] = C_{D_N}([f]), \quad N \geq 1.$$

We have expressed the N^{th}-partial sum of the Fourier series of $[f] \in L_1(\mathbb{T}, \mathbb{C})$ as the convolution of the Dirichlet kernel D_N of order N with $[f]$. *Suddenly our peripatetic meanderings through the land of convolution do not seem as unwarranted!*

The question of whether or not these partial sums converge to $[f]$ in $L_1(\mathbb{T}, \mathbb{C})$ now becomes a question of whether or not $\lim_{N\to\infty} C_{D_N}([f]) = [f]$ in $L_1(\mathbb{T}, \mathbb{C})$. In order to answer this question, we shall examine the nature of the Dirichlet kernel a bit closer, and borrow a couple of results from our previous real analysis courses.

10.4. Theorem. *Let $N \geq 1$ be an integer, and let D_N denote the Dirichlet kernel of order N. Then*

(a) $D_N(-\theta) = D_N(\theta) \in \mathbb{R}$ *for all* $\theta \in \mathbb{R}$.

(b) $\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(\theta) \; d\theta = 1$.

(c) *For* $0 \neq \theta \in [-\pi, \pi)$,

$$D_N(\theta) = \frac{\sin((N + \frac{1}{2})\theta)}{\sin(\frac{1}{2}\theta)}.$$

Also, $D_N(0) = 2N + 1$.

(d) $\|[D_N]\|_1 = \nu_1(D_N) \geq \dfrac{4}{\pi^2} \displaystyle\sum_{n=1}^{N} \frac{1}{n}$.

Proof.

(a) For all $\theta \in \mathbb{R}$ and $n \geq 1$, $\xi_{-n}(\theta) + \xi_n(\theta) = e^{-in\theta} + e^{in\theta} = 2\cos(n\theta)$. Thus

$$D_N(\theta) = \sum_{n=-N}^{N} \xi_n(\theta) = 1 + 2\sum_{n=1}^{N} \cos(n\theta) \in \mathbb{R}.$$

From this it is also clear that $D_N(-\theta) = D_N(\theta)$ for all $\theta \in \mathbb{R}$.

(b) Now

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(\theta)\ d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} 1\ d\theta + 2\sum_{n=1}^{N} \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(n\theta)\ d\theta.$$

We leave it as an exercise for the reader to show that $\int_{-\pi}^{\pi} \cos(n\theta)\ d\theta = 0$, $n \geq 1$, and so

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(\theta)\ d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} 1\ d\theta = 1.$$

(c) Let $\theta \in \mathbb{R}$. Set $\rho(\theta) := e^{-i\theta/2} - e^{i\theta/2} = (-2i)\sin(\frac{1}{2}\theta)$. A routine calculation shows that $\rho \cdot D_N$ involves a telescoping sum, so that

$$\rho(\theta)D_N(\theta) = e^{-i(N+\frac{1}{2})\theta} - e^{i(N+\frac{1}{2})\theta} = (-2i)\sin\Big((N + \frac{1}{2})\theta\Big).$$

If $0 \neq \theta \in [-\pi, \pi)$, then

$$D_N(\theta) = \frac{(-2i)\sin\big((N+\frac{1}{2})\theta\big)}{\rho(\theta)} = \frac{\sin((N+\frac{1}{2})\theta)}{\sin(\frac{1}{2}\theta)}.$$

Meanwhile,

$$D_N(0) = 1 + 2\sum_{n=1}^{N} \cos(0) = 2N + 1.$$

(d) Since D_N is an even, continuous function, and since $|\sin(x)| \leq |x|$, $0 \leq x \leq \pi$,

$$\begin{aligned}
\nu_1(D_N) &= \frac{1}{2\pi}\int_{[-\pi,\pi)} |D_N| \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_N(\theta)|\ d\theta \\
&= \frac{1}{\pi}\int_{0}^{\pi} |D_N(\theta)|\ d\theta \\
&= \frac{1}{\pi}\int_{0}^{\pi} \left|\frac{\sin((N+\frac{1}{2})\theta)}{\sin(\frac{1}{2}\theta)}\right|\ d\theta \\
&\geq \frac{1}{\pi}\int_{0}^{\pi} \frac{|\sin((N+\frac{1}{2})\theta)|}{|\frac{1}{2}\theta|} \\
&= \frac{2}{\pi}\int_{0}^{\pi} \frac{|\sin((N+\frac{1}{2})\theta)|}{|\theta|}.
\end{aligned}$$

We are dealing with Riemann integration here, and so substitution of variables is permissible. Let $\lambda = (N+\frac{1}{2})\theta$, so that $d\lambda = (N+\frac{1}{2})d\theta$. Then

$$\begin{aligned}
\nu_1(D_N) &\geq \frac{2}{\pi}\int_0^{\pi} \frac{|\sin((N+\frac{1}{2})\theta)|}{|\theta|} \\
&\geq \frac{2}{\pi}\int_0^{(N+\frac{1}{2})\pi} \frac{|\sin\lambda|}{|\lambda/(N+\frac{1}{2})|}\ \frac{1}{(N+\frac{1}{2})} d\lambda \\
&\geq \frac{2}{\pi}\int_0^{N\pi} \frac{|\sin\lambda|}{|\lambda|}\ d\lambda \\
&= \frac{2}{\pi}\sum_{n=1}^{N}\int_{(n-1)\pi}^{n\pi} \frac{|\sin\lambda|}{|\lambda|}\ d\lambda \\
&\geq \frac{2}{\pi}\sum_{n=1}^{N}\int_{(n-1)\pi}^{n\pi} \frac{|\sin\lambda|}{n\pi}\ d\lambda \\
&\geq \frac{2}{\pi^2}\sum_{n=1}^{N}\frac{1}{n}\int_{(n-1)\pi}^{n\pi} |\sin\lambda|\ d\lambda
\end{aligned}$$

But on any interval of the form $[(n-1)\pi, n\pi]$, the sine function does not change sign (i.e. it is either always non-positive on such an interval, or always non-negative), and so an easy calculation shows that

$$\int_{(n-1)\pi}^{n\pi} |\sin\lambda|\ d\lambda = \left|\int_{(n-1)\pi}^{n\pi} \sin\lambda\ d\lambda\right| = 2.$$

From this we see that

$$\nu_1(D_N) \geq \frac{4}{\pi^2}\sum_{n=1}^{N}\frac{1}{n},$$

as required.

□

We invite the reader to skip to the Appendix to see the graphs of D_2, D_5 and D_{10}.

The next result follows immediately from Theorems 9.18 and 9.20, together with Theorem 10.4 (d) and the fact that the harmonic series $\sum_{n=1}^{\infty}\frac{1}{n}$ diverges.

10.5. Corollary. *For each $N \geq 1$, let D_N denote the Dirichlet kernel of order N.*

(a) *If $C_{D_N} \in \mathcal{B}(([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty))$ is the convolution operator corresponding to D_N, $N \geq 1$, then*
$$\lim_{N\to\infty} \|C_{D_N}\| = \infty.$$

(b) *If $C_{D_N} \in \mathcal{B}((L_1(\mathbb{T},\mathbb{C}), \|\cdot\|_1))$ is the convolution operator corresponding to D_N, $N \geq 1$, then*
$$\lim_{N\to\infty} \|C_{D_N}\| = \infty.$$

As mentioned above, in order to exploit the connection between the Dirichlet kernel and convolution, we shall require a couple of results from a previous real analysis course. We start by recalling a definition.

10.6. Definition. *Let (X, d) be a metric space and $H \subseteq X$. We say that H is* ***nowhere dense*** *(or* ***meager****, or* ***thin****) if $G := X \setminus \overline{H}$ is dense in X. In other words, the interior of $\overline{H}$ is empty.*

Note: Here, $\overline{H}$ refers to the closure of H in X.

10.7. Examples. We think of nowhere dense subsets of metric spaces as being "small", as the alternate terminology "meager" and "thin" suggest.

(a) The set $H = \mathbb{Z}$ is nowhere dense in $\mathbb{R}$, as is easily verified.
(b) The Cantor set C is nowhere dense in $X = [0, 1]$, equipped with the standard metric inherited from $\mathbb{R}$. This is left as an exercise for the reader.
(c) The set $H = \mathbb{Q}$ of rational numbers is not nowhere dense in $\mathbb{R}$, as $X \setminus \overline{H} = \mathbb{R} \setminus \mathbb{R} = \varnothing$, which is as far from being dense in $\mathbb{R}$ as any set can get.

10.8. Definition. *We say that a subset H of a metric space (X, d) is of the* ***first category*** *in (X, d) if there exists a sequence $(F_n)_{n=1}^\infty$ of closed, nowhere dense sets in X such that*

$$H \subseteq \cup_{n=1}^\infty F_n.$$

Otherwise, H is said to be of the ***second category****.*

We assume that the reader is familiar with the following result, whose proof we consign to the Appendix.

10.9. Theorem. ***The Baire Category Theorem***

A complete metric space (X, d) is of the second category. That is, X is not a countable union of closed, nowhere dense sets in X.

10.10. Examples.

(a) It follows from the Baire Category Theorem that $(\mathbb{R}, d)$ is of the second category, where d represents the standard metric $d(x, y) = |x - y|$, $x, y \in \mathbb{R}$.
(b) Let δ denote the **discrete metric** on $\mathbb{Q}$, so that

$$\delta(p, q) = \begin{cases} 1 & \text{if } p \neq q \\ 0 & \text{if } p = q. \end{cases}$$

Writing $\mathbb{Q} = \{q_n\}_{n=1}^\infty$, which we may do since $\mathbb{Q}$ is denumerable, we find that for each $n \geq 1$, $F_n := \{q_n\}$ is a closed set, and clearly $\mathbb{Q} = \cup_{n=1}^\infty F_n$.

However, F_n is *not* nowhere dense, since as well as being closed, F_n is open, and $q_n \in F_n = \text{int}(\overline{F_n}) \neq \varnothing$.

This is just as well, since $(\mathbb{Q}, \delta)$ is a complete metric space, and hence of the second category by the Baire Category Theorem. In other words, we knew that $\mathbb{Q}$ was not a countable union of closed, nowhere dense sets in $(\mathbb{Q}, \delta)$.

(c) The sets $F_n = \{q_n\}$ from (b) are closed and nowhere dense in $(\mathbb{R}, d)$, where d is the standard metric from (a). Since $\mathbb{Q} = \cup_{n=1}^\infty F_n$, we see that $\mathbb{Q}$ is is of the first category in $(\mathbb{R}, d)$.

10.11. Remark. An alternate form of the Baire Category Theorem says that if (X, d) is a complete metric space, and if $(G_n)_{n=1}^\infty$ is a countable collection of dense, open sets in X, then

$$\cap_{n=1}^\infty G_n \neq \varnothing.$$

We shall not require this below.

The second result from real analysis which we shall recall is the following.

