Lecture Notes on F.Riesz's Approach to the Lebesgue Integral

Douglas S. Bridges
Department of Mathematics & Statistics,
University of Canterbury,
Private Bag 4800,
Christchurch, New Zealand.

January 12, 2006

Preface

In the following, I introduce the Lebesgue integral on the real line $\mathbb{R}$ using the method of F. Riesz. Working with increasing sequences of step functions whose integrals are uniformly bounded above, this method, which is essentially a special case of the Daniell approach to abstract integration, avoids the somewhat tedious technical detail about measures that is required in the standard measure-theoretic introductions to the Lebesgue integral, and thereby enables us rapidly to reach the key results about convergence of sequences and series of integrable functions.

The later sections of the notes contain material about the spaces $L_p(\mathbb{R})$ of p–power integrable functions on $\mathbb{R}$; a development of the Lebesgue double integral, including Fubini's theorem about the equivalence of double and repeated integrals; and a discussion of topics in advanced differentiation theory, such as Fubini's series theorem and the Lebesgue-integral form of the fundamental theorem of calculus.

Acknowledgement. These notes arose from material prepared by F.F. Bonsall and presented by J.C. Alexander to an honours course at the University of Edinburgh in 1967. I am grateful to those two gentlemen for inspiring my interest in Riesz's work on the Lebesgue integral.

1 Compactness

Many results in real-variable theory depend on two fundamental properties of the real line described in this introductory section.

By a **cover** of a subset S of $\mathbb{R}$ we mean a family $\mathcal{U}$ of subsets of $\mathbb{R}$ such that $S \subset \bigcup \mathcal{U}$; we then say that S is **covered** by $\mathcal{U}$ and that $\mathcal{U}$ **covers** S. If also each $U \in \mathcal{U}$ is an open subset of $\mathbb{R}$, we refer to $\mathcal{U}$ as an **open cover** of S (in $\mathbb{R}$). By a **subcover** of a cover $\mathcal{U}$ of S we mean a family $\mathcal{F} \subset \mathcal{U}$ that covers S; if also $\mathcal{F}$ is a finite family, then it is called a **finite subcover** of $\mathcal{U}$.

Although there exist shorter proofs of the next theorem (see the next set of exercises), the one we present is readily adapted to work in a more general context.

Theorem 1 ***The Heine–Borel–Lebesgue theorem.*** *Every open cover of a bounded closed interval I in $\mathbb{R}$ contains a finite subcover of I.*

Proof. Suppose there exists an open cover $\mathcal{U}$ of I that contains no finite subcover of I. Either the closed right half of I or the closed left half (or both) cannot be covered by a finite subfamily of $\mathcal{U}$: otherwise each half, and therefore I itself, would be covered by a finite subfamily. Let I_1 be a closed half of I that is not covered by a finite subfamily of $\mathcal{U}$. In turn, at least one closed half, say I_2, of I_1 cannot be covered by a finite subfamily of $\mathcal{U}$. Carrying on in this way, we construct a nested sequence $I \supset I_1 \supset I_2 \supset \cdots$ of closed subintervals of I such that for each n,

(a) $|I_n| = 2^{-n}|I|$ and

(b) no finite subfamily of $\mathcal{U}$ covers I_n.

By the nested intervals principle, there exists a point $\xi \in \bigcap_{n \geqslant 1} I_n$. Clearly $\xi \in I$, so there exists $U \in \mathcal{U}$ such that $\xi \in U$. Since U is open, there exists $r > 0$ such that if $|x - \xi| < r$, then $x \in U$. Using (a), we can find N such that if $x \in I_N$, then $|x - \xi| < r$ and therefore $x \in U$; thus $I_N \subset U$. This contradicts (b). ■

We follow tradition by referring to a bounded closed interval as a **compact interval.**

A real number a is called a **limit point** of a subset S of $\mathbb{R}$ if each neighbourhood of a intersects $S \backslash \{a\}$; or, equivalently, if for each $\varepsilon > 0$ there exists $x \in S$ with $0 < |x - a| < \varepsilon$. By a **limit point of a sequence** (a_n) we mean a limit point of the set $\{a_1, a_2, \ldots\}$ of terms of the sequence.

A nonempty subset A of $\mathbb{R}$ is said to have the **Bolzano–Weierstrasz property** if each infinite subset S of A has a limit point belonging to A.

Theorem 2 ***The Bolzano–Weierstrasz theorem.*** *Every compact interval in $\mathbb{R}$ has the Bolzano–Weierstrasz property.*

Proof. See Exercise (1.1, 2) below. ■

Exercises (1.1)

.1 Fill in the details of the following alternative proof of the Heine–Borel–Lebesgue theorem. Let $\mathcal{U}$ be an open cover of the compact interval $I = [a, b]$, and define

$$A = \{x \in I : [a, x] \text{ is covered by finitely many elements of } \mathcal{U}\}.$$

Then A is nonempty (it contains a) and is bounded above; let $\xi = \sup A$. Suppose that $\xi \neq b$, and derive a contradiction.

.2 Fill in the details of the following proof of the Bolzano–Weierstrasz theorem. Given a compact interval I, construct a nested sequence $I \supset I_1 \supset I_2 \supset \cdots$ of closed subintervals of I such that for each n,

(a) $|I_n| = 2^{-n} |I|$, and

(b) $S \cap I_n$ is an infinite set.

Let $\xi \in \bigcap_{n \geqslant 1} I_n$, and show that ξ is a limit point of S. (This is one of the commonest proofs of the Bolzano–Weierstrasz theorem in textbooks.)

.3 Here is a sketch of another proof of the Bolzano–Weierstrasz theorem for you to complete. Let I be a compact interval, and S an infinite subset of I; then the supremum of the set

$$A = \{x \in I : S \cap (-\infty, x) \text{ is finite or empty}\}$$

is a limit point of S in I.

.4 Let S be a subset of $\mathbb{R}$ with the Bolzano–Weierstrasz property, and let $(x_n)_{n \geqslant 1}$ be a sequence of points of S. Prove that there exists a subsequence of (x_n) that converges to a limit in S.

.5 Let S be a subset of $\mathbb{R}$ with the Bolzano–Weierstrasz property. Prove that S is closed and bounded. (For boundedness, use a proof by contradiction.)

.6 Show that the Bolzano–Weierstrasz theorem can be proved as a consequence of the Heine–Borel–Lebesgue theorem. (Let I be a compact interval in $\mathbb{R}$, and suppose that there exists an infinite subset S of I that has no limit point in I. First show that for each $s \in \overline{S}$ there exists $r_s > 0$ such that $S \cap (s - r_s,\, s + r_s) = \{s\}$.)

.7 Let f be a real-valued function defined on an interval I. We say that f is **uniformly continuous** on I if to each $\varepsilon > 0$ there corresponds $\delta > 0$ such that $|f(x) - f(x')| < \varepsilon$ whenever $x, x' \in I$ and $|x - x'| < \delta$. Show that a uniformly continuous function is continuous. Give an example of I and f such that f is continuous, but not uniformly continuous, on I.

.8 Use the Heine–Borel–Lebesgue theorem to prove the **uniform continuity theorem:** a continuous real-valued function f on a compact interval $I \subset \mathbb{R}$ is uniformly continuous. (For each $\varepsilon > 0$ and each $x \in I$, choose $\delta_x > 0$ such that if $x' \in I$ and $|x - x'| < 2\delta_x$, then $|f(x) - f(x')| < \varepsilon/2$. The intervals $(x - \delta_x, x + \delta_x)$ form an open cover of I.)

.9 Prove the uniform continuity theorem (see the previous exercise) using the Bolzano–Weierstrasz theorem. (If $f : I \to \mathbb{R}$ is not uniformly continuous, then there exists $\alpha > 0$ with the following property: for each $n \in \mathbb{N}^+$ there exist $x_n, y_n \in I$ such that $|x_n - y_n| < 1/n$ and $|f(x_n) - f(y_n)| \geqslant \alpha$.)

The proof of the following result about boundedness of real-valued functions illustrates well the application of the Heine–Borel–Lebesgue theorem.

Theorem 3 *A continuous real-valued function f on a compact interval I is bounded; moreover, f* ***attains its bounds*** *in the sense that there exist points ξ, η of I such that $f(\xi) = \inf f$ and $f(\eta) = \sup f$.*

Proof. For each $x \in I$ choose $\delta_x > 0$ such that if $x' \in I$ and $|x - x'| < \delta_x$, then $|f(x) - f(x')| < 1$. The intervals $(x - \delta_x, x + \delta_x)$, where $x \in I$, form an open cover of I. By Theorem 1, there exist finitely many points $x_1, \dots, x_N$ of I such that

$$I \subset \bigcup_{k=1}^{N} (x_k - \delta_{x_k}, x_k + \delta_{x_k})\,.$$

Let

$$c = 1 + \max\{|f(x_1)|, \dots, |f(x_N)|\}\,,$$

and consider any point $x \in I$. Choosing k such that $x \in (x_k - \delta_{x_k}, x_k + \delta_{x_k})$, we have

$$\begin{aligned} |f(x)| &\leqslant |f(x) - f(x_k)| + |f(x_k)| \\ &< 1 + |f(x_k)| \\ &\leqslant c, \end{aligned}$$

so f is bounded on I.

Now write

$$m = \inf f, \quad M = \sup f.$$

Suppose that $f(x) \neq M$, and therefore $f(x) < M$, for all $x \in I$. Then $x \rightsquigarrow 1/(M - f(x))$ is a continuous mapping of I into $\mathbb{R}^+$, and so, by the first part of this proof, has a supremum $G > 0$. For each $x \in I$ we then have $M - f(x) \geqslant 1/G$ and therefore $f(x) \leqslant M - 1/G$. This contradicts our choice of M as the supremum of f. ■

Exercises (1.2)

.1 Prove both parts of Theorem 3 using the Bolzano–Weierstrasz theorem and contradiction arguments.

.2 Let f be a continuous function on $\mathbb{R}$ such that $f(x) \to \infty$ as $x \to \pm\infty$. Prove that there exists $\xi \in \mathbb{R}$ such that $f(x) \geqslant f(\xi)$ for all $x \in \mathbb{R}$.

.3 Let f be a continuous function on $\mathbb{R}$ such that $f(x) \to 0$ as $x \to \pm\infty$. Prove that f is both bounded and uniformly continuous.

Theorems 1, 2, and 3 apply also to sets of the form $I \times J \subset \mathbb{R}^2$ where I, J are compact intervals in $\mathbb{R}$: every open cover of $I \times J$ contains a finite subcover; $I \times J$ has the Bolzano–Weierstrasz property; and every continuous function $f : I \times J \to \mathbb{R}$ is bounded and attains its bounds.

2 The Lebesgue Integral

By a **step function** we mean a function $f : \mathbb{R} \to \mathbb{R}$ with the following property: there exist finitely many points $x_0 < x_1 < \cdots < x_n$ and finitely many real numbers $c_0, \ldots, c_{n-1}$ such that

- ▷ $f(x) = 0$ for all $x < x_0$ and for all $x > x_n$,
- ▷ $f(x) = c_k$ for each k and all $x \in (x_k, x_{k+1})$.

Note that the value of f at x_k is not constrained in any way. The **integral** of this step function is defined to be

$$\int f = \sum_{k=0}^{n-1} c_k \, |I_k| \, ,$$

where $|I_k| = x_{k+1} - x_k$ denotes the length of the interval (x_k, x_{k+1}).

Proposition 4 *Let f and g be two step functions such that $f(x) = g(x)$ for all but finitely many points x. Then $\int f = \int g$.*

The proof is left as an exercise. It follows from this proposition that the integral of a step function is well defined, in that its value does not depend on the choice of the finitely many points x_k that are the endpoints of intervals on which the function is constant.