10.12. Theorem. ***The Banach-Steinhaus Theorem, aka The Uniform Boundedness Principle***

Let (X, d) be a complete metric space and $\varnothing \neq \mathcal{F} \subseteq \mathcal{C}(X, \mathbb{R})$. Suppose that for all $x \in X$, there exists a constant $\kappa_x > 0$ such that

$$|f(x)| \leq \kappa_x, \qquad f \in \mathcal{F}.$$

Then there exists a non-empty open set $G \subseteq X$ and $\kappa > 0$ such that

$$|f(x)| \leq \kappa, \qquad x \in G, \ f \in \mathcal{F}.$$

Proof. See the Appendix to this Chapter.

□

There is a stronger version of the Banach-Steinhaus Theorem which applies to the setting of linear operators on Banach spaces. Before stating and proving it, we remind the reader that a Banach space is a complete metric space under the metric induced by the norm.

10.13. Theorem. ***The Banach-Steinhaus Theorem, aka The Uniform Boundedness Principle for Operators***

Let $(\mathfrak{X}, \|\cdot\|_{\mathfrak{X}})$ and $(\mathfrak{Y}, \|\cdot\|_{\mathfrak{Y}})$ be Banach spaces and suppose that $\varnothing \neq \mathcal{F} \subseteq \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. Let $H \subseteq \mathfrak{X}$ be a subset of the second category in $\mathfrak{X}$, and suppose that for each $x \in H$, there exists a constant $\kappa_x > 0$ such that

$$\|Tx\|_{\mathfrak{Y}} \leq \kappa_x, \qquad T \in \mathcal{F}.$$

Then $\mathcal{F}$ is bounded; that is,

$$\sup_{T \in \mathcal{F}} \|T\| < \infty.$$

Proof. For each $n \geq 1$, set

$$F_n := \{x \in \mathfrak{X} : \|Tx\|_{\mathfrak{Y}} \leq n \text{ for all } T \in \mathcal{F}\}.$$

If $(x_k)_{k=1}^\infty$ is a sequence in F_n and if $x = \lim_{k \to \infty} x_k$ exists in $\mathfrak{X}$, then by the continuity of each $T \in \mathcal{F}$,

$$Tx = \lim_{k \to \infty} Tx_k,$$

from which it easily follows that

$$\|Tx\| = \lim_{k\to\infty} \|Tx_k\| \le n.$$

Hence $x \in F_n$, proving that F_n is closed.

By hypothesis,

$$H \subseteq \cup_{n=1}^{\infty} F_n.$$

Our hypothesis says that H is of the second category, and hence there must exist $1 \le N$ such that $\operatorname{int}(\overline{F_N}) = \operatorname{int}(F_N) \ne \varnothing$.

Let $x_0 \in \operatorname{int}(F_N)$, and choose $\delta > 0$ such that

$$B(x_0, \delta) := \{x \in \mathfrak{X} : \|x - x_0\|_{\mathfrak{X}} < \delta\} \subseteq F_N.$$

Then $\|Tx\|_{\mathfrak{Y}} \le N$ for all $x \in B(x_0, \delta)$ and $T \in \mathcal{F}$.

Suppose that $w \in \mathfrak{X}$ and $\|w\| \le 1$. Then $x_0 - \frac{\delta}{2} w \in B(x_0, \delta)$, and so for all $T \in \mathcal{F}$,

$$\|T(x_0 - \frac{\delta}{2} w)\|_{\mathfrak{Y}} \le N.$$

By the triangle inequality, $\frac{\delta}{2}\|Tw\|_{\mathfrak{Y}} - \|Tx_0\|_{\mathfrak{Y}} \le N$. Thus

$$\|Tw\|_{\mathfrak{Y}} \le \frac{2}{\delta}(N + \|Tx_0\|_{\mathfrak{Y}}).$$

Setting $K := \frac{2}{\delta}(N + \|Tx_0\|_{\mathfrak{Y}})$, we see that

$$\sup_{T\in\mathcal{F}} \|T\| \le K < \infty.$$

□

10.14. Corollary. *Let $(\mathfrak{X}, \|\cdot\|_{\mathfrak{X}})$ and $(\mathfrak{Y}, \|\cdot\|_{\mathfrak{Y}})$ be Banach spaces and let $(T_n)_{n=1}^{\infty}$ be an unbounded sequence in $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$, i.e. $\sup_{n\ge1} \|T_n\| = \infty$.*

Let $H = \{x \in \mathfrak{X} : \sup_{n\ge1} \|T_n x\| < \infty\}$. Then H is of the first category in $\mathfrak{X}$, and $J := \mathfrak{X} \setminus H$ is of the second category.

Proof. Note that $0 \in H$, so that $H \ne \varnothing$.

If H were of the second category, then – by the Banach-Steinhaus Theorem for Operators, Theorem 10.13 above – $\{T_n\}_{n=1}^{\infty}$ would be bounded, a contradiction.

Thus H is of the first category. Let $J = \mathfrak{X} \setminus H$. By definition, for all $x \in J$,

$$\sup_{n\ge1} \|T_n x\|_{\mathfrak{Y}} = \infty.$$

We claim that J is of the second category.

Indeed, suppose otherwise. Then we could choose sequences $(K_n)_{n=1}^{\infty}$ and $(L_n)_{n=1}^{\infty}$ of closed, nowhere dense sets in $\mathfrak{X}$ such that

$$H \subseteq \cup_{n=1}^{\infty} K_n \qquad \text{and} \qquad J \subseteq \cup_{n=1}^{\infty} L_n.$$

But then

$$\mathfrak{X} = H \cup J = \cup_{n=1}^{\infty} K_n \bigcup \cup_{n=1}^{\infty} L_n$$

must be of the first category, contradicting the Baire Category Theorem, as $\mathfrak{X}$ is a complete metric space.

□

10.15. It is perhaps worth noting that this result is much, much better than it might appear on the surface. The statement that $\sup_{n\ge1} \|T_n\| = \infty$ is the statement that for each $n \ge 1$, there exists $x_n \in \mathfrak{X}$ with $\|x_n\|_{\mathfrak{X}} = 1$ such that $\lim_{n\to\infty} \|T_n x_n\|_{\mathfrak{Y}} = \infty$. *A priori*, it is not clear that there should exist *any* $x \in \mathfrak{X}$ such that

$$\lim_{n\to\infty} \|T_n x\|_{\mathfrak{Y}} = \infty.$$

The above Corollary not only says that such a vector $x \in \mathfrak{X}$ exists; it asserts that this is true for a very large set of x's, in the sense that the set H of x's for which it fails is a set of the first category in $\mathfrak{X}$.

We are finally prepared to answer the question of whether or not the partial sums of the Fourier series of an element $[f]$ of $L_1(\mathbb{T},\mathbb{C})$ necessarily converge to $[f]$ in the $\|\cdot\|_1$-norm. As we shall now see, an easy application of Corollary 10.14 shows that this *almost never happens*. (Here we use "almost never" informally, to refer to the notion of sets of the first category, and not in the sense of sets of Lebesgue measure zero!). What is more, essentially the same argument shows that the partial sums of the Fourier series of an element $[f]$ of $[\mathcal{C}(\mathbb{T},\mathbb{C})]$ rarely converge in the $\|\cdot\|_\infty$ norm. These are dark times indeed for the sequence of partial sums of a Fourier series.

10.16. Theorem. ***The unbearable lousiness of being a Dirichlet kernel***

(a) *Let*

$$\mathfrak{K}_\infty := \{[f] \in [\mathcal{C}(\mathbb{T},\mathbb{C})] : [f] = \lim_{N\to\infty} \Delta_N([f]) \text{ in } ([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty)\}.$$

Then $\mathfrak{K}_\infty$ is a set of the first category in $([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty)$, whose complement $[\mathcal{C}(\mathbb{T},\mathbb{C})] \setminus \mathfrak{K}_\infty$ is a set of the second category.

(b) *Let*

$$\mathfrak{K}_1 := \{[f] \in L_1(\mathbb{T},\mathbb{C}) : [f] = \lim_{N\to\infty} \Delta_N([f]) \text{ in } (L_1(\mathbb{T},\mathbb{C}), \|\cdot\|_1)\}.$$

Then $\mathfrak{K}_1$ is a set of the first category in $(L_1(\mathbb{T},\mathbb{C}), \|\cdot\|_1)$, whose complement $L_1(\mathbb{T},\mathbb{C}) \setminus \mathfrak{K}_1$ is a set of the second category in $L_1(\mathbb{T},\mathbb{C})$.

Proof. The proofs of both of these parts are almost identical. We shall prove (b), and leave the proof of (a) as an exercise.

Recall that $\Delta_N([f]) = D_N * [f] = C_{D_N}([f])$, $N \ge 1$, where $C_{D_N} \in \mathcal{B}(L_1(\mathbb{T},\mathbb{C}))$ is the convolution operator corresponding to D_N, as described in Theorem 9.20. Furthermore, by Corollary 10.5,

$$\lim_{N\to\infty} \|C_{D_N}\| = \infty.$$

Of course, if $[h] \in L_1(\mathbb{T},\mathbb{C})$ and $[h] = \lim_{N\to\infty} \Delta_N([h]) = \lim_{N\to\infty} C_{D_N}([h])$, then $(C_{D_N}([h]))_{N=1}^\infty$ is bounded in $L_1(\mathbb{T},\mathbb{C})$, and thus it is clear that $\mathfrak{K}_1 \subseteq \mathfrak{H}_1$, where

$$\mathfrak{H}_1 := \{[f] \in L_1(\mathbb{T},\mathbb{C}) : \sup_{N\ge1} \|C_{D_N}([f])\|_1 < \infty\}.$$

By Corollary 10.14, $\mathfrak{H}_1$ is a set of the first category in $L_1(\mathbb{T},\mathbb{C})$, and $\mathfrak{J}_1 := L_1(\mathbb{T},\mathbb{C}) \setminus \mathfrak{H}_1$ is a set of the second category in $L_1(\mathbb{T},\mathbb{C})$.

In particular, for any $[f] \in \mathfrak{J}_1$, we have that the partial sums $(\Delta_N([f]))_{N=1}^{\infty}$ fail to converge to $[f]$, as the sequence is not even bounded.

$\square$

Appendix to Section 10.

10.17. It is instructive to look at the graph of the Dirichlet kernels of various orders. Two things worth noticing are that

- first, the amplitude of the function is increasing near 0; this is clear since each D_N is continuous, and $D_N(0) = 2N + 1$, $N \geq 1$.
- Each of the functions D_N spends a lot of time being negative, and a lot of time being positive. This accounts for the fact that the integrals of D_N are bounded, but that the integrals of $|D_N|$ are not.

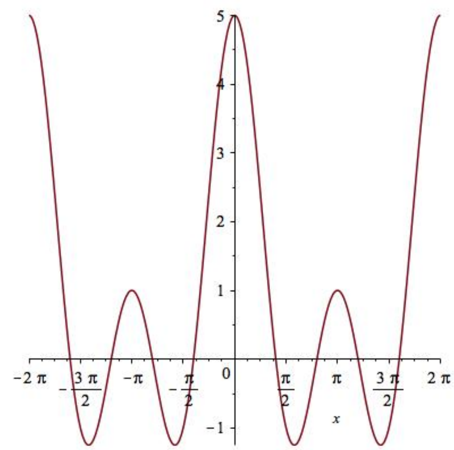


FIGURE 3. The graph of D_2.

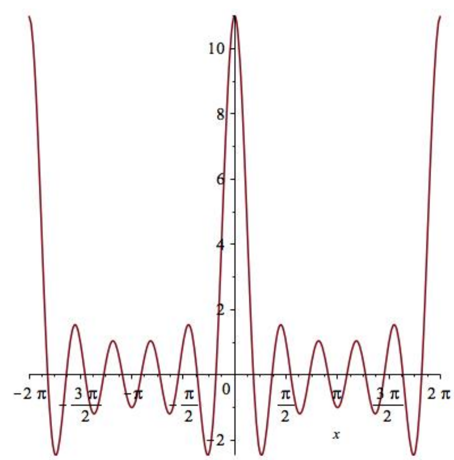


FIGURE 4. The graph of D_5.

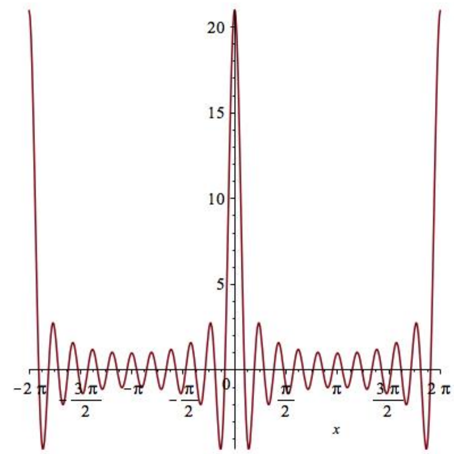


FIGURE 5. THE GRAPH OF D_{10}.

Exercises for Section 10.

Exercise 10.1.

Prove that the Cantor set is nowhere dense in $[0,1]$, where $[0,1]$ is equipped with the standard metric $d(x,y) = |x-y|$ inherited from $\mathbb{R}$.

Exercise 10.2.

Fill in the details of the proof of Theorem 10.4(c) by proving that with $\rho(\theta) := (-2i)\sin(\frac{1}{2}\theta)$,

$$\rho(\theta) D_N(\theta) = e^{-i(N+\frac{1}{2})\theta} - e^{i(N+\frac{1}{2})\theta} = (-2i)\sin\left((N+\frac{1}{2})\theta\right).$$

Exercise 10.3.

Fill in the details of the proof of Theorem 10.16 (a), namely: let

$$\mathfrak{K}_\infty := \{[f] \in [\mathcal{C}(\mathbb{T},\mathbb{C})] : [f] = \lim_{N\to\infty} \Delta_N([f]) \text{ in } ([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty)\}.$$

Prove that $\mathfrak{K}_\infty$ is a set of the first category in $([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty)$, whose complement $[\mathcal{C}(\mathbb{T},\mathbb{C})] \setminus \mathfrak{K}_\infty$ is a set of the second category.

11. The Féjer kernel

I've never really wanted to go to Japan. Simply because I don't like eating fish. And I know that's very popular out there in Africa.

Britney Spears

11.1. We have seen in the last Chapter that if $[f] \in L_1(\mathbb{T},\mathbb{C})$, and if $\Delta_N([f]) = \sum_{n=-N}^{N} \alpha_n^{[f]}[\xi_n]$, then $(\Delta_N([f]))_{N=1}^{\infty}$ almost never converges to $[f]$.

Not all is lost. In this Chapter we shall replace partial sums $(\Delta_N([f]))_{N=1}^{\infty}$ of the Fourier series of $[f]$ by weighted partial sums which *will* converge to $[f]$.

11.2. Definition. *Let $(\mathfrak{X}, \|\cdot\|_{\mathfrak{X}})$ be a Banach space, and $(x_n)_{n=0}^{\infty}$ be a sequence in $\mathfrak{X}$. The N^{th}-**Cesàro mean** of the sequence is*

$$\sigma_N := \frac{1}{N}(x_0 + x_1 + \cdots + x_{N-1}), \quad N \ge 1.$$

The next Proposition is routine, and its proof is left to the exercises.

11.3. Proposition. *Suppose that $\mathfrak{X}$ is a Banach space and $(x_n)_{n=0}^{\infty}$ is a sequence in $\mathfrak{X}$. Let $(\sigma_N)_{N=1}^{\infty}$ denote the sequence of Cesàro means of $(x_n)_{n=0}^{\infty}$.*

If $x = \lim_{n\to\infty} x_n$ exists, then $x = \lim_{N\to\infty} \sigma_N$.

11.4. Definition. *Let $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$. The N^{th}-**Cesàro sum** of the Fourier series of f is the N^{th}-Cesàro mean of the sequence $(\Delta_n^{\circ}(f))_{n=0}^{\infty}$. Thus*

$$\begin{aligned}\sigma_N^{\circ}(f) &= \frac{1}{N}(D_0 \diamond f + D_1 \diamond f + \cdots + D_{N-1} \diamond f)\\ &= F_N \diamond f,\end{aligned}$$

*where $F_N := \frac{1}{N}(D_0 + D_1 + \cdots + D_{N-1})$ is called the **Féjer kernel of order** N.*

*We also define the N^{th}-**Cesàro sum** of the Fourier series of $[f]$ is the N^{th}-Cesàro mean of the sequence $(\Delta_n[f]))_{n=0}^{\infty}$, namely*

$$\begin{aligned}\sigma_N[f] &:= \frac{1}{N}(D_0 * [f] + D_1 * [f] + \cdots + D_{N-1} * [f])\\ &= F_N * [f]\\ &= [F_N \diamond f]\\ &= [\sigma_N^{\circ}(f)].\end{aligned}$$

11.5. Remark. We remark that the fact that $D_n \in \mathcal{C}(\mathbb{T},\mathbb{C})$ for all $n \ge 0$ implies that $F_N \in \mathcal{C}(\mathbb{T},\mathbb{C})$ for all $N \ge 1$. By Proposition 9.6, it follows that $\sigma_N^\circ(f) \in \mathcal{C}(\mathbb{T},\mathbb{C}) \subseteq \mathcal{L}_1(\mathbb{T},\mathbb{C})$ for all $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$.

Furthermore, for all $\theta \in \mathbb{R}$,

$$\begin{aligned}\sigma_N^\circ(f)(\theta) &= \frac{1}{2\pi}\int_{[-\pi,\pi)} F_N(s) f(\theta - s)\ dm(s)\\ &= \frac{1}{2\pi}\int_{[-\pi,\pi)} F_N(\theta - s) f(s)\ dm(s)\end{aligned}$$

As seen in Theorem 9.18, the fact that $F_N \in \mathcal{C}(\mathbb{T},\mathbb{C})$ also implies that for every homogeneous Banach space $\mathfrak{B}$ and $[f] \in \mathfrak{B}$, we have

$$\sigma_N[f] = F_N * [f] \in \mathfrak{B}.$$

As a special case of this phenomenon, $\sigma_N[f] = F_N * [f] \in L_p(\mathbb{T},\mathbb{C})$ for all $[f] \in L_p(\mathbb{T},\mathbb{C})$.

11.6. Theorem. *For each $1 \le N \in \mathbb{N}$,*

(a) *F_N is a 2π-periodic, even, continuous function;*

(b) *If $0 \ne \theta \in [-\pi,\pi)$, then*

$$F_N(\theta) = \frac{1}{N}\left(\frac{1-\cos(N\theta)}{1-\cos\theta}\right) = \frac{1}{N}\left(\frac{\sin(\frac{N}{2}\theta)}{\sin(\frac{1}{2}\theta)}\right)^2,$$

while $F_N(0) = N$. In particular, $F_N(\theta) \ge 0$ for all $\theta \in \mathbb{R}$;

(c) $\nu_1(F_N) = \frac{1}{2\pi}\int_{-\pi}^{\pi} |F_N(\theta)|\, d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} F_N(\theta)\, d\theta = 1.$

(d) *For all $0 < \delta \le \pi$,*

$$\lim_{N\to\infty}\left(\int_{-\pi}^{-\delta} |F_N(\theta)|\ d\theta + \int_{\delta}^{\pi} |F_N(\theta)|\ d\theta\right) = 0;$$

(e) *For $0 < |\theta| < \pi$,*

$$0 \le F_N(\theta) \le \frac{\pi^2}{N\theta^2}.$$

Proof.

(a) That F_N is 2π-periodic, even and continuous follows immediately from the fact that each D_n is, $0 \le n \le N-1$.

(b) First,

$$\begin{aligned}F_N(0) &= \frac{1}{N}(\sum_{n=0}^{N-1} D_n(0))\\ &= \frac{1}{N}(\sum_{n=0}^{N-1}(2n+1))\\ &= \frac{1}{N}(2\frac{(N-1)N}{2} + N)\\ &= N.\end{aligned}$$

For $0 \neq \theta \in [-\pi, \pi)$,

$$\begin{aligned} F_N &= \frac{1}{N} \sum_{n=0}^{N-1} D_n \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \Big(\sum_{k=-n}^{n} \xi_k \Big) \\ &= \frac{1}{N} \big(\xi_{-N+1} + 2\xi_{-N+2} + \cdots + (N-1)\xi_{-1} + N\xi_0 \\ &\qquad\qquad + (N-1)\xi_1 + \cdots + 2\xi_{N-2} + \xi_{N-1} \big) . \end{aligned}$$

Let $\rho = (2 - (\xi_{-1} + \xi_1))N$. As was the case with the Dirichlet kernel, we observe that the product $\rho \cdot F_N$ involves a telescoping sum, and that

$$\rho(\theta)\, F_N(\theta) = 2 - (\xi_{-N}(\theta) + \xi_N(\theta)) = 2 - 2\cos(N\theta).$$

That is,

$$\begin{aligned} F_N(\theta) &= \frac{1}{N} \, \frac{1 - \cos(N\theta)}{1 - \cos\theta} \\ &= \frac{1}{N} \, \frac{\left(e^{i\frac{N}{2}\theta} - e^{-i\frac{N}{2}\theta} \right)^2}{\left(e^{i\frac{\theta}{2}} - e^{-i\frac{\theta}{2}} \right)^2} \\ &= \frac{1}{N} \left(\frac{\sin(\frac{N}{2}\theta)}{\sin(\frac{1}{2}\theta)} \right)^2 . \end{aligned}$$

In particular, $F_N \geq 0$.