Proposition 5 *Let f and g be step functions, and let $\alpha, \beta \in \mathbb{R}$. Then $\alpha f + \beta g$ is a step function, and the integral is a linear function:*

$$\int (\alpha f + \beta g) = \alpha \int f + \beta \int g.$$

Moreover, fg, $\max\{f,g\}$, $\min\{f,g\}$,

$$f^{+} = \max\{f, 0\},$$

and

$$f^{-} = \max\{-f, 0\}$$

are step functions. Finally, if $f(x) \leqslant g(x)$ *for all but finitely many (possibly no) points* x*, then* $\int f \leqslant \int g$.

Exercises (2.1)

.1 Prove Proposition 4.

.2 Prove Proposition 5.

A sequence $(f_n)_{n\geqslant 1}$ of step functions is called a **Riesz sequence** if

- ▷ $f_n(x) \leqslant f_{n+1}(x)$ for each n and all $x \in \mathbb{R}$, and
- ▷ the sequence $\left(\int f_n\right)_{n\geqslant 1}$ is bounded above.

Since the sequence of integrals $\int f_n$ is then both increasing and bounded above, it converges to a limit $l \in \mathbb{R}$. For each x the sequence $(f_n(x))_{n\geqslant 1}$ of real numbers is also increasing, but it may diverge to infinity. However, the convergence of the sequence of integrals suggests that the set of points x at which $(f_n(x))$ diverges to infinity is "small" in some sense. A subset A of $\mathbb{R}$ is called a **null set** if there exists a Riesz sequence (f_n) of step functions such that $f_n(x) \uparrow \infty$ for each $x \in A$. It follows immediately that every subset of a null set is also a null set. We say that a property $P(x)$ of real numbers holds **almost everywhere** if there exists a null set A such that $P(x)$ holds for all $x \in \mathbb{R}\backslash A$; we then write

$$P(x) \quad \text{a.e.}$$

If S is a subset of $\mathbb{R}$, A is a null set, and $P(x)$ is a property that holds for all x in $S\backslash A$, then we say that $P(x)$ holds **almost everywhere in** S, and we write

$$P(x) \quad \text{a.e. in } S.$$

By the definition of "null set" a Riesz sequence of step functions converges almost everywhere.

Lemma 6 *Let* A *be a null set. Then for each* $\varepsilon > 0$ *there exists a Riesz sequence* (f_n) *of nonnegative step functions such that* $f_n(x) \uparrow \infty$ *for each* $x \in A$*, and such that* $0 \leqslant \int f_n \leqslant \varepsilon$ *for each* n.

Proof. Choose a Riesz sequence (g_n) of step functions and $M > 0$ such that

(i) $g_n(x) \uparrow \infty$ for all $x \in A$ and

(ii) $\int g_n \leqslant M$ for all n.

Then $(g_n - g_1)_{n\geqslant 1}$ is a Riesz sequence of nonnegative step functions such that $g_n(x) - g_1(x) \uparrow \infty$ for each $x \in A$, and such that $\int (g_n - g_1) \leqslant M + \int |g_1|$ for each n. It remains to take

$$f_n = \frac{\varepsilon}{M + \int |g_1|}\,(g_n - g_1)\,.$$

■

How small is a null set?

Proposition 7 *A countable union of null sets is a null set.*

Proof. Let (A_n) be a sequence of null sets. By the preceding lemma, for each n there exists a Riesz sequence $\left(f_k^{(n)}\right)_{k\geqslant 1}$ of step functions that diverges to ∞ on A_n and that satisfies $\int f_k^{(n)} \leqslant 2^{-n}$ for all n and k. Define step functions

$$f_k = f_k^{(1)} + f_k^{(2)} + \cdots + f_k^{(k)} \quad (k \geqslant 1)\,.$$

Then

$$\begin{aligned} f_k(x) &= f_k^{(1)}(x) + \cdots + f_k^{(k)}(x) \\ &\leqslant f_{k+1}^{(1)}(x) + \cdots + f_{k+1}^{(k)}(x) + f_{k+1}^{(k+1)}(x) = f_{k+1}(x) \end{aligned}$$

and

$$\int f_k \leqslant 2^{-1} + 2^{-2} + \cdots + 2^{-k} < 1.$$

So $(f_k)_{k\geqslant 1}$ is a Riesz sequence of step functions. Finally, if $x \in A_n$, then for each $k \geqslant n$,

$$f_k(x) \geqslant f_k^{(n)}(x) \to \infty \text{ as } k \to \infty.$$

Hence $f_k(x) \to \infty$ for all $x \in \bigcup_{n\geqslant 1} A_n$. ■

Proposition 8 *Let (f_n) be a sequence of step functions that vanish outside a proper compact interval I and that decrease to 0 pointwise on I. Then $\int f_n \to 0$ as $n \to \infty$.*

Proof. For each n there is a finite (possibly empty) set D_n of points where f_n is not continuous. Then

$$D = \bigcup_{n\geqslant 1} D_n$$

is countable, say

$$D_n = \{x_1, x_2, x_3, \ldots\}\,.$$

We may assume that the endpoints of I belong to D. Choose $M>0$ such that $|f_1(x)| \leqslant M$ for all $x \in \mathbb{R}$. Given $\varepsilon>0$, for each n let J_n be an open interval of length $\leqslant \left(2^{n+1}M\right)^{-1}\varepsilon$ that contains x_n. Then $D \subset \bigcup_{n\geqslant 1} J_n$ and $\sum_{n=1}^{\infty} |J_n| \leqslant \varepsilon/M$.

Next, given $x \in I\backslash D$, choose a positive integer ν_x such that $f_n(x) \leqslant \varepsilon/2\,|I|$ for all $n \geqslant \nu_x$. Since $x \notin D$, f_{ν_x} is continuous at x and hence, being a step function, is constant in some open subinterval V_x of (the interior of) I containing x. Thus for all $t \in V_x$ and all $n \geqslant \nu_x$,

$$f_n(t) \leqslant f_{\nu_x}(t) \leqslant \frac{\varepsilon}{2\,|I|}.$$

The open intervals J_n $(n \geqslant 1)$ and V_x $(x \in I\backslash D)$ form an open cover of I, from which, by Theorem 1, we can extract a finite subcover, consisting, say, of the intervals J_n $(1 \leqslant n \leqslant N)$ and V_{ξ_k} $(1 \leqslant k \leqslant m)$. Define step functions g_1, g_2 as follows:

$$g_1(x) = \begin{cases} M & \text{if } x \in \bigcup_{n=1}^{N} J_n \\ 0 & \text{otherwise,} \end{cases}$$

$$g_2(x) = \begin{cases} \varepsilon/2\,|I| & \text{if } x \in \bigcup_{k=1}^{m} V_{\xi_k} \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\int g_1 \leqslant M\sum_{n=1}^{N} |J_n| \leqslant M\sum_{n=1}^{\infty} |J_n| \leqslant M\sum_{n=1}^{\infty} \left(2^n M\right)^{-1}\varepsilon = \frac{\varepsilon}{2}$$

and, as each V_{ξ_k} is in the interior of I,

$$\int g_2 \leqslant \frac{\varepsilon}{2\,|I|}\left|\bigcup_{k=1}^{m} V_{\xi_k}\right| \leqslant \frac{\varepsilon}{2}.$$

Let

$$\nu = \max\left\{\nu_{\xi_1}, \ldots, \nu_{\xi_m}\right\}.$$

Then for each $n \geqslant \nu$, since $f_n(x) \leqslant g_1(x) + g_2(x)$ for each $x \in I$, we have

$$\int f_n \leqslant \int g_1 + \int g_2 \leqslant \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the desired conclusion follows. ∎

Proposition 9 *Let (f_n) be a decreasing sequence of step functions such that* $\lim_{n\to\infty} f_n(x) = 0$ *a.e. Then* $\lim_{n\to\infty}\int f_n = 0$.

Proof. Choose a null set A such that $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in \mathbb{R}\backslash A$. Given $\varepsilon > 0$, use Lemma 6 to find a Riesz sequence (g_n) of step functions such that $g_n \geqslant 0$ and $0 \leqslant \int g_n \leqslant \varepsilon$ for each n, and such that $g_n(x) \uparrow \infty$ for each $x \in A$. Consider the step functions

$$h_n = (f_n - g_n)^+ .$$

For each n,

$$(f_{n+1} - g_{n+1}) - (f_n - g_n) = (f_{n+1} - f_n) + (g_n - g_{n+1}) \leqslant 0,$$

and so $0 \leqslant h_{n+1} \leqslant h_n$. Also, since $g_n \geqslant 0$, for each $x \in \mathbb{R}\backslash A$ we have

$$0 \leqslant \lim_{n\to\infty} h_n(x) \leqslant \lim_{n\to\infty} f_n(x) = 0$$

and therefore $h_n(x) \to 0$. On the other hand, for each $x \in A$ we have

$$f_n(x) - g_n(x) \leqslant f_1(x) - g_n(x) < 0$$

for all sufficiently large n; whence $h_n(x) = 0$ for all sufficiently large n. It now follows that $h_n(x) \to 0$ for all $x \in \mathbb{R}$. We can now apply Proposition 8 to show that $\int h_n \to 0$. Since

$$\int f_n = \int (f_n - g_n) + \int g_n \leqslant \int h_n + \varepsilon,$$

it follows that $\lim_{n\to\infty} \int f_n \leqslant \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we conclude that $\lim_{n\to\infty} \int f_n = 0$. ■

Let $E^{\uparrow}$ denote the set of all functions that are limits almost everywhere of some Riesz sequence (f_n) of step functions. If $f \in E^{\uparrow}$, and (f_n) is a Riesz sequence of step functions such that $f(x) = \lim_{n\to\infty} f_n(x)$ a.e., then we define the **(Lebesgue) integral** of f to be

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

In that case, if $g(x) = f(x)$ a.e., then, as the union of two null sets is a null set, $g(x) = \lim_{n\to\infty} f_n(x)$ a.e.; so $g \in E^{\uparrow}$ and $\int g = \lim_{n\to\infty} \int f_n = \int f$.

Exercises (2.2)

.1 Prove that (i) any nonempty finite set and (ii) the set of rational numbers is a null set. Is the empty subset of $\mathbb{R}$ a null set (explain your answer)?

.2 Let (f_n) and (g_n) be Riesz sequences of step functions such that

$$\lim_{n\to\infty} f_n(x) \geqslant \lim_{n\to\infty} g_n(x)$$

almost everywhere. Prove that

$$\lim_{n\to\infty} \int f_n \geqslant \lim_{n\to\infty} \int g_n.$$

(Consider the step functions $\phi_{m,n} = (g_m - f_n)^+$. Show that for each m, $\int \phi_{m,n} \to 0$ as $n \to \infty$.)

.3 Prove that the foregoing definition of the Lebesgue integral of an element of $E^\uparrow$ is a good one—in other words, that

(i) if (g_n) is another Riesz sequence of step functions converging to f almost everywhere, then

$$\lim_{n\to\infty} \int g_n = \lim_{n\to\infty} \int f_n$$

and

(ii) if f is a step function, then its Lebesgue integral coincides with the integral as we originally defined it.

(For (i) use the preceding exercise.)

.4 Prove that the function defined by $f(x) = 1$ for all $x \in \mathbb{R}$ is not in $E^\uparrow$. Hence prove that the complement $\mathbb{R}\backslash A$ of a null set A is not a null set.

.5 Let $f, g \in E^\uparrow$ and let $\lambda \geqslant 0$. Prove that $f+g$, λf, $\max\{f, g\}$ and $\min\{f, g\}$ belong to $E^\uparrow$, and that $\int (\lambda f + g) = \lambda \int f + \int g$.

Now let

$$L = \left\{ f_1 - f_2 : f_1, f_2 \in E^\uparrow \right\}.$$

The elements of L are called **(Lebesgue) summable** functions. If $f_1, f_2 \in E^\uparrow$ and $f = f_1 - f_2$, define the (**Lebesgue) integral** of f to be

$$\int f = \int f_1 - \int f_2.$$

Exercises (2.3)

.1 Prove that the foregoing is a good definition of the integral of an element f of L.

.2 Let $f, g \in L$ and $\alpha, \beta \in \mathbb{R}$. Prove the following.