(c) Keeping in mind that $0 \neq n$ implies that $\int_{-\pi}^{\pi} \xi_n(\theta)\, d\theta = 0$, we get:

$$\begin{aligned} \nu_1(F_N) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |F_N(\theta)| \, d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} F_N(\theta) \, d\theta \\ &= \frac{1}{2\pi} \frac{1}{N} \int_{-\pi}^{\pi} \left(\sum_{n=0}^{N-1} \left(\sum_{k=-n}^{n} \xi_k(\theta) \right) \right) d\theta \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=-n}^{n} \frac{1}{2\pi} \int_{-\pi}^{\pi} \xi_k(\theta) \, d\theta \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \xi_0(\theta) \, d\theta \\ &= \frac{1}{N} \sum_{n=0}^{N-1} 1 \\ &= 1. \end{aligned}$$

(d) Let $0 < \delta \le \pi$. For $\delta \le |\theta| \le \pi$, we have that

$$\left|\sin\frac{\theta}{2}\right| \ge \sin\frac{\delta}{2},$$

and so

$$\int_{-\pi}^{-\delta} F_N(\theta)\ d\theta = \frac{1}{N}\int_{-\pi}^{-\delta}\left(\frac{\sin(\frac{N}{2}\theta)}{\sin(\frac{1}{2}\theta)}\right)^2\ d\theta \le \frac{1}{N}(\pi-\delta)\frac{1}{(\sin\frac{\delta}{2})^2}.$$

Thus

$$\lim_{N\to\infty}\int_{-\pi}^{-\delta} F_N(\theta)\, d\theta = 0.$$

Similarly,

$$\lim_{N\to\infty}\int_{\delta}^{\pi} F_N(\theta)\, d\theta = 0,$$

from which (d) easily follows.

(e) Finally, we leave it as a routine calculus exercise for the reader to verify that on the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$, we have

$$\frac{2}{\pi}|\theta| \le |\sin\theta|,$$

and so

$$\begin{aligned} F_N(\theta) &= \frac{1}{N}\left(\frac{\sin(\frac{N}{2}\theta)}{\sin(\frac{1}{2}\theta)}\right)^2 \\ &\le \frac{1}{N}\left|\frac{1}{\frac{2}{\pi}\frac{\theta}{2}}\right|^2 \\ &= \frac{\pi^2}{N\theta^2}. \end{aligned}$$

□

Our goal is to show that the Cesàro sums of the Fourier series of an element $[f]$ of $L_1(\mathbb{T},\mathbb{C})$ converge in the $\|\cdot\|_1$-norm back to $[f]$. There is nothing unique about the Féjer kernel, however. Let us examine a more general phenomenon, of which the Féjer kernel is but an example.

11.7. Definition. *A **summability kernel** is a sequence $(k_n)_{n=1}^\infty$ of 2π-periodic, continuous, complex-valued functions on $\mathbb{R}$ satisfying:*

(a) $\frac{1}{2\pi}\int_{-\pi}^{\pi} k_n = 1$ *for all* $n \ge 1$*;*

(b) $\sup_{n\ge1}\nu_1(k_n) = \sup_{n\ge1}\frac{1}{2\pi}\int_{-\pi}^{\pi}|k_n| < \infty$*; and*

(c) *for all* $0 < \delta \le \pi$*,*

$$\lim_{n\to\infty}\left(\int_{-\pi}^{-\delta}|k_n| + \int_{\delta}^{\pi}|k_n|\right) = 0.$$

*If, in addition, $k_n \ge 0$ for all $n \ge 1$, we say that $(k_n)_{n=1}^\infty$ is a **positive summability kernel**.*

11.8. Theorem. *The Féjer kernel $(F_N)_{N=1}^\infty$ is a positive summability kernel.*
Proof. This is the content of Theorem 11.6.

□

11.9. Examples.

(a) For each $1 \le n \in \mathbb{N}$, consider the piecewise linear function

$$\begin{array}{rccc} k_n^\bullet : & [-\pi, \pi) & \to & \mathbb{R} \\ & \theta & \mapsto & \begin{cases} 0 & \text{if } \theta \in [-\pi, \frac{-1}{n}] \cup [\frac{1}{n}, \pi) \\ n + n^2\theta & \text{if } \theta \in (\frac{-1}{n}, 0] \\ n - n^2\theta & \text{if } \theta \in (0, \frac{1}{n}). \end{cases} \end{array}$$

For $1 \le n \in \mathbb{N}$, let k_n be the 2π-periodic function on $\mathbb{R}$ whose restriction to the interval $[-\pi, \pi)$ coincides with $k_n^\bullet$.

Then $(k_n)_{n=1}^\infty$ is a positive summability kernel. The details are left to the reader.

(b) For each $1 \le n \in \mathbb{N}$, consider the piecewise linear function

$$\begin{array}{rccc} r_n^\bullet : & [-\pi, \pi) & \to & \mathbb{R} \\ & \theta & \mapsto & \begin{cases} 0 & \text{if } \theta \in [-\pi, 0] \cup [\frac{2}{n}, \pi) \\ n^2\theta & \text{if } \theta \in (0, \frac{1}{n}] \\ n - n^2(\theta - \frac{1}{n}) & \text{if } \theta \in (\frac{1}{n}, \frac{2}{n}). \end{cases} \end{array}$$

For $1 \le n \in \mathbb{N}$, let r_n be the 2π-periodic function on $\mathbb{R}$ whose restriction to the interval $[-\pi, \pi)$ coincides with $r_n^\bullet$.

Then $(r_n)_{n=1}^\infty$ is a positive summability kernel. The details are left to the reader.

11.10. Theorem. *Let $(\mathfrak{B}, \|\cdot\|_{\mathfrak{B}})$ be a homogeneous Banach space over $\mathbb{T}$ and let $(k_n)_{n=1}^\infty$ be a summability kernel. If $[f] \in \mathfrak{B}$, then*

$$\lim_{n\to\infty} \|k_n * [f] - [f]\|_{\mathfrak{B}} = 0,$$

and so $[f] = \lim_{n\to\infty} k_n * [f]$ in $\mathfrak{B}$.
Proof. The result is trivial if $[f] = 0$. Let $0 \ne [f] \in \mathfrak{B}$. Recall that

(a) the function

$$\begin{array}{rccc} \Psi_{[f]} : & \mathbb{R} & \to & \mathfrak{B} \\ & s & \mapsto & \tau_s[f] \end{array}$$

is continuous, and that

(b) τ_s is isometric for all $s \in \mathbb{R}$. In particular, $\|\tau_s[f]\|_{\mathfrak{B}} = \|[f]\|_{\mathfrak{B}}$ for all $s \in \mathbb{R}$.

That $\tau_0[f] = [f]$ is clear from the definition of τ_0. Let $M := \sup_{n\ge1} \nu_1(k_n) < \infty$, as $(k_n)_{n=1}^\infty$ is a summability kernel.

Let $\varepsilon > 0$ and choose $\delta > 0$ such that $|s - 0| < \delta$ implies that

$$\|\tau_s[f] - \tau_0[f]\|_{\mathfrak{B}} < \frac{\varepsilon}{2M}.$$

This is possible by (a) above. Next, choose $1 \le N \in \mathbb{N}$ such that $n \ge N$ implies that

$$\frac{1}{2\pi}\int_{-\pi}^{-\delta} |k_n(s)| ds + \frac{1}{2\pi}\int_{\delta}^{\pi} |k_n(s)| ds < \frac{\varepsilon}{4\|[f]\|_{\mathfrak{B}}}.$$

Then $n \ge N$ implies that

$$\begin{aligned}\|k_n * [f] - [f]\|_{\mathfrak{B}} &= \|\frac{1}{2\pi}\int_{-\pi}^{\pi} k_n(s)(\tau_s[f] - \tau_0[f]) ds\|_{\mathfrak{B}} \\ &\le \frac{1}{2\pi}\int_{-\pi}^{-\delta} |k_n(s)| \; \|\tau_s[f] - \tau_0[f]\|_{\mathfrak{B}} \; ds \\ &\quad + \frac{1}{2\pi}\int_{-\delta}^{\delta} |k_n(s)| \; \|\tau_s[f] - \tau_0[f]\|_{\mathfrak{B}} \; ds \\ &\quad + \frac{1}{2\pi}\int_{\delta}^{\pi} |k_n(s)| \; \|\tau_s[f] - \tau_0[f]\|_{\mathfrak{B}} \; ds.\end{aligned}$$

Now $\|\tau_s[f] - \tau_0[f]\|_{\mathfrak{B}} \le \|\tau_s[f]\|_{\mathfrak{B}} + \|\tau_0[f]\|_{\mathfrak{B}} = 2\|[f]\|_{\mathfrak{B}}$, and thus for $n \ge N$, we have

$$\begin{aligned}\|k_n * [f] - [f]\|_{\mathfrak{B}} &\le \frac{1}{2\pi}\int_{-\pi}^{-\delta} |k_n(s)|(2\|[f]\|_{\mathfrak{B}}) \; ds \\ &\quad + \frac{1}{2\pi}\int_{-\delta}^{\delta} |k_n(s)| \frac{\varepsilon}{2M} \; ds \\ &\quad + \frac{1}{2\pi}\int_{-\pi}^{-\delta} |k_n(s)|(2\|[f]\|_{\mathfrak{B}}) \; ds \\ &\le 2\|[f]\|_{\mathfrak{B}} \frac{\varepsilon}{4\|[f]\|_{\mathfrak{B}}} + M \frac{\varepsilon}{2M} \\ &= \varepsilon.\end{aligned}$$

In other words,

$$\lim_{n\to\infty} k_n * [f] = [f].$$

□

11.11. Corollary.

(a) *For each* $f \in (\mathcal{C}(\mathbb{T}, \mathbb{C}), \|\cdot\|_{\sup})$,

$$\lim_{N\to\infty} \sigma_N^{\circ}(f) = f.$$

(b) *Let* $1 \le p < \infty$*. For each* $[g] \in (L_p(\mathbb{T}, \mathbb{C}), \|\cdot\|_p)$,

$$\lim_{N\to\infty} \sigma_N[g] = [g].$$

Proof.

(a) By Example 9.11, we know that $([\mathcal{C}(\mathbb{T}, \mathbb{C})], \|\cdot\|_\infty)$ is a homogeneous Banach space and that the map

$$\begin{array}{rccc}\Gamma : & (\mathcal{C}(\mathbb{T}, \mathbb{C}), \|\cdot\|_{\sup}) & \to & ([\mathcal{C}(\mathbb{T}, \mathbb{C})], \|\cdot\|_\infty) \\ & f & \mapsto & [f]\end{array}$$

is an isometric isomorphism of Banach spaces.

Let $f \in \mathcal{C}(\mathbb{T}, \mathbb{C})$. By Theorem 11.8, $(F_N)_{N=1}^{\infty}$ is a (positive) summability kernel, and by definition, $\sigma_N([f]) = F_N * [f]$. By Theorem 11.10 above,

$$\lim_{N\to\infty} \|F_N \diamond f - f\|_{\sup} = \lim_{N\to\infty} \|F_N * [f] - [f]\|_{\infty} = 0.$$

(b) Again, by Example 9.12, for each $1 \le p < \infty$, $(L_p(\mathbb{T}, \mathbb{C}), \|\cdot\|_p)$ is a homogeneous Banach space over $\mathbb{T}$. Let $1 \le p < \infty$ and $[f] \in L_p(\mathbb{T}, \mathbb{C})$. Since $(F_N)_{N=1}^{\infty}$ is a (positive) summability kernel, and since $\sigma_N([f]) = F_N * [f]$ for all $N \ge 1$,

$$\lim_{N\to\infty} \sigma_N([f]) = \lim_{N\to\infty} F_N * [f] = [f]$$

by Theorem 11.10.

□

We are now in a position to show that the Fourier coefficients of an $\mathcal{L}_p(\mathbb{T}, \mathbb{C})$-function completely determine that function (almost everywhere).

11.12. Corollary. *Let $1 \le p < \infty$. If $[f], [g] \in L_p(\mathbb{T}, \mathbb{C})$ and $\alpha_n^{[f]} = \alpha_n^{[g]}$ for all $n \in \mathbb{Z}$, then $[f] = [g]$.*

Proof. It is clear that if $\alpha_n^{[f]} = \alpha_n^{[g]}$ for all $n \in \mathbb{Z}$, then $\sigma_N[f] = \sigma_N([g])$ for all $N \ge 1$.

By Corollary 11.11,

$$[f] = \lim_{N\to\infty} \sigma_N([f]) = \lim_{N\to\infty} \sigma_N([g]) = [g].$$

□

11.13. Local structure and Féjer kernels. Corollaries 11.11 and 11.12 tell us that Féjer kernels are nice enough to recover an element of $L_p(\mathbb{T}, \mathbb{C})$ from its Fourier coefficients. As we have emphasized in these notes, however, these elements are *equivalence classes* of functions in $\mathcal{L}_p(\mathbb{T}, \mathbb{C})$, and not functions themselves. In other words, the aforementioned Corollaries say that we can recover functions in $\mathcal{L}_p(\mathbb{T}, \mathbb{C})$ *almost everywhere* on $\mathbb{R}$. We now turn our attention to the functions themselves, and study in what sense (if any) the convolution of $\mathcal{L}_p(\mathbb{T}, \mathbb{C})$ functions with Féjer kernels converge *pointwise*.

11.14. Definition. *Given $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$ and $\theta \in \mathbb{R}$, we set*

$$\omega_f(\theta) := \lim_{t\to 0^+} \frac{f(\theta - t) + f(\theta + t)}{2},$$

*provided that the limit exists. When it does exist, we shall refer to this value as the **average value of f at θ**.*

11.15. Theorem. [**Féjer's Theorem.**] *Let* $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$.

(a) *If* $\theta \in \mathbb{R}$ *and* $\omega_f(\theta)$ *exists, then*

$$\lim_{N \to \infty} F_N \diamond f(\theta) = \omega_f(\theta).$$

(b) *Suppose that there exists a closed interval* $[a, b] \subseteq [-\pi, \pi)$ *such that* f *is continuous on* $[a, b]$ *(in particular,* f *is continuous from the right at* a *and from the left at* b*), then*

$$(F_N \diamond f)_{N=1}^{\infty}$$

converges uniformly to f *on* $[a, b]$.

Proof.

(a) Let $\varepsilon > 0$ and choose $\delta > 0$ such that $0 < |s| < \delta$ implies that

$$|\omega_f(\theta) - \frac{f(\theta - s) + f(\theta + s)}{2}| < \varepsilon.$$

Now

$$\begin{aligned} |\sigma_N(f)(\theta) - \omega_f(\theta)| &= \left| \frac{1}{2\pi} \int_{[-\pi,\pi)} F_N(s) f(\theta - s)\ dm(s) - \omega_f(\theta) \right| \\ &= \left| \frac{1}{2\pi} \int_{[-\pi,\pi)} F_N(s)(f(\theta - s) - \omega_f(\theta))\ dm(s) \right| \\ &\le \left| \frac{1}{2\pi} \int_{[-\delta,\delta]} F_N(s)(f(\theta - s) - \omega_f(\theta))\ dm(s) \right| \\ &\qquad + \left| \frac{1}{2\pi} \int_{[-\pi,-\delta)\cup(\delta,\pi]} F_N(s)(f(\theta - s) - \omega_f(\theta))\ dm(s) \right|. \end{aligned}$$

By our work in the Assignments,

$$\int_{[-\delta,\delta]} F_N(s)(f(\theta - s) - \omega_f(\theta))\ dm(s) = \int_{[-\delta,\delta]} F_N(-s)(f(\theta + s) - \omega_f(\theta))\ dm(s).$$

But F_N is even, so $F_N(-s) = F_N(s)$, $s \in [-\delta, \delta]$. Thus

$$\frac{1}{2\pi} \int_{[-\delta,\delta]} F_N(s)(f(\theta - s) - \omega_f(\theta))\ dm(s) = \frac{1}{2\pi} \int_{[-\delta,\delta]} F_N(s) \left(\frac{f(\theta - s) + f(\theta + s)}{2} - \omega_f(\theta) \right)\ dm(s).$$

Thus

$$\left|\frac{1}{2\pi}\int_{[-\delta,\delta]} F_N(s)(f(\theta-s)-\omega_f(\theta))\ dm(s)\right| \le$$

$$\frac{1}{2\pi}\int_{[-\delta,\delta]} F_N(s)\left|\frac{f(\theta-s)+f(\theta+s)}{2}-\omega_f(\theta)\right|\ dm(s)$$

$$\le \frac{1}{2\pi}\int_{[-\delta,\delta]} F_N(s)\varepsilon\ dm(s)$$

$$\le \varepsilon\left(\frac{1}{2\pi}\int_{[-\pi,\pi)} F_N(s)\ dm(s)\right)$$

$$= \varepsilon.$$

Meanwhile, for $\delta \le |s| \le \pi$,

$$0 \le F_N(s) \le \frac{\pi^2}{Ns^2} \le \frac{\pi^2}{N\delta^2},$$

and so

$$\left|\frac{1}{2\pi}\int_{[-\pi,-\delta)\cup(\delta,\pi]} F_N(s)(f(\theta-s)-\omega_f(\theta))\ dm(s)\right|$$

$$\le \frac{\pi^2}{N\delta^2}\left(\frac{1}{2\pi}\int_{[-\pi,-\delta)\cup(\delta,\pi]} |f(\theta-s)|+|\omega_f(\theta)|\ dm(s)\right)$$

$$\le \frac{\pi^2}{N\delta^2}\left(\frac{1}{2\pi}\int_{[-\pi,\pi)} |f(\theta-s)|\ dm(s)+|\omega_f(\theta)|\right)$$

$$\le \frac{\pi^2}{N\delta^2}\left(\|[f]\|_1+|\omega_f(\theta)|\right),$$

which converges to 0 as N tends to ∞.

Thus $\lim_{N\to\infty}\sigma_N(f)(\theta)=\omega_f(\theta)$.

(b) The proof is essentially identical to that above.

First note that the continuity of f at θ for $\theta\in[a,b]$ implies that $\omega_f(\theta) = f(\theta)$ for all $\theta\in[a,b]$. But f continuous on $[a,b]$ implies that f is uniformly continuous on $[a,b]$, and so we may find a single $\delta>0$ such that $0\le|s|>\delta$ implies that

$$|f(\theta)-\frac{f(\theta-s)+f(\theta+s)}{2}|<\varepsilon, \quad \theta\in[a,b].$$

For this $\delta>0$, and for *all* $\theta\in[a,b]$, the estimates from part (a) above show that

$$|\sigma_N(f)(\theta)-f(\theta)| \le \frac{1}{2\pi}\int_{[-\delta,\delta]} F_N(s)\left|\frac{f(\theta-s)+f(\theta+s)}{2}-f(\theta)\right|\ dm(s)$$

$$+\frac{1}{2\pi}\int_{[-\pi,-\delta)\cup(\delta,\pi]} F_N(s)\big(|f(\theta-s)|+|f(\theta)|\big)\ dm(s)$$

$$<\varepsilon+\frac{\pi^2}{N\delta^2}\left(\|[f]\|_1+|f(\theta)|\right),$$

and thus $(\sigma_N(f))_{N=1}^\infty$ converges uniformly to f on $[a,b]$.

□

11.16. Corollary.

(a) *If* $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ *and* f *is continuous at* $\theta_0 \in \mathbb{R}$*, then*

$$\lim_{N\to\infty} F_N \diamond f(\theta_0) = f(\theta_0).$$

(b) *If* $f \in \mathcal{C}(\mathbb{T},\mathbb{C})$*, then*

$$(F_N \diamond f)_{N=1}^\infty$$

converges uniformly to f *on* $\mathbb{R}$*.*

(c) *Let* $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ *and* $\theta_0 \in \mathbb{R}$*. If* f *is continuous at* θ_0 *and* $(D_N \diamond f(\theta_0))_{N=1}^\infty$ *converges, then*

$$\lim_{N\to\infty} D_N \diamond f(\theta_0) = f(\theta_0).$$

Proof.

(a) If f is continuous at θ_0, then $f(\theta_0) = \omega_f(\theta_0)$ and so by Theorem 11.15,

$$\lim_{N\to\infty} F_N \diamond f(\theta_0) = \omega_f(\theta_0) = f(\theta_0).$$

(b) This is exactly Corollary 11.11 (a).

(c) Again, if f is continuous at θ_0, then $f(\theta_0) = \omega_f(\theta_0)$.

Moreover, if $(D_N \diamond f(\theta_0))_{N=1}^\infty$ converges to some value $\beta \in \mathbb{C}$, then $(F_N \diamond f(\theta_0))_{N=1}^\infty$ converges to β by Proposition 11.3. But (a) shows that $\beta = f(\theta_0)$, completing the proof.

□

11.17. Culture. Let $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$. A point $\theta \in \mathbb{R}$ is called a **Lebesgue-point** of f if

$$\lim_{h\to 0^+} \frac{1}{h} \int_{[0,h]} \left| \frac{f(\theta - s) + f(\theta + s)}{2} - f(\theta) \right| dm(s) = 0.$$

It can be shown that almost every real number θ is a Lebesgue point of f.

A modification of the proof of Féjer's Theorem yields

The Lebesgue-Féjer Theorem.

If $\theta \in \mathbb{R}$ is a Lebesgue point for $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$, then

$$f(\theta) = \lim_{N\to\infty} F_N \diamond f(\theta).$$

In particular,

$$f(\theta) = \lim_{N\to\infty} F_N \diamond f(\theta)$$

almost everywhere on $\mathbb{R}$.

11.18. So far we have seen that while it is extremely rare for the partial sums of the Fourier series of a continuous function f, or of an element $[g]$ of $L_1(\mathbb{T},\mathbb{C})$ to converge to f uniformly (or to $[g]$ in the $\|\cdot\|_1$-norm), nevertheless, this is always the case regarding the Cesàro sums of the Fourier series.

Furthermore, Féjer's Theorem shows that if $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ and if the average value $\omega_f(\theta)$ exists at $\theta \in \mathbb{R}$, then

$$\omega_f(\theta) = \lim_{N\to\infty} \sigma_N(f)(\theta) = \lim_{N\to\infty} F_N \diamond f(\theta).$$

Our next goal is to show that if f is sufficiently "*smooth*" at a point θ_0 (in a sense which we shall now make explicit), then in fact the *partial sums* of f at θ_0 converge to $f(\theta_0)$.