(i) $\alpha f + \beta g \in L$ and $\int (\alpha f + \beta g) = \alpha \int f + \beta \int g$.

(ii) If $f(x) \geqslant 0$ a.e., then $\int f \geqslant 0$.

.3 Let f, g be summable functions. Prove that $\max\{f, g\}$, $\min\{f, g\}$, and $|f|$ are summable, and that $\left|\int f\right| \leqslant \int |f|$. (First consider the case where f, g are both nonnegative.)

.4 Prove that if f is summable, then the functions

$$\min\{f, n\}, \max\{\min\{f, n\}, -n\}, \text{ and } \min\left\{|f|, \frac{1}{n}\right\}$$

are summable for each positive integer n.

.5 Prove that the Lebesgue integral is **translation invariant**: if $f \in L$ and $y \in \mathbb{R}$, then the function f_y defined by $f_y(x) = f(x+y)$ is summable, and $\int f_y = \int f$. (Reduce to the case where $f \in E^{\uparrow}$.)

.6 Prove that if f is summable and α is a nonzero real number, then $h(x) = f(\alpha x)$ defines a summable function and $|\alpha| \int h = \int f$.

We say that $f \in L$ is **summable over the subset** A **of** $\mathbb{R}$ if the function $f\chi_A$ belongs to L, where $\chi_A : \mathbb{R} \to \{0, 1\}$ is the **characteristic function**, or **indicator**, or A :

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases}$$

We then write

$$\int_A f = \int f\chi_A.$$

If $A = [a, b]$ is a compact interval, we write

$$\int_a^b f = \int_{[a,b]} f;$$

and if, for example, $A = (-\infty, b]$, we write

$$\int_{-\infty}^b f = \int_{(-\infty,b]} f.$$

Note that

$$\begin{aligned} \chi_{A\cup B} &= \max\{\chi_A, \chi_B\} \text{ and} \\ \chi_{A\cap B} &= \min\{\chi_A, \chi_B\}. \end{aligned}$$

For example, we have

$$\begin{aligned} \chi_{A\cap B}(x) = 1 &\Leftrightarrow x \in A \cap B \\ &\Leftrightarrow x \in A \text{ and } x \in B \\ &\Leftrightarrow \chi_A(x) = 1 \text{ and } \chi_B(x) = 1 \\ &\Leftrightarrow \min\{\chi_A(x), \chi_B(x)\} = 1. \end{aligned}$$

Proposition 10 *If f is summable over the subsets A, B of $\mathbb{R}$, then it is summable over $A \cup B$ and $A \cap B$.*

Proof. First consider the case where $f \geqslant 0$. We have

$$f\chi_{A\cup B} = \max\{f\chi_A, f\chi_B\} \text{ and}$$
$$f\chi_{A\cap B} = \min\{f\chi_A, f\chi_B\},$$

both of which are summable by Exercise (2.3, 3). In the general case, set $f = f^+ - f^-$. Then $f^+ \in L$, by Exercise (2.3, 3). Since $f^+\chi_A = \max\{f\chi_A, 0\}$, the same exercise shows that f^+ is summable over A; similarly, f^+ is summable over B. It follows from the first part of this proof that f^+ is summable over both $A \cup B$ and $A \cap B$. Similarly, f^- is summable over $A \cup B$ and $A \cap B$. Hence

$$f\chi_{A\cup B} = f^+\chi_{A\cup B} - f^-\chi_{A\cup B}$$

is summable, as, likewise, is $f\chi_{A\cap B}$. ■

Exercises (2.4)

.1 Let f be summable over A and B. Is f summable over $A \backslash B$?

.2 Prove that if $f \in L$, then f is summable over any bounded interval I in $\mathbb{R}$. (First consider the case where f is a step function.)

Can we extend the class of summable functions by a method similar to the one that led us from the step functions to the set L? To consider this question properly, we call a sequence (f_n) in L a **Riesz sequence of summable functions** if $f_1 \leqslant f_2 \leqslant \cdots$ and the sequence $\left(\int f_n\right)_{n\geqslant 1}$ is bounded above.

Proposition 11 *Let (f_n) be a sequence of elements of $E^\uparrow$ such that $f_n(x) \geqslant 0$ a.e. and $\sum_{n=1}^{\infty} \int f_n$ converges in $\mathbb{R}$. Then there exists $f \in E^\uparrow$ such that $\sum_{n=1}^{\infty} f_n(x) = f(x)$ a.e. and $\int f = \sum_{n=1}^{\infty} \int f_n$.*

Proof. For each n choose[1] a Riesz sequence $\left(f_k^{(n)}\right)_{k\geqslant 1}$ of step functions and a null set A_n such that

(i) $\lim_{k\to\infty} f_k^{(n)}(x) = f_n(x)$ for all $x \in \mathbb{R}\backslash A_n$,

(ii) $\lim_{k\to\infty} \int f_k^{(n)} = \int f_n$, and

(iii) $f_k^{(n)} \geqslant 0$ for each k.

[1] This requires a little thought, though not much. Let $(g_n)_{n\geqslant 1}$ be a Riesz sequence of step functions converging almost everywhere to f, and take $f_n = \max\{0, g_n\}$.

Define the step functions

$$g_i = f_i^{(1)} + f_i^{(2)} + \cdots + f_i^{(i)} \quad (i \geqslant 1)\,.$$

Then (cf. the proof of Proposition 7) $g_i \leqslant g_{i+1}$. Also, since $f_i^{(n)} \uparrow f_n$ as $i \to \infty$,

$$\int g_i \leqslant \int f_1 + \int f_2 + \cdots + \int f_i \leqslant \sum_{n=1}^{\infty} \int f_n.$$

Hence $(g_i)_{i \geqslant 1}$ is a Riesz sequence of step functions, and so there exist $f \in E^{\uparrow}$ and a null set A_0 such that $g_i(x) \uparrow f(x)$ for all $x \in \mathbb{R}\backslash A_0$ and $\int g_i \uparrow \int f$ as $i \to \infty$. Let

$$A = \bigcup_{n=0}^{\infty} A_n,$$

which, being a countable union of null sets, is a null set. For each $x \in \mathbb{R}\backslash A$,

$$\begin{aligned} f(x) &\geqslant g_i(x) \\ &= f_i^{(1)}(x) + f_i^{(2)}(x) + \cdots + f_i^{(i)}(x) \\ &\geqslant f_i^{(1)}(x) + f_i^{(2)}(x) + \cdots + f_i^{(n)}(x) \quad (n \leqslant i)\,. \end{aligned}$$

Keeping n fixed and letting $i \to \infty$, we obtain

$$f(x) \geqslant f_1(x) + \cdots + f_n(x).$$

So the series $\sum\limits_{n=1}^{\infty} f_n(x)$ has partial sums bounded above by $f(x)$.Since

$$f_n(x) = \lim_{i\to\infty} f_i^{(n)}(x) \geqslant 0 \quad (x \in \mathbb{R}\backslash A)\,,$$

that series converges to a sum at most $f(x)$ for each $x \in \mathbb{R}\backslash A$. For each such x,

$$g_n(x) \leqslant f_1(x) + \cdots + f_n(x) \leqslant f(x) \quad (n \geqslant 1)\,,$$

so

$$f(x) = \lim_{n\to\infty} g_n(x) \leqslant \sum_{n=1}^{\infty} f_n(x) \leqslant f(x)$$

and therefore

$$f(x) = \sum_{n=1}^{\infty} f_n(x).$$

Finally,

$$\int g_i \leqslant \sum_{n=1}^{i} \int f_n \leqslant \int f,$$

so, letting $i \to \infty$, we have $\int f = \sum\limits_{n=1}^{\infty} \int f_n$. ∎

Before extending this result to the set L, we need a lemma.

Lemma 12 *Let $f \in L$, and let $f(x) \geqslant 0$ a.e. Then for each $\varepsilon > 0$ there exist $g, h \in E^{\uparrow}$ such that $f = g - h$, $g(x) \geqslant 0$ a.e, $h(x) \geqslant 0$ a.e., and $\int h < \varepsilon$.*

Proof. Choose $g^{(1)}, h^{(1)}$ in $E^{\uparrow}$ such that $f = g^{(1)} - h^{(1)}$. There exists a Riesz sequence (h_n) of step functions such that

$$\begin{aligned} h^{(1)}(x) &= \lim_{n\to\infty} h_n(x) \text{ a.e. and} \\ \int h^{(1)} &= \lim_{n\to\infty} \int h_n. \end{aligned}$$

Given $\varepsilon > 0$, choose N such that $\int h^{(1)} - \int h_N < \varepsilon$. Take $h = h^{(1)} - h_N$ and $g = g^{(1)} - h_N$. Since h_N is a step function, so is $-h_N$, and therefore both g and h belong to $E^{\uparrow}$. Also, $\int h < \varepsilon$, $h(x) \geqslant 0$ a.e., and $g(x) = f(x) + h(x) \geqslant 0$ a.e. Finally, $f = g - h$ a.e. ■

Theorem 13 ***Lebesgue's series theorem (first form):*** *Let (f_n) be a sequence of summable functions such that $f_n(x) \geqslant 0$ a.e. for each n, and $\sum_{n=1}^{\infty} \int f_n$ converges in $\mathbb{R}$. Then there exists $f \in L$ such that $\sum_{n=1}^{\infty} f_n(x)$ converges* a.e. *to $f(x)$ and $\sum_{n=1}^{\infty} \int f_n = \int f$.*

Proof. Using Lemma 12, choose sequences $(\phi_n), (\psi_n)$ in $E^{\uparrow}$ such that for each n, $f_n = \phi_n - \psi_n$, $\phi_n(x) \geqslant 0$ a.e., $\psi_n(x) \geqslant 0$ a.e., and $0 \leqslant \int \psi_n \leqslant 2^{-n}$. Then the series $\sum_{n=1}^{\infty} \int \psi_n$ converges by comparison with $\sum_{n=1}^{\infty} 2^{-n}$. Hence, by Proposition 11, there exists $\psi \in E^{\uparrow}$ such that $\sum_{n=1}^{\infty} \psi_n(x) = \psi(x)$ a.e. and $\sum_{n=1}^{\infty} \int \psi_n = \int \psi$. Since $\phi_n = f_n + \psi_n$, $\int \phi_n = \int f_n + \int \psi_n$ and the convergence of the series $\sum_{n=1}^{\infty} \int f_n$, $\sum_{n=1}^{\infty} \int \psi_n$ entails that of $\sum_{n=1}^{\infty} \int \phi_n$. Applying Proposition 11 once more, we obtain $\phi \in E^{\uparrow}$ such that $\sum_{n=1}^{\infty} \phi_n(x) = \phi(x)$ a.e. and $\sum_{n=1}^{\infty} \int \phi_n = \int \phi$. Since the union of two null sets is a null set, it remains to take $f = \phi - \psi$. ■

Exercises (2.5)

.1 Prove Lebesgue's series theorem (second form): If (f_n) is a sequence of summable functions such that $\sum_{n=1}^{\infty} \int |f_n|$ converges, then there exists $f \in L$ such that $\sum_{n=1}^{\infty} f_n(x) = f(x)$ a.e. and $\sum_{n=1}^{\infty} \int f_n = \int f$.

.2 Let f be summable. Prove that $\int |f| = 0$ if and only if $f(x) = 0$ a.e. (*Hint* for "only if": take $f_n = |f|$ in Theorem 13.)

.3 Prove that a null set cannot contain a proper interval.

.4 Let (A_n) be a sequence of subsets of $\mathbb{R}$, and f a function that is summable over each A_n, such that $\sum\limits_{n=1}^{\infty} \int_{A_n} |f|$ converges. Prove that

(i) f is summable over $A = \bigcup\limits_{n \geqslant 1} A_n$, and $\int_A |f| \leqslant \sum\limits_{n=1}^{\infty} \int_{A_n} |f|$;

(ii) if also the sets A_n are pairwise disjoint, then $\int_A f = \sum\limits_{n=1}^{\infty} \int_{A_n} f$.