11.19. Definition. *Suppose that $f : \mathbb{R} \to \mathbb{C}$ is measurable and that $\theta_0 \in \mathbb{R}$. We say that f is* ***locally Lipschitz*** *at θ_0 if there exist $M > 0$ and $\delta > 0$ such that*

$$|f(\theta_0 + s) - f(\theta_0)| < M|s| \qquad \textit{for all } 0 \le |s| < \delta.$$

11.20. Example. Suppose that $f : \mathbb{R} \to \mathbb{C}$ is measurable and that f admits left- and right-sided derivatives at $\theta_0 \in \mathbb{R}$, in the sense that there exist $y_1, y_2 \in \mathbb{C}$ such that

$$\lim_{s\to 0^-} \left| \frac{f(\theta_0 + s) - f(\theta_0)}{s} - y_1 \right| = 0 = \lim_{s\to 0^+} \left| \frac{f(\theta_0 + s) - f(\theta_0)}{s} - y_1 \right|.$$

Let $0 < \varepsilon < 1$, and choose $\delta > 0$ such that

- $-\delta < s < 0$ implies that $\left| \frac{f(\theta_0 + s) - f(\theta_0)}{s} - y_1 \right| < \varepsilon$, and
- $0 < s < \delta$ implies that $\left| \frac{f(\theta_0 + s) - f(\theta_0)}{s} - y_2 \right| < \varepsilon$.

(That is, find $\delta_1 > 0$, $\delta_2 > 0$ that work for y_1 and y_2 respectively, and let $\delta = \min(\delta_1, \delta_2)$.)

Then $-\delta < s < \delta$ implies that

$$|f(\theta_0 + s) - f(\theta_0)| < (\varepsilon + \max(|y_1|, |y_2|))\ |s| < M|s|,$$

where $M = 1 + \max(|y_1|, |y_2|)$ is a fixed constant.

Thus f is locally Lipschitz at θ_0.

In particular, if f is actually differentiable at $\theta_0 \in \mathbb{R}$, then f is locally Lipschitz at θ_0.

11.21. Theorem. *Let $f \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$ and suppose that f is locally Lipschitz at $\theta_0 \in \mathbb{R}$. Then*

$$\lim_{N\to\infty} s_N(f)(\theta_0) = \lim_{N\to\infty} D_N \diamond f(\theta_0) = f(\theta_0).$$

Proof. Fix $M > 0$ and $0 < \delta < \frac{\pi}{2}$ as in the definition of locally Lipschitz such that $0 \le |s| < \delta$ implies that

$$|f(\theta_0 + s) - f(\theta_0)| \le M|s|.$$

Now,

$$s_N(f)(\theta_0) = \frac{1}{2\pi}\int_{[-\pi,\pi)} D_N(s) f(\theta_0 - s) dm(s),$$

and

$$1 = \frac{1}{2\pi}\int_{[-\pi,\pi)} D_N,$$

and so

$$\begin{aligned}|s_N(f)(\theta_0) - f(\theta_0)| &= \left|\frac{1}{2\pi}\int_{[-\pi,\pi)} (f(\theta_0+s) - f(\theta_0)) D_N(s)\ dm(s)\right| \\ &\le \frac{1}{2\pi}\int_{[-\delta,\delta]} |f(\theta_0+s) - f(\theta_0)||D_N(s)|\ dm(s) \\ &\quad \frac{1}{2\pi}\left|\int_{[-\pi,-\delta]\cup[\delta,\pi)} (f(\theta_0+s) - f(\theta_0)) D_N(s)\ dm(s)\right|\end{aligned}$$

Now, since $0 \le |s| < \delta \le \frac{\pi}{2}$, we have that $\frac{2}{\pi}|s| \le |\sin s|$, and thus

$$\frac{1}{\pi}|s| \le |\sin\frac{s}{2}|, \quad 0 \le |s| < \delta.$$

Thus

$$\begin{aligned}&\frac{1}{2\pi}\int_{[-\delta,\delta]} |f(\theta_0+s) - f(\theta_0)||D_N(s)|\ dm(s) \\ &\qquad \le \frac{1}{2\pi}\int_{[-\delta,\delta]} \frac{|f(\theta_0+s) - f(\theta_0)|}{|s|/\pi} |\sin(N+\frac{1}{2})s|\ dm(s) \\ &\qquad \le \frac{1}{2}\int_{[-\delta,\delta]} \frac{M|s|}{|s|}\ dm(s) \\ &\qquad = M\delta,\end{aligned}$$

independent of N!!!

Next we consider $\delta \le |s| \le \pi$. Observe that

$$\begin{aligned}(f(\theta_0+s) - f(\theta_0)) D_N(s) &= \frac{f(\theta_0+s) - f(\theta_0)}{\sin(s/2)} \sin((N+\frac{1}{2})s) \\ &= \frac{f(\theta_0+s) - f(\theta_0)}{\sin(s/2)} \left(\frac{e^{is/2}}{2i} e^{iNs} - \frac{e^{-is/2}}{2i} e^{-iNs}\right).\end{aligned}$$

Define

$$g_1(s) = \chi_{[-\pi,-\delta)\cup(\delta,\pi)} \frac{f(\theta_0+s)-f(\theta_0)}{\sin(s/2)} \frac{e^{is/2}}{2i}$$

and

$$g_2(s) = \chi_{[-\pi,-\delta)\cup(\delta,\pi)} \frac{f(\theta_0+s)-f(\theta_0)}{\sin(s/2)} \frac{e^{-is/2}}{2i}.$$

Then g_1, g_2 are measurable (why?).

Also,

$$\begin{aligned}
\frac{1}{2\pi}\int_{[-\pi,\pi)} |g_1| &= \frac{1}{2\pi}\int_{[-\pi,-\delta)\cup(\delta,\pi)} \frac{|f(\theta_0+s)-f(\theta_0)|}{|\sin s/2|}\frac{1}{2} \\
&\le \frac{1}{4\pi}\frac{|f(\theta_0+s)|+|f(\theta_0)|}{|\sin\delta/2|} \\
&= \frac{1}{2\pi|\sin\delta/2|}\left(\|[f]\|_1 + 2\pi|f(\theta_0)|\right) \\
&< \infty.
\end{aligned}$$

Thus $g_1 \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$, and similarly, $g_2 \in \mathcal{L}_1(\mathbb{T},\mathbb{C})$.

By the Riemann-Lebesgue Lemma 8.9, we have that

$$\lim_{N\to\infty} \frac{1}{2\pi}\int_{[-\pi,\pi)} g_1(s)e^{iNs}\ dm(s) = 0$$

and

$$\lim_{N\to\infty} \frac{1}{2\pi}\int_{[-\pi,\pi)} g_2(s)e^{-iNs}\ dm(s) = 0.$$

From this it easily follows that

$$\lim_{N\to\infty} \frac{1}{2\pi}\int_{[-\pi,-\delta)\cup(\delta,\pi)} (f(\theta_0+s)-f(\theta_0))D_N(s)\ dm(s) = 0.$$

Thus, given $\varepsilon > 0$, we may choose $\delta > 0$ and $N \ge 1$ such that

(i) $M\delta < \dfrac{\varepsilon}{2}$, and

(ii) for all $n \ge N$ we have

$$\left|\frac{1}{2\pi}\int_{[-\pi,-\delta)\cup(\delta,\pi)} (f(\theta_0+s)-f(\theta_0))D_N(s)\ dm(s)\right| < \frac{\varepsilon}{2}.$$

Then $n \ge N$ implies that

$$|s_N(f)(\theta_0) - f(\theta_0)| < \varepsilon,$$

or in other words,

$$\lim_{N\to\infty} s_N(f)(\theta_0) = f(\theta_0).$$

□

11.22. Example. Consider the function $f(\theta) = |\theta|$, $\theta \in [-\pi, \pi)$, extended 2π-periodically to all of $\mathbb{R}$. Clearly f is continuous on $\mathbb{R}$.

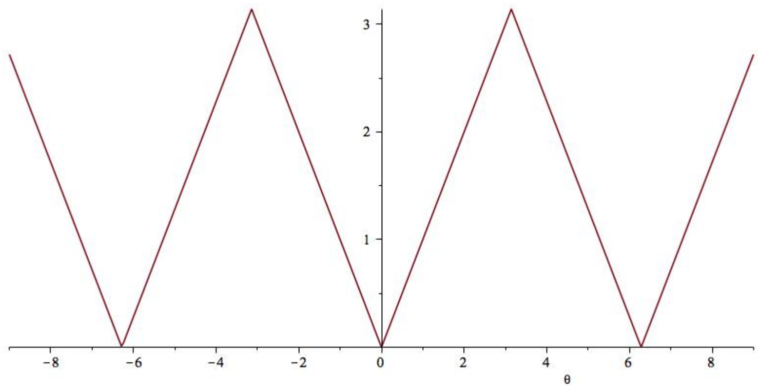


FIGURE 6. THE GRAPH OF f.

Since f is continuous and linear on $(-\pi, 0) \cup (0, \pi)$, it is locally Lipschitz there. It is also easy to see that f is locally Lipschitz at $\theta = n\pi$, $n \in \mathbb{Z}$ with Lipschitz constant $M = 1$.

We are preparing to do the unthinkable: we will calculate $s_N(f)$, $N \geq 1$. To do this, we must first calculate $\alpha_n^{[f]}$, $n \in \mathbb{Z}$. To quote Fourier himself, "buckle up, cupcake".

We first point out that f is Riemann integrable and bounded over $[-\pi, \pi)$, and that multiplying f by $\xi_n = e^{in\theta}$ (for any value of $n \in \mathbb{Z}$) doesn't change this fact. Because of this, and by virtue of Theorem 5.24, in calculating the Fourier coefficients of $[f]$, we may always replace the Lebesgue integrals which appear by Riemann integrals.

CASE 1: $n = 0$. Observe that

$$\alpha_0^{[f]} = \frac{1}{2\pi}\int_{[-\pi,\pi)} f = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\ d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\theta|\ d\theta = \frac{1}{2\pi}\pi^2 = \frac{\pi}{2}.$$

CASE 1: $n \neq 0$.

In this case,

$$\begin{aligned}\alpha_n^{[f]} &= \frac{1}{2\pi}\int_{[-\pi,\pi)} f\overline{\xi_n} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |\theta| e^{-in\theta}\ d\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{0} (-\theta) e^{-in\theta}\ d\theta + \frac{1}{2\pi}\int_{0}^{\pi} \theta\, e^{-in\theta}\ d\theta.\end{aligned}$$

These integrals are easily found using integration by parts, and we leave them as exercises for the reader. The answers are:

$$\alpha_n^{[f]} = \begin{cases} 0 & 0 \neq n \text{ an even integer} \\ \frac{-2}{n^2\pi} & 0 \neq n \text{ an odd integer.} \end{cases}$$

Note that $e^{-in\theta} + e^{in\theta} = 2\cos(n\theta)$ for all $n \geq 1$. Thus – taking into account that we only wish to sum over odd n's below –

$$\begin{aligned} s_N([f]) &= \frac{\pi}{2} + \sum_{1\leq 2n-1\leq N} \left(\frac{-2}{(2n-1)^2\pi}\right) 2\cos((2n-1)\theta) \\ &= \frac{\pi}{2} - \frac{4}{\pi} \sum_{1\leq 2n-1\leq N} \frac{1}{(2n-1)^2} \cos((2n-1)\theta). \end{aligned}$$

But

$$\begin{aligned} \left| \sum_{1\leq 2n-1\leq N} \frac{1}{(2n-1)^2} \cos((2n-1)\theta) \right| &\leq \sum_{1\leq 2n-1\leq N} \frac{1}{(2n-1)^2} \\ &\leq \sum_{n=1}^{N} \frac{1}{n^2}. \end{aligned}$$

But this last series converges, which implies that $(s_N([f]))_{n=1}^\infty$ converges in the $\|\cdot\|_\infty$-norm, as $([\mathcal{C}(\mathbb{T},\mathbb{C})], \|\cdot\|_\infty)$ is complete. In other words,

$$s_N(f) = \sum_{n=-N}^{N} \alpha_n^{[f]} \xi_n$$

converges uniformly in $(\mathcal{C}(\mathbb{T},\mathbb{C}), \|\cdot\|_{\sup})$ to a continuous function.