We are now in a position to show that the extension process that took us from the set of step functions to $E^{\uparrow}$ and then L cannot be applied, beginning with L, to extend the class of summable functions.

Theorem 14 ***Beppo Levi's theorem:*** *Let (f_n) be a Riesz sequence of summable functions. Then there exists a summable function f such that*

- ▷ $\lim_{n\to\infty} f_n(x) = f(x)$ *a.e. and*
- ▷ $\lim_{n\to\infty} \int f_n = \int f$.

Proof. Choose M such that

$$M \geqslant \int f_n - \int f_1 \quad (n \geqslant 1)\,.$$

Define the summable functions

$$g_n = f_{n+1} - f_n \quad (n \geqslant 1)\,.$$

Then $g_n(x) \geqslant 0$ a.e. and $\sum\limits_{n=1}^{k} g_n = f_{k+1} - f_1$; so

$$\sum_{n=1}^{k} \int g_n = \int f_{k+1} - \int f_1 \leqslant M \quad (k \geqslant 1).$$

Thus the series $\sum\limits_{n=1}^{\infty} \int g_n$ of nonnegative terms has bounded partial sums and so is convergent. Applying Theorem 13, we obtain $g \in L$ such that $\sum\limits_{n=1}^{\infty} g_n(x) = g(x)$ a.e. and $\sum\limits_{n=1}^{\infty} \int g_n = \int g$. Now let $f = g + f_1 \in L$. Then

$$\lim_{n\to\infty} f_n(x) = f_1(x) + \sum_{n=1}^{\infty} g_n(x) = f_1(x) + g(x) = f(x) \quad \text{a.e.}$$

and

$$\lim_{n\to\infty}\int f_n = \int f_1 + \sum_{n=1}^{\infty}\int g_n = \int f_1 + \int g = \int (f_1+g) = \int f,$$

as we required. ■

Corollary 15 *Let (f_n) be a sequence of summable functions such that $f_1 \geqslant f_2 \geqslant \cdots \geqslant 0$ a.e. Then there exists a summable function f such that $f_n(x) \to f(x)$ a.e. and $\int f_n \to \int f$ as $n \to \infty$.*

Proof. Apply Beppo Levi's theorem to the sequence $(-f_n)_{n\geqslant 1}$ of summable functions. ■

Exercises (2.6)

.1 Define $f(x) = \mathrm{e}^{-\alpha x}$, where α is a positive constant. Prove that f is summable over $[0,\infty)$, and calculate $\int f$. (You may assume something that we will prove in Section 4: namely, that if $f : [a,b] \to \mathbb{R}$ is Riemann integrable, then it is summable and the Lebesgue integral of f equals its Riemann integral.)

.2 Prove that the series $\sum\limits_{n=1}^{\infty} \mathrm{e}^{-n^2 x}$ converges for each $x > 0$. Define

$$f(x) = \begin{cases} \sum\limits_{n=1}^{\infty} \mathrm{e}^{-n^2 x} & \text{if } x > 0 \\ \\ 0 & \text{if } x \leqslant 0. \end{cases}$$

Prove that f is summable and that $\int f = \sum\limits_{n=1}^{\infty} 1/n^2$.

.3 Let f be a summable function, and J a bounded interval. Use Beppo Levi's theorem to prove that f is summable over J. (Consider the sequence $(\min\{f, g_n\})_{n\geqslant 1}$, where $g_n(x) = n$ if $x \in J$, and $g_n(x) = 0$ otherwise.) Extend this result to an unbounded interval J. (First take $f \geqslant 0$. Consider the sequence (f_n), where $f_n(x) = f(x)$ if $x \in J \cap [-n,n]$, and $f_n(x) = 0$ otherwise.)

.4 Let f be summable, and $\varepsilon > 0$. Show that there exists a bounded interval J such that $\int_{\mathbb{R}\setminus J} |f| < \varepsilon$. (Consider $|f|\,\chi_n$, where χ_n is the characteristic function of $[-n,n]$.)

.5 Prove **Fatou's lemma**: If (f_n) is a sequence of nonnegative summable functions that converges almost everywhere to a function f, and if the sequence $\left(\int f_n\right)$ is bounded above, then f is summable and

$$\int f \leqslant \liminf \int f_n.$$

(Apply Beppo Levi's theorem to the functions $g_n = \inf_{k\geqslant n} f_k$.)

.6 Let $(f_n)_{n=0}^{\infty}$ be a sequence of nonnegative summable functions such that $f_n(x) \leqslant f_0(x)$ a.e. for each n. Prove that $\sup_{k \geqslant n} f_k$ is summable for each n, that

$$\limsup f_n = \lim_{k \to \infty} \sup_{n \geqslant k} f_n$$

is summable, and that

$$\limsup_{n \geqslant 1} \int f_n \leqslant \int \limsup f_n.$$

.7 Let (f_n) be a sequence of summable functions such that $0 \leqslant f_{n+1}(x) \leqslant f_n(x)$ a.e. Prove that $\lim_{n \to \infty} \int f_n = 0$ if and only if $\lim_{n \to \infty} f_n(x) = 0$ a.e. (For "only if", use Corollary 15. For "if", assume that $\int f_n \to 0$, choose a subsequence $(f_{n_k})_{k \geqslant 1}$ such that $\int f_{n_k} < 2^{-k}$ for each k, and, using Lebesgue's series theorem, show that $\lim_{k \to \infty} f_{n_k}(x) = 0$ a.e.)

Let f, g be functions defined almost everywhere. We say that g **dominates** f if $|f(x)| \leqslant g(x)$ a.e.

Theorem 16 ***Lebesgue's dominated convergence theorem:*** *Let (f_n) be a sequence of summable functions that converges almost everywhere to a function f, and suppose that there exists a summable function g that dominates each f_n. Then f is summable and $\int f = \lim_{n \to \infty} \int f_n$.*

Proof. By Exercise (2.6, 6), the functions

$$g_n = \sup_{k \geqslant n} f_k$$

are summable. Moreover, (g_n) is a decreasing sequence that converges to f almost everywhere. Since $\int (-g_n) \leqslant \int g$, we can apply Beppo Levi's Theorem to the sequence $(-g_n)$, to show that f is summable and that $\int g_n \to \int f$. Replacing f by $-f$ in this argument, we see that $\int h_n \to \int f$, where

$$h_n = \inf_{k \geqslant n} f_k.$$

Finally, $h_n \leqslant f_n \leqslant g_n$, so

$$\int h_n \leqslant \int f_n \leqslant \int g_n$$

and therefore $\int f_n \to \int f$. ∎

Exercises (2.7)

.1 Prove that if f is summable, then $\int \min\{f, n\} \to \int f$ as $n \to \infty$. (Note that $\min\{f, n\}$ is summable, by Exercise (2.3, 4).)

.2 Let f be a summable function, and for each positive integer n define

$$f_n = \max\{\min\{f, n\}, -n\}.$$

Prove that $\int |f - f_n| \to 0$ as $n \to \infty$.

.3 Give two proofs that if f is a summable function, then $\int \min\left\{|f|, \frac{1}{n}\right\} \to 0$ as $n \to \infty$.

.4 Give an example of a sequence (f_n) of summable functions such that $f_n(x) \to 0$ a.e., $\lim_{n\to\infty} \int f_n = 0$, but there is no summable function that dominates each f_n.

.5 Let $f \in L$ and define $g(x) = \int_{-\infty}^{x} f$. Prove that g is continuous on $\mathbb{R}$. (*Hint*: Given $x_0 \in \mathbb{R}$, show that g is sequentially continuous at x_0 : that is, that if (x_n) is a sequence converging to x_0, then $g(x_n) \to g(x)$. To do this, consider f_n where $f_n(x) = f(x)$ if $x \leqslant x_n$, and $f_n(x) = 0$ otherwise.)

.6 Let f be summable, and let $a < b$. Suppose that $\int_a^x f = 0$ for each $x \in [a, b]$. Prove that $f(x) = 0$ a.e. in $[a, b]$. (*Hint*: choose a strictly increasing sequence (x_n) of points of (a, b) converging to b, and consider $f_n = f\chi_{[a,x_n]}$.)

.7 Let f be a real-valued function on $\mathbb{R}$ that has a bounded derivative f' on an open interval containing the compact interval $[a, b]$. Prove that f' is summable over $[a, b]$ and that $\int_a^b f' = f(b) - f(a)$. (Consider the functions

$$x \rightsquigarrow n\left(f\left(x + \frac{1}{n}\right) - f(x)\right),$$

where n is a positive integer. Once again, you will have to anticipate the connection between Riemann and Lebesgue integrals. Also, you will need to use some results from elementary calculus.)

The following result provides an alternative criterion for summability.

Proposition 17 *Let f be a real-valued function defined almost everywhere. Then f is summable if and only if there exists a sequence (f_n) of step functions such that*

(i) *For each $\varepsilon > 0$ there exists N such that $\int |f_m - f_n| < \varepsilon$ whenever $m, n \geqslant N$; and*

(ii) *$f_n(x) \to f(x)$ a.e.*

In that case, $\int |f - f_n| \to 0$ as $n \to \infty$.

Proof. We prove "if"; the converse is left as an exercise. Accordingly, assume (i) and (ii), and choose a null set A_1 such that $f_n(x) \to f(x)$ for all $x \in \mathbb{R} \backslash A_1$. Compute a strictly increasing sequence $(n_k)_{k \geqslant 1}$ of positive integers such that for each k,

$$\int |f_m - f_n| < 2^{-k} \quad (m, n \geqslant n_k)\,. \tag{1}$$

Then $\int \left|f_{n_{k+1}} - f_{n_k}\right| < 2^{-k}$ and so the series $\sum\limits_{k=1}^{\infty} \int \left|f_{n_{k+1}} - f_{n_k}\right|$ converges, by comparison with $\sum\limits_{n=1}^{\infty} 2^{-k}$. It follows from Lebesgue's series theorem that there exist a summable function ϕ and a null set A_2 such that $\sum\limits_{k=1}^{\infty} \left|f_{n_{k+1}}(x) - f_{n_k}(x)\right|$ converges to $\phi(x)$ for all $x \in \mathbb{R} \backslash A_2$. Then $A = A_1 \cup A_2$ is a null set, and for all $x \in \mathbb{R} \backslash A$ the series $\sum\limits_{k=1}^{\infty} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right)$ converges and

$$f_{n_1}(x) + \sum_{k=1}^{\infty} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right) = \lim_{k \to \infty} f_{n_k}(x) = f(x).$$

Hence f is the limit a.e. of the summable functions

$$x \rightsquigarrow f_{n_1}(x) + \sum_{k=1}^{K} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right) \quad (K = 1, 2, \cdots).$$

Since each of these functions is dominated by the summable function $|f_{n_1}| + |\phi|$, we conclude from Lebesgue's dominated convergence theorem that f is summable.

The proof of the final part of the proposition is left to the next exercise. ■

Exercises (2.8)

.1 Under the hypotheses of Proposition 17, prove that $\int |f - f_n| \to 0$ as $n \to \infty$.

.2 Prove the following converse of Proposition 17: If f is summable, then there exists a sequence (f_n) of step functions such that $\int |f - f_n| \to 0$ and $f_n(x) \to f(x)$ a.e.

3 Null Sets Revisited

A subset of $\mathbb{R}$ is said to **have measure zero**, or to **be of measure zero**, if for each $\varepsilon > 0$ it can be covered by a sequence $(I_n)_{n\geqslant 1}$ of (possibly empty) intervals such that $\sum_{n=1}^{\infty} |I_n| \leqslant \varepsilon$.

Proposition 18 *Every null set has measure zero.*

Proof. Let A be a null subset of $\mathbb{R}$, and choose a Riesz sequence (f_n) of step functions such that $f_n(t) \uparrow \infty$ for each $t \in A$, $f_n \geqslant 0$ for each n, and $0 \leqslant \int f_n \leqslant 1$ for each n. Given $\varepsilon > 0$, consider the sets

$$A_n = \left\{ x \in \mathbb{R} : f_n(x) \geqslant \frac{1}{\varepsilon} \right\} \quad (n \geqslant 1).$$

Clearly,

$$A \subset \bigcup_{n\geqslant 1} A_n \text{ and } A_n \subset A_{n+1} \tag{1}$$

Also, since each f_n is a step function, each A_n is the union of a finite number of disjoint (possibly degenerate) intervals. For each of the intervals I that make up A_n in this way we have $\chi_I \leqslant \varepsilon f_n$; so the total length of these intervals is

$$|A_n| \leqslant \varepsilon \int f_n \leqslant \varepsilon.$$

Next, define

$$B_1 = A_1, \quad B_n = A_n \backslash A_{n-1} \quad (n \geqslant 2).$$

Then B_n is a finite union of disjoint (possibly degenerate) intervals, and

$$\bigcup_{n=1}^{k} B_n = A_k \quad (k \geqslant 1).$$

Hence, by (1),

$$\bigcup_{n\geqslant 1} B_n = \bigcup_{n\geqslant 1} A_n \supset A.$$

Further, the total length of the disjoint subintervals that compose B_n is

$$|B_n| = |A_n| - |A_{n-1}|,$$

so

$$\sum_{n=1}^{k} |B_n| = |A_k| \leqslant \varepsilon \quad (k \geqslant 1)$$

and therefore $\sum_{n=1}^{\infty} |B_n|$ converges to a sum $\leqslant \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we conclude that A has measure zero. ■

Why do we introduce the sets B_n in the foregoing proof? We do so because although A is covered by the sets A_n, the series of lengths of the disjoint intervals making up all the A_n need not converge: since the set A_n are not necessarily disjoint, some of those intervals might appear infinitely often in the series.

Proposition 19 *Every set of measure zero is a null set.*

Proof. Let A be a set of measure zero. For each positive integer n, A is covered by a set A_n that is a union of a sequence $(J_{n,k})_{k\geqslant 1}$ of disjoint intervals with total length at most $1/n$. Define

$$S = \bigcap_{n\geqslant 1} A_n,$$

$$S_k = \bigcap_{n=1}^{k} A_n \quad (k \geqslant 1).$$

The characteristic function of S_k is

$$\chi_k = \min\{\chi_{A_1}, \ldots, \chi_{A_k}\},$$

and it is easily seen that

$$\chi_S(x) = \lim_{k\to\infty} \chi_k(x). \tag{1}$$

Now let $\chi_{n,i}$ be the characteristic function of $\bigcup_{k=1}^{i} J_{n,k}$, which is a step function. We have

$$\chi_{n,i} \leqslant \chi_{n,i+1} \quad (i \geqslant 1),$$

$$\lim_{i\to\infty} \chi_{n,i}(x) = \chi_{A_n}(x) \quad (x \in \mathbb{R}),$$

and

$$\int \chi_{n,i} \leqslant \sum_{k=1}^{i} |J_{n,k}| \leqslant 1/n \quad (i \geqslant 1).$$

Hence $(\chi_{n,i})_{i\geqslant 1}$ is a Riesz sequence of step functions converging almost everywhere to χ_{A_n}; so $\chi_{A_n} \in E^{\uparrow}$ and

$$0 \leqslant \int \chi_{A_n} = \lim_{i\to\infty} \int \chi_{n,i} \leqslant \frac{1}{n}.$$

It follows that χ_n is summable and that

$$0 \leqslant \int \chi_n \leqslant \int \chi_{A_n} \leqslant \frac{1}{n}.$$

Since also $\chi_k \geqslant \chi_{k+1}$, we can apply Corollary 15 to show, with reference to (1), that $\chi_S \in L$ and that

$$0 \leqslant \int \chi_S = \lim_{k\to\infty} \int \chi_k = 0.$$

As $\chi_S \geqslant 0$, it follows that $\chi_S(x) = 0$ almost everywhere. So S is a null set; whence A, being a subset of a null set, is a null set. ■

The following exercises are designed to show that not all null sets are countable.

Exercises (3.1)

.1 Let C be the **Cantor set**—that is, the subset of $[0,1]$ consisting of all numbers that have a ternary (base 3) expansion $\sum\limits_{n=1}^{\infty} a_n 3^{-n}$ with $a_n \in \{0,2\}$ for each n. Prove that if a, b are two elements of C that differ in their mth ternary places, then $|a-b| \geqslant 3^{-m}$.

.2 With C as in the previous exercise, prove that $[0,1]\backslash C$ is the union of a sequence $(J_n)_{n\geqslant 1}$ of non-overlapping open intervals whose lengths sum to 1, and that C has measure zero.

.3 For each $x = \sum\limits_{n=1}^{\infty} a_n 3^{-n} \in C$ define

$$F(x) = \sum_{n=1}^{\infty} a_n 2^{-n-1}.$$

Prove that this is a good definition of a function $F : C \to \mathbb{R}$. Prove also that F is increasing, continuous, and maps C into $[0,1]$. Hence prove that C is uncountable.

4 Riemann and Lebesgue Integrals

We now recall the basic definitions of Riemann integration theory. By a **partition** of a compact interval $I = [a, b]$ we mean a finite sequence $P = (x_0, x_1, \ldots, x_n)$ of points of I such that

$$a = x_0 \leqslant x_1 \leqslant \cdots \leqslant x_n = b.$$

The real number

$$\max \{x_{i+1} - x_i : 0 \leqslant i \leqslant n-1\}$$

is called the **mesh** of the partition. Loosely, we identify P with the set $\{x_0, \ldots, x_n\}$. A partition Q is called a **refinement** of P if $P \subset Q$.

Now let $f : I \to \mathbb{R}$ a bounded function, and for $0 \leqslant i \leqslant n-1$ define

$$\begin{aligned} m_i(f) &= \inf \{f(x) : x_i \leqslant x \leqslant x_{i+1}\}, \\ M_i(f) &= \sup \{f(x) : x_i \leqslant x \leqslant x_{i+1}\}. \end{aligned}$$

The real numbers

$$\begin{aligned} L(f, P) &= \sum_{i=0}^{n-1} m_i(f)\,(x_{i+1} - x_i), \\ U(f, P) &= \sum_{i=0}^{n-1} M_i(f)\,(x_{i+1} - x_i) \end{aligned}$$

are called the *lower sum* and *upper sum*, respectively, for f and P. Since

$$(b-a) \inf f \leqslant L(f, P) \leqslant U(f, P) \leqslant (b-a) \sup f,$$

the **lower integral** of f,

$$\underline{\int_a^b} f = \sup \{L(f, P) : P \text{ is a partition of } [a, b]\},$$

and the **upper integral** of f,

$$\overline{\int_a^b} f = \inf \{U(f, P) : P \text{ is a partition of } [a, b]\},$$

exist. Moreover,

$$\underline{\int_a^b} f \leqslant \overline{\int_a^b} f. \tag{1}$$

To see this, first observe that if P' is a refinement of a partition P, then[2]

$$L(f, P) \leqslant L(f, P') \leqslant U(f, P') \leqslant U(f, P);$$

[2] This is easily proved: first consider the case where P' is obtained by adding one point to the partition P.

so if Q is any partition, then

$$L(f,P) \leqslant L(f,P \cup Q) \leqslant U(f,P \cup Q) \leqslant U(f,Q).$$

Thus every lower sum for f is less than or equal to every upper sum, from which (1) follows.

We say that f is **Riemann integrable over** I if its lower and upper integrals coincide, in which case we define the **Riemann integral of** f **over** I to be

$$\underset{R}{\int_a^b} f = \underline{\int_a^b} f = \overline{\int_a^b} f.$$

We also define

$$\underset{R}{\int_b^a} f = -\underset{R}{\int_a^b} f$$

when f is Riemann integrable over $[a,b]$.

Theorem 20 *Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable, and extend f to $\mathbb{R}$ by setting $f(x) = 0$ if $x < a$ or $x > b$. Then f is summable, and $\int f = \underset{R}{\int_b^a} f$.*

Proof. With the partition P as above, define step functions g_P, h_P on $\mathbb{R}$ as follows:

$$g_P(x) = \begin{cases} m_i(f) & \text{if } x_i \leqslant x < x_{i+1} \\ m_{n-1}(f) & \text{if } x = x_n = b \\ 0 & \text{if } x < a \text{ or } x > b, \end{cases}$$

$$h_P(x) = \begin{cases} M_i(f) & \text{if } x_i \leqslant x < x_{i+1} \\ M_{n-1}(f) & \text{if } x = x_n = b \\ 0 & \text{if } x < a \text{ or } x > b. \end{cases}$$

Then $g_P \leqslant f \leqslant h_P$ and

$$\int g_P = L(f,P) \leqslant \underline{\int_a^b} f = \overline{\int_a^b} f \leqslant U(f,P) = \int h_P.$$

Since f is Riemann integrable, we can choose a sequence $(P_n)_{n\geqslant 1}$ of partitions of I such that for each n,

$$\int g_{P_n} = L(f,P_n) \leqslant \underline{\int_a^b} f = \overline{\int_a^b} f \leqslant U(f,P_n) = \int h_{P_n}$$

and

$$\int (h_{P_n} - g_{P_n}) < \frac{1}{n}.$$

Define step functions

$$\begin{aligned}\phi_n &= \max\{g_{P_1}, g_{P_2}, \ldots, g_{P_n}\},\\ \psi_n &= \min\{h_{P_1}, h_{P_2}, \ldots, h_{P_n}\}.\end{aligned}$$

Then for each n,

$$g_{P_n} \leqslant \phi_n \leqslant \phi_{n+1} \leqslant f \leqslant \psi_{n+1} \leqslant \psi_n \leqslant h_{P_n},$$

$$\int \phi_n \leqslant \underline{\int_a^b} f = \overline{\int_a^b} f \leqslant \int \psi_n, \tag{1}$$

and

$$0 \leqslant \int (\psi_n - \phi_n) \leqslant \int (h_{P_n} - g_{P_n}) < \frac{1}{n}. \tag{2}$$

It follows from (1) and (2) that

$$\lim_{n\to\infty} \int \phi_n = \underset{R}{\int_b^a} f \; = \lim_{n\to\infty} \int \psi_n.$$

Since $\int \phi_n \leqslant \int h_{P_1}$ for each n, we see that (ϕ_n) is a Riesz sequence of step functions; whence there exist a null set A and a function $\phi \in E^{\uparrow}$ such that $\lim_{n\to\infty} \phi_n(x) = \phi(x)$ for each $x \in \mathbb{R}\backslash A$, and

$$\lim_{n\to\infty} \int \phi_n = \int \phi. \tag{3}$$

Similarly, $(-\psi_n)$ is a Riesz sequence of step functions, and so there exists a null set $B \subset \mathbb{R}$ and a functions ψ such that $-\psi \in E^{\uparrow}$, $\lim_{n\to\infty} \psi_n(x) = \psi(x)$ for each $x \in \mathbb{R}\backslash B$, and $\lim_{n\to\infty} \int \psi_n = \int \psi$. Now, $\psi_n - \phi_n$ is summable,

$$0 \leqslant \psi_{n+1} - \phi_{n+1} \leqslant \psi_n - \phi_n,$$

and $\int (\psi_n - \phi_n) < 1/n$. By Exercise $(2.6, 7)$, there exists a null set C such that $\lim_{n\to\infty} (\psi_n(x) - \phi_n(x)) = 0$ for all $x \in \mathbb{R}\backslash C$. Then

$$S = A \cup B \cup C$$

is a null set, and for all $x \in \mathbb{R}\backslash S$,

$$\psi(x) = \phi(x) \leqslant f(x) \leqslant \psi(x)$$

and therefore $f(x) = \phi(x)$. Thus f is summable (in fact, both f and $-f$ are in $E^{\uparrow}$ and, by (3),

$$\int f = \int \phi = \lim_{n\to\infty} \int \phi_n = \underset{R}{\int_b^a} f.$$

■

The converse of Theorem 20 is false. Define $f : [0,1] \to \{0,1\}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \in [0,1] \cap \mathbb{Q} \\ 1 & \text{if } x \in [0,1] \setminus \mathbb{Q}. \end{cases}$$

Then $\underline{\int_a^b} f = 0$ and $\overline{\int_a^b} f = 1$, so f is not Riemann integrable. However, extending f by setting $f(x) = 0$ if $x < 0$ or $x > 1$, we see that $f(x) = 0$ a.e.; whence f is summable and $\int f = 0$.