Since f is globally Lipschitz, by Theorem 11.21,

$$f(\theta) = |\theta| = \lim_{N\to\infty} s_N(f)(\theta)$$

for all $\theta \in [-\pi, \pi)$. That is,

$$f(\theta) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \cos((2n-1)\theta).$$

It is time to have fun:

(a) Let $\theta = 0$. Then

$$\begin{aligned} 0 = |0| &= \frac{\pi}{2} - \frac{4}{\pi} \left(\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \cos((2n-1)0) \right) \\ &= \frac{\pi}{2} - \frac{4}{\pi} \left(\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \right). \end{aligned}$$

Thus

$$\frac{\pi^2}{8} = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} = \frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots.$$

(b) Let $\theta = \frac{\pi}{2}$. Then

$$\frac{\pi}{2} = |\frac{\pi}{2}| = \frac{\pi}{2} - \frac{4}{\pi}\left(\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \cos((2n-1)\frac{\pi}{2})\right)$$
$$= \frac{\pi}{2} - \frac{4}{\pi}\left(\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \cdot 0\right).$$

Thus

$$\frac{\pi}{2} = \frac{\pi}{2},$$

a relatively well-known result.

(c) Let $\theta = \frac{\pi}{4}$. Now $\cos((2n-1)\frac{\pi}{4}) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } n = 1 \mod 4 \text{ or } n = 0 \mod 4 \\ \frac{-1}{\sqrt{2}} & \text{if } n = 2 \mod 4 \text{ or } n = 3 \mod 4 \end{cases}$.

Thus

$$\frac{\pi}{4} = |\frac{\pi}{4}| = \frac{\pi}{2} - \frac{4}{\pi}\left(\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \cos((2n-1)\frac{\pi}{4})\right)$$
$$= \frac{\pi}{2} - \frac{4}{\pi}\frac{1}{\sqrt{2}}\left(\frac{1}{1^2} - \frac{1}{3^2} + \frac{1}{5^2} - \frac{1}{7^2} + \frac{1}{9^2} - \frac{1}{11^2} + \cdots\right).$$

In other words,

$$\frac{\sqrt{2}\pi^2}{16} = \left(\frac{1}{1^2} - \frac{1}{3^2} + \frac{1}{5^2} - \frac{1}{7^2} + \frac{1}{9^2} - \frac{1}{11^2} + \cdots\right).$$

Clearly life does not get better than this. Alas.

Appendix to Section 11.

11.23. It is also instructive to look at the graph of the Féjer kernels of various orders. Two things worth noticing are that

- first, the amplitude of the function is increasing near 0; this is clear since each F_N is continuous, and $F_N(0) = N$, $N \geq 1$.
- For each $\delta > 0$, the functions are becoming uniformly close to zero when $\delta < |\theta| < \pi$.

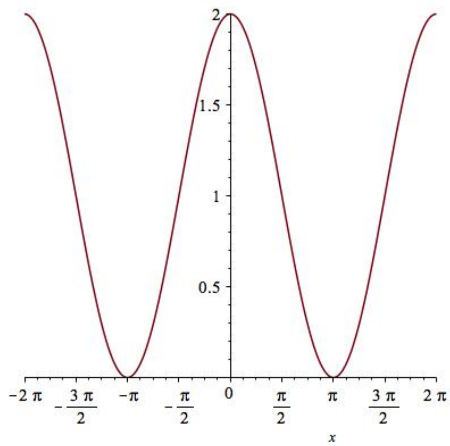


Figure 7. The graph of K_2.

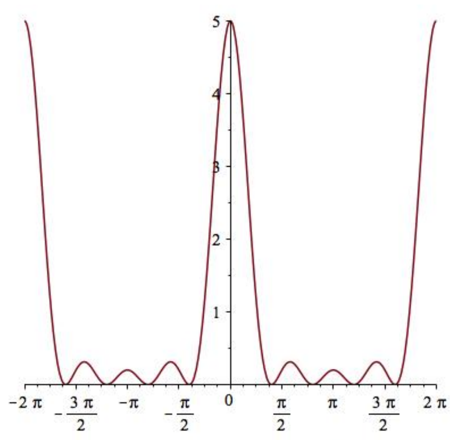


Figure 8. The graph of K_5.

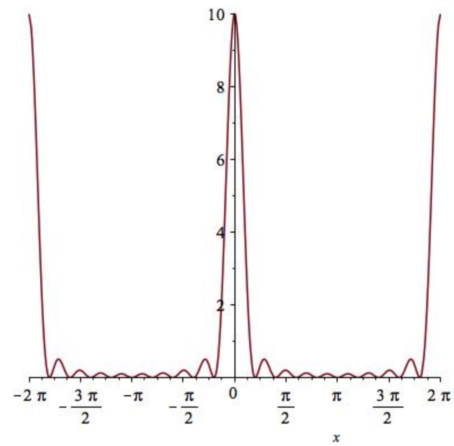


FIGURE 9. THE GRAPH OF K_{10}.

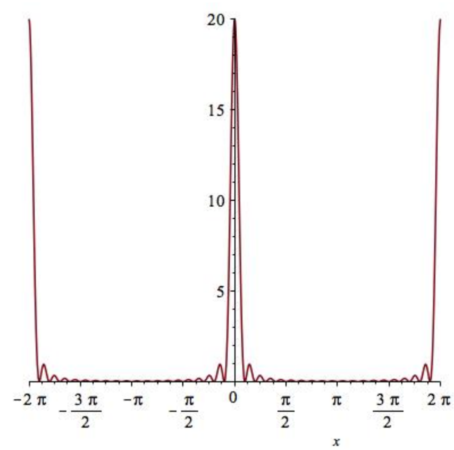


FIGURE 10. THE GRAPH OF K_{20}.

11.24. Remark. In fact, it can be shown that if $k : \mathbb{R} \to \mathbb{C}$ is *piecewise-continuous* on $[-\pi, \pi)$ and k is 2π-periodic, then for all $f \in \mathcal{L}_1(\mathbb{T}, \mathbb{C})$, $k \diamond f$ is continuous on $\mathbb{R}$.

Using this, one may extend the definition of a summability kernel to include piecewise-continuous functions, as opposed to continuous functions.

11.25. Examples.

(a) For $1 \le n \in \mathbb{N}$, let k_n be the 2π-periodic function on $\mathbb{R}$ whose restriction to the interval $[-\pi, \pi)$ coincides with

$$n\pi \chi_{[-\frac{1}{n}, \frac{1}{n}]}.$$

Then $(k_n)_{n=1}^{\infty}$ is a positive summability kernel. The details are left to the reader.

(b) For $1 \le n \in \mathbb{N}$, let k_n be the 2π-periodic function on $\mathbb{R}$ whose restriction to the interval $[-\pi, \pi)$ coincides with

$$2n\pi\chi_{[0,\frac{1}{n}]}.$$

Then $(k_n)_{n=1}^{\infty}$ is a positive summability kernel. Again, the details are left to the reader.

Exercises for Section 11.

Exercise 11.1.

Prove the claim of Proposition 11.3, namely: suppose that $\mathfrak{X}$ is a Banach space and $(x_n)_{n=0}^\infty$ is a sequence in $\mathfrak{X}$. Let $(\sigma_N)_{N=1}^\infty$ denote the sequence of Cesàro means of $(x_n)_{n=0}^\infty$.

Prove that if $x = \lim_{n\to\infty} x_n$ exists, then $x = \lim_{N\to\infty} \sigma_N$.

Exercise 11.2.

Prove that the sequences $(k_n)_n$ and $(r_n)_n$ listed in Examples 11.9 are indeed positive summability kernels.

12. Which sequences are sequences of Fourier coefficients?

I've learned about his illness. Let's hope it's nothing trivial.

Irving S. Cobb

12.1. Given $[f] \in L_1(\mathbb{T}, \mathbb{C})$, we have defined the Fourier series of $[f]$ to be

$$\sum_{n \in \mathbb{Z}} \alpha_n^{[f]} [\xi_n].$$

The Riemman-Lebesgue Lemma is the statement that

$$(\alpha_n^{[f]})_{n \in \mathbb{Z}} \in c_0(\mathbb{Z}, \mathbb{C}).$$

It is a natural question to ask, therefore, whether *every* sequence $(\beta_n)_{n \in \mathbb{Z}} \in c_0(\mathbb{Z}, \mathbb{C})$ is the sequence of coefficients of some $[f] \in L_1(\mathbb{T}, \mathbb{C})$. What *is* clear from our work on Hilbert spaces is that every $(\gamma_n)_{n \in \mathbb{Z}} \in \ell_2(\mathbb{Z}, \mathbb{C})$ is the set of Fourier coefficients of some $[f] \in L_2(\mathbb{T}, \mathbb{C})$, namely $[f] = \sum_{n \in \mathbb{Z}} \gamma_n [\xi_n]$.

Our approach to this problem will be via Operator Theory. Recall from Chapter 8 that we defined the map

$$\begin{array}{rccc} \Lambda: & (L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1) & \to & (c_0(\mathbb{Z}, \mathbb{C}), \|\cdot\|_\infty) \\ & [f] & \mapsto & (\alpha_n^{[f]})_{n \in \mathbb{Z}}. \end{array}$$

Since Lebesgue integration is linear, so is Λ. Also, as was shown in paragraph 8.8

$$|\alpha_n^{[f]}| \le \|[f]\|_1 \quad \text{for all } n \in \mathbb{Z},$$

so

$$\|\Lambda([f])\|_\infty = \sup\{|\alpha_n^{[f]}| : n \in \mathbb{Z}\} \le \|[f]\|_1.$$

This is precisely the statement that the operator Λ is bounded, with $\|\Lambda\| \le 1$.

By Corollary 11.12, if $[f], [g] \in L_1(\mathbb{T}, \mathbb{C})$ and $\Lambda([f]) = \Lambda([g])$, then $[f] = [g]$, and thus Λ is injective.

The question of whether or not every sequence in $c_0(\mathbb{Z}, \mathbb{C})$ is the sequence of Fourier coefficients of some element of $L_1(\mathbb{T}, \mathbb{C})$ is therefore the question of whether or not Λ is surjective.