Exercises (4.1)

.1 Let $f : \mathbb{R} \to \mathbb{R}$ be nonnegative, bounded, and (extended) Riemann integrable over $\mathbb{R}$ in the sense that

$$\underset{R}{\int_{-\infty}^{\infty}} f = \lim_{n\to\infty} \underset{R}{\int_{-n}^{n}} f$$

exists. Prove that f is summable and that $\int f = \int_{-\infty}^{\infty} f$.

.2 Show that the hypothesis that f is nonnegative cannot be omitted from the preceding exercise. (*Hint*: Take $f(x) = (-1)^n /n$ for $n - 1 \leqslant x < n$.)

.3 Let F be continuous on $\mathbb{R}$, and let $a < b$. Prove that for each $x \in [a,b]$, $n \int_x^{x+\frac{1}{n}} F \to F(x)$ as $n \to \infty$.

.4 Let f and f_n $(n \geqslant 1)$ be Riemann integrable over $[a,b]$. Suppose that $f_n \leqslant f_{n+1}$ for each n, and that $\lim_{n\to\infty} f_n(x) = f(x)$ for each $x \in [a,b]$. Prove that $\underset{R}{\int_a^b} f_n \to \underset{R}{\int_a^b} f$ as $n \to \infty$. (Note that we need to include the hypothesis that f is *Riemann* integrable here; without it, we could prove is that f is summable, but not necessarily that it is Riemann integrable.)

Is there a convenient characterisation of Riemann integrability? To find one, we need some definitions and lemmas. We first introduce the **limit inferior** $\underline{f} : I \to \mathbb{R}$ and the **limit superior** $\overline{f} : I \to \mathbb{R}$ of a bounded function f, defined by

$$\underline{f}(x) = \lim_{\delta\downarrow 0} \inf \{f(t) : t \in I \cap (x - \delta, x + \delta)\},$$
$$\overline{f}(x) = \lim_{\delta\downarrow 0} \sup \{f(t) : t \in I \cap (x - \delta, x + \delta)\}.$$

These are well defined. For example, if $\delta > 0$, then for each $x \in I$, as f is bounded,

$$M_\delta(x) = \sup \{f(t) : t \in I \cap (x - \delta, x + \delta)\}$$

exists; moreover, $M_\delta(x)$ is an decreasing function of δ and is bounded below by $\inf \{f(t) : t \in I\}$. Hence $\overline{f}(x) = \lim_{\delta\downarrow 0} M_\delta(x)$ exists; a similar argument shows that $\underline{f}(x)$ exists.

Lemma 21 *Let $f : [a,b] \to \mathbb{R}$ be bounded, and $x_0 \in [a,b]$. Then f is continuous at x_0 if and only if $\underline{f}(x_0) = f(x_0) = \overline{f}(x_0)$.*

Proof. Exercise. ■

Lemma 22 *If $f : [a,b] \to \mathbb{R}$ is bounded, then $\overline{f}, \underline{f}$ are summable over $[a,b]$, $\int_a^b \underline{f} = \underline{\int_a^b} f$, and $\int_a^b \underline{f} = \overline{\int_a^b} f$.*

Proof. Let (P_n) be a sequence of partitions of $I = [a,b]$ whose meshes tend to 0 as $n \to \infty$, and let A be the set of all partition points in $\bigcup\limits_{n \geqslant 1} P_n$. Then A, being countable, is a null set. Given $x \in I \backslash A$ and $n \geqslant 1$, we can find an interval J_n of the partition P_n such that $x \in J_n^\circ$. Let δ_n be the maximum, and let δ_n' be the minimum, of the distance from x to the end points of J_n. For each positive δ set

$$\begin{aligned} M_\delta(x) &= \sup\{f(t) : t \in I \cap (x - \delta, x + \delta)\}, \\ m_\delta(x) &= \inf\{f(t) : t \in I \cap (x - \delta, x + \delta)\}. \end{aligned}$$

Then

$$\begin{aligned} M_{\delta_n'}(x) &\leqslant \sup\{f(t) : t \in J_n\} \leqslant M_{\delta_n}(x), \\ m_{\delta_n}(x) &\leqslant \inf\{f(t) : t \in J_n\} \leqslant m_{\delta_n'}(x). \end{aligned}$$

Since $\delta_n \to 0$ and $\delta_n' \to 0$ as $n \to \infty$, we conclude that

$$\begin{aligned} \overline{f}(x) &= \lim_{n \to \infty} \sup\{f(x) : x \in J_n\}, \\ \underline{f}(x) &= \lim_{n \to \infty} \inf\{f(x) : x \in J_n\}. \end{aligned}$$

For each partition $P : a = x_0 < x_1 < \ldots < x_n = b$ define step functions ϕ_P, ψ_P by

$$\phi_P(x) = \begin{cases} m_i(f) & \text{if } 1 \leqslant i < n \text{ and } x_{i-1} < x < x_i \\ 0 & \text{if } x = x_i \text{ for some } i \text{ or } x < a \text{ or } x < b, \end{cases}$$

$$\psi_P(x) = \begin{cases} M_i(f) & \text{if } 1 \leqslant i < n \text{ and } x_{i-1} < x < x_i \\ 0 & \text{if } x = x_i \text{ for some } i \text{ or } x < a \text{ or } x < b. \end{cases}$$

Then for each $x \in I \backslash A$ we have

$$\begin{aligned} \underline{f}(x) &= \lim_{n \to \infty} \phi_{P_n}(x), \\ \overline{f}(x) &= \lim_{n \to \infty} \psi_{P_n}(x). \end{aligned}$$

To see this, let $\varepsilon > 0$ and choose $\delta_\varepsilon > 0$ such that if $0 < \delta < \delta_\varepsilon$, then $0 \leqslant M_\delta(x) - \overline{f}(x) < \varepsilon$. There exists a positive integer N such that $\operatorname{mesh}(P_n) < \delta_\varepsilon$

for all $n \geqslant N$. Let $n \geqslant N$, and let (x_{k-1}, x_k) be the partition interval for P_n that contains x. Then $(x_{k-1}, x_k) \subset (x - \delta_\varepsilon, x + \delta_\varepsilon)$, so

$$M_{\delta_\varepsilon}(x) \geqslant M_k \geqslant \overline{f}(x)$$

and therefore

$$0 \leqslant \psi_{P_n}(x) - \overline{f}(x) = M_k - \overline{f}(x) \leqslant M_{\delta_\varepsilon}(x) - \overline{f}(x) < \varepsilon.$$

The summable functions ϕ_{P_n}, ψ_{P_n} are dominated by the summable step function F defined by

$$F(x) = \begin{cases} \sup\{|f(t)| : t \in I\} & \text{if } x \in I \\ 0 & \text{if } x \notin I. \end{cases}$$

Hence, by Lebesgue's dominated convergence theorem, the functions $\underline{f}$ and $\overline{f}$, extended by setting them equal to 0 outside I, are summable,

$$\int f = \lim_{n\to\infty} \int \phi_{P_n} = \underline{\int_a^b} f,$$

and

$$\int f = \lim_{n\to\infty} \int \psi_{P_n} = \overline{\int_a^b} f.$$

■

Theorem 23 *A bounded function f is Riemann integrable over the compact interval $I = [a,b]$ if and only if it is continuous almost everywhere in I.*

Proof. We see that since $\overline{f} \geqslant \underline{f}$,

$$\begin{aligned} f \text{ is Riemann integrable} &\Leftrightarrow \underline{\int_a^b} f = \overline{\int_a^b} f \\ &\Leftrightarrow \int_a^b \underline{f} = \int_a^b \overline{f} \quad \text{by Lemma 22} \\ &\Leftrightarrow \int_a^b \left(\overline{f} - \underline{f}\right) = 0 \\ &\Leftrightarrow \overline{f}(x) - \underline{f}(x) = 0 \text{ a.e.in } I, \end{aligned}$$

which, by Lemma 21, holds if and only if f is continuous a.e. in I. ■

Exercises (4.2)

.1 Prove Lemma 21.

.2 Use Theorem 23 to give an alternative proof of Theorem 20.

Proposition 24 ***Integration by Parts:*** *Let f, g be summable over the compact interval $[a, b]$, and let*

$$F(x) = \int_a^x f, \quad G(x) = \int_a^x g \quad (a \leqslant x \leqslant b).$$

Then Fg and fG are summable over $[a, b]$, and

$$\int_a^b Fg + \int_a^b fG = F(b)G(b).$$

Proof. In the trivial case where $f(x) = c_1$ and $g(x) = c_2$ are constant for all $x \in [a, b]$, we use Theorem 20 to show that

$$\int_a^b Fg = \underset{R}{\int_a^b} c_1 c_2 (x - a) \, \mathrm{d}x = \tfrac{1}{2} c_1 c_2 (b - a)^2 = \int_a^b fG$$

and hence that

$$\int_a^b Fg + \int_a^b fG = c_1 c_2 (b - a)^2 = F(b)G(b).$$

The desired conclusion now follows in the case where f, g are step functions that vanish outside $[a, b]$. In the case of general summable functions f, g we use Proposition 17 and Exercise (2.8, 1) to construct sequences $(f_n), (g_n)$ of step functions such that $f_n(x) \to f(x)$ a.e., $g_n(x) \to g(x)$ a.e., $\int |f - f_n| \to 0$, and $\int |g - g_n| \to 0$ as $n \to \infty$. For each n write

$$F_n(x) = \begin{cases} \int_a^x f_n & \text{if } a \leqslant x \leqslant b \\ 0 & \text{otherwise,} \end{cases},$$

$$G_n(x) = \begin{cases} \int_a^x g_n & \text{if } a \leqslant x \leqslant b \\ 0 & \text{otherwise.} \end{cases}.$$

By (an obvious variant of) Exercise (2.7, 5), F_n and G_n are both continuous, and hence Riemann integrable, on $[a, b]$. Since g_n is a step function and therefore continuous a.e., the function $F_n g_n$ is Riemann integrable, and hence summable (Theorem 20), over $[a, b]$. Similarly, $f_n G_n$ is summable over $[a, b]$. Now, for each $x \in [a, b]$,

$$\begin{aligned} |F(x) - F_n(x)| &= \left| \int_a^x (f - f_n) \right| \\ &\leqslant \int_a^x |f - f_n| \leqslant \int |f - f_n| \to 0 \text{ as } n \to \infty \end{aligned}$$

and

$$|F_n(x)| = \left| \int_a^x f_n \right| \leqslant \int |f_n|,$$

with similar inequalities holding for $|G(x)-G_n(x)|$ and $|G_n(x)|$. Thus

$$\begin{aligned}
&|F(x)g(x)+f(x)G(x)-(F_n(x)g_n(x)+f_n(x)G_n(x))| \\
&\leqslant |F(x)-F_n(x)|\,|g(x)|+|F_n(x)|\,|g(x)-g_n(x)| \\
&\quad +|f(x)|\,|G(x)-G_n(x)|+|f_n(x)-f(x)|\,|G_n(x)| \\
&\leqslant |g(x)|\int |f-f_n|+|g(x)-g_n(x)|\int |f_n|+|f(x)|\int |g-g_n| \\
&+|f(x)-f_n(x)|\int |g_n|
\end{aligned}$$

and so

$$\begin{aligned}
&\left|\int_a^b (Fg+fG)-\int_a^b (F_ng_n+f_nG_n)\right| \\
&\leqslant \left(\int |g|\right)\int |f-f_n|+\left(\int |g-g_n|\right)\int |f_n| \\
&+\left(\int |f|\right)\int |g-g_n|+\left(\int |f-f_n|\right)\int |g_n| \\
&=\left(\int |g|+|g_n|\right)\int |f-f_n|+\left(\int |f|+|f_n|\right)\int |g-g_n|.
\end{aligned}$$

But $||g|-|g_n||\leqslant |g-g_n|$, so

$$\left|\int |g_n|-|g|\right|\leqslant \int |g-g_n|\to 0$$

—that is, $\int |g_n|\to\int |g|$. Similarly, $\int |f_n|\to\int |f|$. It follows that for all sufficiently large n we have

$$\begin{aligned}
&\left(\int |g|+|g_n|\right)\int |f-f_n|+\left(\int |f|+|f_n|\right)\int |g-g_n| \\
&\leqslant \left(1+2\int |g|\right)\int |f-f_n|+\left(1+2\int |f|\right)\int |g-g_n|\to 0 \text{ as } n\to\infty.
\end{aligned}$$

Thus

$$\left|\int_a^b (Fg+fG)-\int_a^b (F_ng_n+f_nG_n)\right|\to 0 \text{ as } n\to\infty,$$

and so

$$\begin{aligned}
\int_a^b (Fg+fG) &= \lim_{n\to\infty}\int_a^b (F_ng_n+f_nG_n) \\
&= \lim_{n\to\infty} F_n(b)G_n(b) \quad \text{(by the step-function case)} \\
&= F(b)G(b).
\end{aligned}$$

■

5 Measurable functions

Let f be a real-valued function defined almost everywhere on $\mathbb{R}$. We say that f is **measurable** if it is the limit a.e. of a sequence of step functions.