The result we shall need from Functional Analysis is the *Inverse Mapping Theorem.* To get this result, we will first require a lemma, and some notation.

Given a Banach space $(\mathfrak{Z}, \|\cdot\|_{\mathfrak{Z}})$ and a real number $r > 0$, we denote the closed ball of radius r centred at the origin by

$$\mathfrak{Z}_r = \{z \in \mathfrak{Z} : \|z\|_{\mathfrak{Z}} \le r\}.$$

For $z_0 \in \mathfrak{Z}$ and $\varepsilon > 0$, we shall denote by $B^{\mathfrak{Z}}(z_0, \varepsilon) = \{z \in \mathfrak{Z} : \|z - z_0\| < \varepsilon\}$ the open ball of radius ε in $\mathfrak{Z}$, centred at z_0.

12.2. Lemma. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and suppose that $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. If $\mathfrak{Y}_1 \subseteq \overline{T\mathfrak{X}_m}$ for some $m \ge 1$, then $\mathfrak{Y}_1 \subseteq T\mathfrak{X}_{2m}$.*

Proof. First observe that $\mathfrak{Y}_1 \subseteq \overline{T\mathfrak{X}_m}$ implies that $\mathfrak{Y}_r \subseteq \overline{T\mathfrak{X}_{rm}}$ for all $r > 0$.

Choose $y \in \mathfrak{Y}_1$. Then there exists $x_1 \in \mathfrak{X}_m$ so that $\|y - Tx_1\| < 1/2$. Since $y - Tx_1 \in \mathfrak{Y}_{1/2} \subseteq \overline{T\mathfrak{X}_{m/2}}$, there exists $x_2 \in \mathfrak{X}_{m/2}$ so that $\|(y - Tx_1) - Tx_2\| < 1/4$. More generally, for each $n \ge 1$, we can find $x_n \in \mathfrak{X}_{m/2^{n-1}}$ so that

$$\|y - \sum_{j=1}^{n} Tx_j\| < \frac{1}{2^n}.$$

Since $\mathfrak{X}$ is complete and $\sum_{n=1}^{\infty} \|x_n\| \le \sum_{n=1}^{\infty} \frac{m}{2^{n-1}} = 2m$, we have $x = \sum_{n=1}^{\infty} x_n \in \mathfrak{X}_{2m}$. By the continuity of T,

$$Tx = T\left(\sum_{n=1}^{\infty} x_n\right) = \lim_{N\to\infty} T\left(\sum_{n=1}^{N} x_n\right) = y.$$

Thus $y \in T\mathfrak{X}_{2m}$; i.e. $\mathfrak{Y}_1 \subseteq T\mathfrak{X}_{2m}$.

□

12.3. Theorem. ***The Open Mapping Theorem.***

Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and suppose that $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is a surjection. Then T is an open map - i.e. if $G \subseteq \mathfrak{X}$ is open, then $TG \subseteq \mathfrak{Y}$ is open.

Proof. Since T is surjective, $\mathfrak{Y} = T\mathfrak{X} = \cup_{n=1}^{\infty} T\mathfrak{X}_n = \cup_{n=1}^{\infty} \overline{T\mathfrak{X}_n}$. Now $\mathfrak{Y}$ is a complete metric space, and so by the Baire Category Theorem 10.9 , there exists $m \ge 1$ so that the interior $\text{int}(\overline{T\mathfrak{X}_m}) \ne \varnothing$. As $T\mathfrak{X}_m$ is dense in $\overline{T\mathfrak{X}_m}$, we can choose $y \in \text{int}(\overline{T\mathfrak{X}_m}) \cap T\mathfrak{X}_m$.

Let $\delta > 0$ be such that $B^{\mathfrak{Y}}(y, \delta) = y + B^{\mathfrak{Y}}(0, \delta) \subseteq \text{int}(\overline{T\mathfrak{X}_m}) \subseteq (\overline{T\mathfrak{X}_m})$. Then $B^{\mathfrak{Y}}(0, \delta) \subseteq -y + \overline{T\mathfrak{X}_m} \subseteq T\mathfrak{X}_m + \overline{T\mathfrak{X}_m} \subseteq \overline{T\mathfrak{X}_{2m}}$. (This last step uses the linearity of T.)

Thus $\mathfrak{Y}_{\delta/2} \subseteq B^{\mathfrak{Y}}(0, \delta) \subseteq \overline{T\mathfrak{X}_{2m}}$. By Lemma 12.2 above,

$$\mathfrak{Y}_{\delta/2} \subseteq T\mathfrak{X}_{4m},$$

or equivalently,

$$T\mathfrak{X}_r \supset \mathfrak{Y}_{r\delta/8m}$$

for all $r > 0$.

Suppose that $G \subseteq \mathfrak{X}$ is open and that $y \in TG$, say $y = Tx$ for some $x \in G$. Since G is open, we can find $\varepsilon > 0$ so that $x + B^{\mathfrak{X}}(0, \varepsilon) \subseteq G$. Thus

$$\begin{aligned} TG &\supseteq Tx + T(B^{\mathfrak{X}}(0, \varepsilon)) \\ &\supseteq y + T\mathfrak{X}_{\varepsilon/2} \\ &\supseteq y + \mathfrak{Y}_{\varepsilon\delta/16m} \\ &\supseteq y + B^{\mathfrak{Y}}(0, \varepsilon\delta/16m) \\ &= B^{\mathfrak{Y}}(y, \varepsilon\delta/16m). \end{aligned}$$

Thus $y \in TG$ implies that $y \in \text{int}\, TG$, and so TG is open.

□

12.4. Corollary. ***The Inverse Mapping Theorem***
Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and suppose that $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is a bijection. Then T^{-1} is continuous, and so T is a homeomorphism.

Proof. That T^{-1} is linear is basic linear algebra.

If $G \subseteq \mathfrak{X}$ is open, then $(T^{-1})^{-1}(G) = TG$ is open in $\mathfrak{Y}$ by the Open Mapping Theorem above. Hence T^{-1} is continuous.

□

12.5. Theorem. *The map*

$$\begin{array}{rccc} \Lambda : & (L_1(\mathbb{T}, \mathbb{C}), \|\cdot\|_1) & \to & (c_0(\mathbb{Z}, \mathbb{C}), \|\cdot\|_\infty) \\ & [f] & \mapsto & (\alpha_n^{[f]})_{n \in \mathbb{Z}}. \end{array}$$

is not surjective.

Proof. As we saw in paragraph 12.1, Λ is continuous, linear and injective. If it were surjective, then by the Inverse Mapping Theorem,

$$\begin{array}{rccc} \Lambda^{-1} : & c_0(\mathbb{Z}, \mathbb{C}) & \to & L_1(\mathbb{T}, \mathbb{C}) \\ & (\alpha_n^{[f]})_{n \in \mathbb{Z}} & \mapsto & [f] \end{array}$$

would be continuous.

Let $D_N = \sum_{n=-N}^{N} \xi_n$ be the Dirichlet kernel of order N, and let $d_N := \Lambda([D_N])$, $N \geq 1$.

Then $d_N = (..., 0, 0, \ldots, 0, 1, 1, \ldots, 1, 1, 0, 0, \ldots)$, with the 1's appearing for the indices $-N \leq k \leq N$. Clearly $\|d_N\|_\infty = 1$, each d_N is finitely supported, but by Theorem 10.4,

$$\lim_{N \to \infty} \|\Lambda^{-1}(d_N)\|_1 = \lim_{N \to \infty} \|[D_N]\|_1 = \infty.$$

Thus Λ^{-1} is not continuous, and so Λ is not surjective.

That is, there exist sequences $(\beta_n)_{n \in \mathbb{Z}} \in c_0(\mathbb{Z}, \mathbb{C})$ which are not the Fourier coefficients of any element of $L_1(\mathbb{T}, \mathbb{C})$.

□

12.6. Of course, the fact that $[f] \in L_2(\mathbb{T}, \mathbb{C})$ if and only if $(\alpha_n^{[f]})_{n \in \mathbb{Z}} \in \ell_2(\mathbb{Z}, \mathbb{C})$ makes it tempting to conjecture that perhaps the range of the map Λ from Theorem 12.5 should be $\ell_1(\mathbb{Z}, \mathbb{C})$. Tempting, but alas, false.

The sequence $\beta_n = \begin{cases} \frac{1}{n} & \text{if } n \geq 1 \\ 0 & \text{if } n \leq 0 \end{cases}$ is clearly in $\ell_2(\mathbb{Z}, \mathbb{C})$, and thus $[f] := \sum_{n \in \mathbb{Z}} \beta_n [\xi_n]$ converges in $L_2(\mathbb{T}, \mathbb{C}) \subseteq L_1(\mathbb{T}, \mathbb{C})$. On the other hand,

$$\Lambda([f]) = (\beta_n)_{n \in \mathbb{Z}}$$

is definitely not in $\ell_1(\mathbb{Z}, \mathbb{C})$.

As stated in the book by Katznelson [**2**], p. 23,

> The only spaces, defined by conditions of size or smoothness of the functions, for which we obtain ... [a] complete characterisation, that is, a necessary and sufficient condition expressed in terms of order of magnitude, for a sequence $\{a_n\}$ to be the Fourier coefficients of a function in the space, are $L_2(\mathbb{T}, \mathbb{C})$ and its "derivatives". (Such as the space of absolutely continuous functions with derivatives in $L_2(\mathbb{T}, \mathbb{C})$.)

So - much like the enigma surrounding the Cadbury Caramilk bar, the mystery persists.

Appendix to Section 12.

Exercises for Section 12.

Exercise 12.1.

Keeping in mind that $\sin n\theta = \dfrac{\xi_n(\theta) - \xi_{-n}(\theta)}{2i}$, $\theta \in \mathbb{R}$, the series $\displaystyle\sum_{n=1}^{\infty} \frac{\sin n\theta}{n}$ can be expressed as the Fourier series of a function in $L_2(\mathbb{T}, \mathbb{C})$. Which function?

Bibliography

[1] I.N. Herstein. *Topics in Algebra, 2nd ed.* John Wiley and Sons, New York, Chichester, Brisbane, Toronto, Singapore, 1975.

[2] Y. Katznelson. *An introduction to harmonic analysis, Third edition.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2004.

[3] L.W. Marcoux. *An introduction to Functional Analysis.* Online notes – available at http://www.math.uwaterloo.ca/~lwmarcou/, 2012.

[4] J.C. Oxtoby. *Measure and Category: A survey of the analogies between topological and measure spaces*, volume 2 of *Graduate Texts in Mathematics.* Springer-Verlag, New York-Berlin, 1971.

[5] R.M. Solovay. A model of set-theory in which every set of reals is lebesgue measurable. *Ann. of Math.*, 92:1–56, 1970.

[6] G. Vitali. Sul problema della misura dei gruppi di punti di una retta. *Bologna, Tip. Gamberini e Parmeggiani*, 1905.

[7] R. Whitley. Projecting m onto c_0. *The Amer. Math. Monthly*, 73:285–286, 1966.

[8] Stephen Willard. *General Topology.* Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1970.

Index