Proposition 25 *A function f is measurable if and only if it is the limit a.e. of a sequence of summable functions.*

Proof. Since step functions are summable, "only if" is trivial. Conversely, suppose that f is the limit a.e. of a sequence (f_n) of summable functions. By Proposition 17, there exists a sequence (g_n) of step functions such that $\int |f_n - g_n| < 2^{-n}$ for each n. Then $\sum\limits_{n=1}^{\infty} \int |f_n - g_n|$ converges, by comparison with $\sum\limits_{n=1}^{\infty} 2^{-n}$, so, by Lebesgue's series theorem, $\sum\limits_{n=1}^{\infty} |f_n(x) - g_n(x)|$ converges a.e. Hence $\lim_{n\to\infty} (f_n(x) - g_n(x)) = 0$ a.e. Since the union of two null sets is null, it follows that $g_n(x) \to f(x)$ a.e.; whence f is measurable. ∎

Exercises (5.1)

.1 Prove that any summable function is measurable. Give an example of a measurable function that is not summable.

.2 Prove that if f and g are measurable, then so are the functions $f + g$, $f - g$, $\max\{f, g\}$, $\min\{f, g\}$, and $|f|$.

.3 Prove that a continuous function $f : \mathbb{R} \to \mathbb{R}$ is measurable.

The following result provides a very useful test for summability.

Theorem 26 *A measurable function dominated by a summable function is summable.*

Proof. Let f be a measurable function, g a summable function, and A_1 a null set such that $|f(x)| \leqslant g(x)$ for all $x \in \mathbb{R} \backslash A_1$. Choose a sequence (f_n) of summable functions and a null set A_2 such that $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in \mathbb{R} \backslash A_2$. Then $A = A_1 \cup A_2$ is a null set. Define

$$g_n = \max\{-g, \min\{f_n, g\}\}.$$

Then g_n is summable, by Exercise (2.3, 3), and for each $x \in \mathbb{R} \backslash A$, $|g_n(x)| \leqslant g(x)$ and

$$\lim_{n\to\infty} g_n(x) = \max\{-g(x), \min\{f(x), g(x)\}\} = f(x).$$

Since g is summable, so is f, by Lebesgue's dominated convergence theorem. ∎

Exercises (5.2)

.1 Prove that a measurable function f is summable if and only if $|f|$ is summable.

.2 Prove that a summable function is summable over any interval.

.3 Let (f_n) be a sequence of measurable functions that converges almost everywhere to a function f. Prove that f is measurable. (For each k define the step function g_k by

$$g_k(x) = \begin{cases} k & \text{if } -k \leqslant x \leqslant k \\ 0 & \text{otherwise.} \end{cases}$$

First prove that $\max\{-g_k, \min\{f, g_k\}\}$ is summable.)

.4 Let f be measurable, and p a positive number. Prove that $|f|^p$ is measurable.

.5 Give an example of a measurable function f which is not summable even though f^2 is summable.

.6 Give two proofs that a product of measurable functions is measurable. (For one proof use Exercises $(5.1, 2)$ and $(5.2, 5)$.)

.7 Let f be measurable, and nonzero almost everywhere. Prove that $1/f$ is measurable. (First consider the case where $f \geqslant c$ almost everywhere for some positive constant c. For general $f \geqslant 0$, consider $f_n = 1/\left(f + n^{-1}\right)$.)

.8 Let $1 < \alpha < 2$, and define

$$f(x) = \begin{cases} x^{-\alpha} \sin x & \text{if } x > 0 \\ 0 & \text{if } x \leqslant 0. \end{cases}$$

Prove that f is summable.

.9 Prove the **Riemann–Lebesgue lemma:** If f is a summable function, then the functions $x \rightsquigarrow f(x) \sin nx$ and $x \rightsquigarrow f(x) \cos nx$ are summable, and

$$\lim_{n\to\infty} \int f(x) \sin nx \, \mathrm{d}x = 0 \quad \text{and} \quad \lim_{n\to\infty} \int f(x) \cos nx \, \mathrm{d}x = 0.$$

.10 For each positive integer n let $f_n : [0,1] \to \mathbb{R}$ be defined by

$$f_n(x) = \begin{cases} n^{-2} x^{-1/2} \cos\left(nx^{-1}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Prove that f_n is summable over $[0,1]$, that the series $\sum\limits_{n=1}^{\infty} f_n(x)$ converges almost everywhere to a function f that is summable over $[0,1]$, and that $\sum\limits_{n=1}^{\infty} \int_0^1 f_n = \int_0^1 f$.

.11 Let f be a summable function, and g a bounded summable function such that $\lim_{x\to-\infty} g(x) = 0 = \lim_{x\to\infty} g(x)$ almost everywhere. Prove that the function $h : x \rightsquigarrow g(x)f(x/n)$ is summable for each positive integer n, and that $\lim_{n\to\infty} \frac{1}{n}\int h = 0$. (*Hint*: See Exercise $(2.3, 5)$.)

Proposition 27 *Let f be a measurable function, and $\phi : \mathbb{R} \to \mathbb{R}$ a continuous function. Then $\varphi \circ f$ is measurable.*

Proof. Choose a sequence (f_n) of step functions, and a null set A, such that $f_n(x) \to f(x)$ for all $x \in \mathbb{R}\backslash A$. For each n define

$$g_n = \phi \circ f_n - \phi(0).$$

Then g_n is a step function and so is measurable. On the other hand, the constant function $x \rightsquigarrow \phi(0)$ is measurable (why?); so $\phi \circ f_n$ is measurable, as the difference of two measurable functions, by Exercise $(5.1, 2)$. The continuity of ϕ ensures that $\phi \circ f_n(x) \to \phi \circ f(x)$ for each $x \in \mathbb{R}\backslash A$. It follows from Exercise $(5.2, 3)$ that $\phi \circ f$ is measurable. ■

A subset A of $\mathbb{R}$ is called a **measurable set** (respectively, **integrable set)** if χ_A is a measurable (respectively, summable) function. A measurable subset of an integrable set is integrable, by Theorem 26.

If $A \subset \mathbb{R}$ is integrable, we define its (Lebesgue) **measure** to be $\mu(A) = \int \chi_A$.

Exercises (5.3)

.1 Let A, B be measurable sets. Prove that $A \cup B$, $A \cap B$, and $A \backslash B$ are measurable.

.2 Let (A_n) be a sequence of pairwise disjoint measurable sets. Prove that

(i) $\bigcup_{n \geqslant 1} A_n$ is measurable;

(ii) if $\sum_{n=1}^{\infty} \mu(A_n)$ is convergent, then $\bigcup_{n \geqslant 1} A_n$ is integrable, and

$$\mu\left(\bigcup_{n \geqslant 1} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

.3 Prove that any interval in $\mathbb{R}$ is measurable.

.4 Let f be a nonnegative summable function. Prove that for each $\varepsilon > 0$ there exists $\delta > 0$ such that if A is an integrable set and $\mu(A) < \delta$, then $\int_A f < \varepsilon$.

.5 Let $\mathcal{B}$ be the smallest collection of subsets of $\mathbb{R}$ that satisfies the following properties.

— Any open subset of $\mathbb{R}$ is in $\mathcal{B}$.

— If $A \in \mathcal{B}$, then $\mathbb{R} \backslash A \in \mathcal{B}$.

— The union of a sequence of elements of $\mathcal{B}$ belongs to $\mathcal{B}$.

The elements of $\mathcal{B}$ are called **Borel sets**. Prove that any Borel set is measurable.

If $\Diamond$ is a binary relation on $\mathbb{R}$ and f, g are functions defined almost everywhere on $\mathbb{R}$, we define

$$[\![f \Diamond g]\!] = \{x \in \mathbb{R} : f(x) \Diamond g(x)\}.$$

So, for example,

$$[\![f > g]\!] = \{x \in \mathbb{R} : f(x) > g(x)\}.$$

We also use analogous notations such as

$$[\![a \leqslant f < b]\!] = \{x \in \mathbb{R} : a \leqslant f(x) < b\}.$$

Just as the measurability of a set is related to that of a corresponding (characteristic) function, so the measurability of a function is related to that of certain associated sets.

Proposition 28 *Let f be a real-valued function defined almost everywhere. Then f is measurable if and only if $[\![f > r]\!]$ is measurable for each $r \in \mathbb{R}$.*

Proof. Suppose that f is measurable, let $r \in \mathbb{R}$, and for each positive integer n define

$$f_n = \frac{(f-r)^+}{\frac{1}{n} + (f-r)^+}.$$

Since the functions $t \rightsquigarrow t^+$ and

$$t \rightsquigarrow \frac{t}{\frac{1}{n} + t}$$

are continuous on $\mathbb{R}$ and $\mathbb{R}^{0+}$, respectively, we see from Exercises $(5.1, 3)$ and Proposition 27 that f_n is measurable. But

$$\lim_{n \to \infty} f_n = \chi_{[\![f > r]\!]} \text{ a.e.},$$

so $[\![f > r]\!]$ is measurable, by Exercise $(5.2, 3)$.

Now assume, conversely, that $[\![f > r]\!]$ is measurable for each $r \in \mathbb{R}$. Given a positive integer n, choose real numbers

$$\ldots, r_{-2}, r_{-1}, r_0, r_1, r_2, \ldots$$

such that $0 < r_{k+1} - r_k < 2^{-n}$ for each k. Then

$$[\![r_{k-1} < f \leqslant r_k]\!] = [\![f > r_{k-1}]\!] \backslash [\![f > r_k]\!]$$

is measurable, by Exercise $(5.3, 1)$; let χ_k denote its characteristic function. The function

$$f_n = \sum_{k=-\infty}^{\infty} r_{k-1} \chi_k$$

is measurable: for it is the limit almost everywhere of the sequence of partial sums of the series on the right-hand side, and Exercises $(5.1, 2)$ and $(5.2, 3)$ apply. To each x in the domain of f there corresponds a unique k such that $r_{k-1} < f(x) \leqslant r_k$; then

$$0 \leqslant f(x) - f_n(x) \leqslant r_k - r_{k-1} < 2^{-n}.$$

Hence the sequence (f_n) converges almost everywhere to f, which is therefore measurable, again by Exercise $(5.2, 3)$. ■

The exercises in the next set extend the ideas used in the proof of Proposition 28.

Exercises (5.4)

.1 Let f be a function defined almost everywhere on $\mathbb{R}$. Prove that the following conditions are equivalent.

(i) f is measurable.

(ii) $[\![f \geqslant r]\!]$ is measurable for each r.

(iii) $[\![f \leqslant r]\!]$ is measurable for each r.

(iv) $[\![f < r]\!]$ is measurable for each r.

(v) $[\![r \leqslant f < R]\!]$ is measurable whenever $r < R$.

.2 In the notation of the second part of the proof of Proposition 28, prove that if f is nonnegative and summable, then each f_n is summable and $\lim_{n\to\infty} \int f_n = \int f$.

.3 Let f be a nonnegative measurable function vanishing outside the interval $[a, b]$. For the purpose of this exercise, we call a sequence $(r_n)_{n=0}^{\infty}$ of real numbers *admissible* if $r_0 = 0$ and there exists $\delta > 0$ such that $r_{n+1} - r_n < \delta$ for all n; and we say that the series $\sum_{n=1}^{\infty} r_n \mu(E_n)$ *corresponds* to the admissible sequence, where E_n, whose characteristic function we denote by χ_n, is the measurable set $[\![r_{n-1} \leqslant f < r_n]\!]$. Suppose that this series converges. Let $(r'_n)_{n=0}^{\infty}$ be any admissible sequence for f, and let χ'_n be the characteristic function of $E'_n = [\![r'_{n-1} \leqslant f < r'_n]\!]$. Prove that

(i) the series $\sum_{n=1}^{\infty} r'_{n-1}\chi'_n$ and $\sum_{n=1}^{\infty} r_n\chi_n$ converge almost everywhere to summable functions,

(ii) $\sum_{n=1}^{\infty} r'_{n-1}\chi'_n \leqslant f \leqslant \sum_{n=1}^{\infty} r_n\chi_n$ almost everywhere,

(iii) the series $\sum_{n=1}^{\infty} r'_{n-1}\mu(E'_n)$ and $\sum_{n=1}^{\infty} r_n\mu(E_n)$ converge, and

(iv) $\sum_{n=1}^{\infty} r'_{n-1}\mu(E'_n) \leqslant \sum_{n=1}^{\infty} r_n\mu(E_n)$.

Hence prove that if $\sum_{n=1}^{\infty} r_n\mu(E_n)$ converges for at least one admissible sequence (r_n), then f is summable, and $\int f$ is both the infimum of the set

$$\left\{\sum_{n=1}^{\infty} r_n\mu(E_n) : (r_n) \text{ is admissible}, \forall n\,(E_n = [\![r_{n-1} \leqslant f < r_n]\!])\right\}$$

and the supremum of the set

$$\left\{\sum_{n=1}^{\infty} r_{n-1}\mu(E_n) : (r_n) \text{ is admissible}, \forall n\,(E_n = [\![r_{n-1} \leqslant f < r_n]\!])\right\}.$$

(Lebesgue's original approach to his new theory of integration [7, 8] was based on the ideas of this exercise. He first defined notions of measure of a set and measurable function. Given a bounded measurable function f on a measurable subset E of a compact interval $[a, b]$, he then formed sums analogous to $\sum_{n=1}^{\infty} r_n\mu(E_n)$ and $\sum_{n=1}^{\infty} r_{n-1}\mu(E_n)$, and showed that the infimum of sums of the first type equals the supremum of sums of the second; the common value of the infimum and supremum is precisely the Lebesgue integral of f.)

.4 By a **simple function** we mean a finite sum of functions of the form $c\chi$, where $c \in \mathbb{R}$ and χ is the characteristic function of an integrable set. Let f be a nonnegative summable function. Show that there exists a sequence (f_n) of simple functions such that

(i) $0 \leqslant f_n \leqslant f$ for each n,

(ii) $f = \sum_{n=1}^{\infty} f_n$ almost everywhere, and

(iii) $\int f = \sum_{n=1}^{\infty} \int f_n$.

(First reduce to the case where f is nonnegative and vanishes outside a compact interval. Then use the preceding exercise to construct f_k inductively such that $\int \left(f - \sum_{n=1}^{k} f_n\right) < 2^{-k}$.)

This exercise relates our development to axiomatic measure theory, which is based on primitive notions of a "measurable subset" of a set X and the "measure" of such a set, and in which the integral is often built up in the following way. First, define a function $f : X \to \mathbb{R}$ to be measurable if $[\![f < \alpha]\!]$ is a measurable set for each $\alpha \in \mathbb{R}$. Next, define the integral of a simple function $\sum_{n=1}^{N} c_n\chi_{A_n}$, where the measurable sets A_n are pairwise disjoint, to be $\sum_{n=1}^{N} c_n\mu(A_n)$. If f is a nonnegative measurable function, then define its integral to be the supremum of the integrals of simple functions s which satisfy $0 \leqslant s \leqslant f$ on the complement of a set whose measure is 0.

Are all subsets of $\mathbb{R}$ measurable? No: the axiom of choice (Appendix) ensures that non-measurable sets exist.[3] To show this, following Zermelo, we define an equivalence relation $\sim$ on $(0, 1)$ by

$$x \sim y \text{ if and only if } x - y \in \mathbb{Q}.$$

Let $\dot{x}$ denote the equivalence class of x under this relation. By the axiom of choice, there exists a function ϕ on the set of these equivalence classes such that

$$\phi(\dot{x}) \in \dot{x} \quad (x \in (0, 1)).$$

[3] Solovay has shown that there is a model of Zermelo-Fraenkel set theory, without the Axiom of Choice, in which every subset of $\mathbb{R}$ is Lebesgue measurable.

Let

$$E = \{\phi(\dot{x}) : x \in (0,1)\}.$$

Note that if $\phi(\dot{x}) \sim \phi(\dot{x}')$, then $\phi(\dot{x}) \in \dot{x}'$; whence $x' \sim \phi(\dot{x}) \in \dot{x}$, $x' \sim x$ and therefore $\dot{x} = \dot{x}'$. Now let $r_1, r_2, \ldots$ be a one-one enumeration of $\mathbb{Q} \cap (-1,1)$, and for each n define

$$E_n = E + r_n = \{y + r_n : y \in E\}.$$

We prove that

(a) for each $x \in (0,1)$ there exists $r \in \mathbb{Q} \cap (-1,1)$ such that $x \in E + r$;

(b) if $r_n < r_m$, then $E_m \cap E_n = \varnothing$.

First, observe that for each $x \in (0,1)$ we have $x \sim \phi(\dot{x})$; whence $x - \phi(\dot{x}) \in \mathbb{Q} \cap (-1,1)$ and therefore (a) holds. On the other hand, if $r_n < r_m$ and $x \in E_m \cap E_n$, then there exist $y, z \in E$ such that

$$y + r_m = x = z + r_n.$$

Hence $y - z = r_n - r_m \neq 0$; so $y \sim z$ and $y \neq z$, which is absurd since $y, z \in E$. This proves (b).

Now suppose that E is measurable and hence integrable; then each E_n is integrable and, by Exercise (2.3, 4), has measure $\mu(E)$. Since $(E_n)_{n \geqslant 1}$ is a sequence of pairwise disjoint subsets of $(-1,2)$, the partial sums of $\sum_{n=1}^{\infty} \mu(E_n)$ are bounded by 3; so $\sum_{n=1}^{\infty} \mu(E_n)$ converges. It follows from Exercise (5.3, 2) that $\bigcup_{n \geqslant 1} E_n$ is integrable and has measure $\sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{\infty} \mu(E)$. This is impossible unless $\mu(E) = 0$, which must be the case. Hence $\mu\left(\bigcup_{n \geqslant 1} E_n\right) = 0$. But, by (a), $(0,1) \subset \bigcup_{n \geqslant 1} E_n$, so $\mu\left(\bigcup_{n \geqslant 1} E_n\right) \geqslant 1$. This final contradiction shows that E is not measurable.

6 The L_p Spaces

In this section we introduce certain infinite-dimensional Banach spaces of summable functions that appear very frequently in many areas of pure and applied mathematics. For convenience, we call real numbers p, q **conjugate exponents** if $p > 1$, $q > 1$, and $1/p + 1/q = 1$.

We begin our discussion with an elementary lemma.

Lemma 29 *If x, y are positive numbers and $0 < \alpha < 1$, then*

$$x^\alpha y^{1-\alpha} \leqslant \alpha x + (1-\alpha)y.$$

Proof. Taking $u = x/y$, consider

$$f(u) = u^\alpha - \alpha u - 1 + \alpha.$$

We have $f'(u) = \alpha(u^{\alpha-1} - 1)$, which is positive if $0 < u < 1$ and negative if $u > 1$. Since $f(1) = 0$, it follows that $f(u) \leqslant 0$ for all $u > 0$. This immediately leads to the desired inequality. ■

Lemma 30 *Let p, q be conjugate exponents, and f, g measurable functions on $\mathbb{R}$ such that $|f|^p$ and $|g|^q$ are summable. Then fg is summable, and **Hölder's inequality***

$$\left|\int fg\right| \leqslant \left(\int |f|^p\right)^{1/p} \left(\int |g|^q\right)^{1/q} \tag{1}$$

holds.

Proof. We first note that if $\int |f|^p = 0$, then $|f|^p = 0$ almost everywhere; so $f = 0$, and therefore $fg = 0$, almost everywhere. Then fg is summable, $\int fg = 0$, and (1) holds trivially, as it does also in the case where $\int |g|^q = 0$. Thus we may assume that $\int |f|^p > 0$ and $\int |g|^q > 0$. We then have, almost everywhere,

$$\begin{aligned}\frac{|fg|}{\left(\int |f|^p\right)^{1/p} \left(\int |g|^q\right)^{1/q}} &= \left(\frac{|f|^p}{\int |f|^p}\right)^{1/p} \left(\frac{|g|^q}{\int |g|^q}\right)^{1/q} \\ &\leqslant \frac{|f|^p}{p\int |f|^p} + \frac{|g|^q}{q\int |g|^q}\end{aligned}$$

(where the last step uses Lemma 29), so

$$|fg| \leqslant \left(\int |f|^p\right)^{1/p} \left(\int |g|^q\right)^{1/q} \left(\frac{|f|^p}{p\int |f|^p} + \frac{|g|^q}{q\int |g|^q}\right). \tag{2}$$

Now, fg is measurable and the right-hand side of (2) is summable. Hence, by Theorem 26, fg is summable and

$$\left|\int fg\right| \leqslant \int |fg| \leqslant \left(\int |f|^p\right)^{1/p} \left(\int |g|^q\right)^{1/q} \left(\frac{1}{p} + \frac{1}{q}\right),$$

from which (1) follows. ■

Proposition 31 *Let $p \geqslant 1$, and let f, g be measurable functions on $\mathbb{R}$ such that $|f|^p$ and $|g|^p$ are summable. Then $|f+g|^p$ is summable, and **Minkowski's inequality***

$$\left(\int |f+g|^p\right)^{1/p} \leqslant \left(\int |f|^p\right)^{1/p} + \left(\int |g|^p\right)^{1/p}$$

holds.

Proof. Clearly, we may assume that $p > 1$. Now, $|f+g|^p$ is measurable, by Exercise (2.3.3: 5). Since

$$|f+g|^p \leqslant (2\max\{|f|, |g|\})^p \leqslant 2^p (|f|^p + |g|^p)$$

and the last function is summable, it follows from Theorem 26 that $|f+g|^p$ is summable. The functions $|f|$ and $|f+g|^{p-1}$ are measurable, by Exercise (5.2, 4), and

$$\left(|f+g|^{p-1}\right)^q = |f+g|^p \in L_1(\mathbb{R}).$$

Thus, by Lemma 30, $|f+g|^{p-1}|f|$ is summable and

$$\int |f+g|^{p-1}|f| \leqslant \left(\int |f+g|^p\right)^{1-p^{-1}} \left(\int |f|^p\right)^{1/p}.$$

Similarly, $|f+g|^{p-1}|g|$ is summable and

$$\int |f+g|^{p-1}|g| \leqslant \left(\int |f+g|^p\right)^{1-p^{-1}} \left(\int |g|^p\right)^{1/p}.$$

It follows that

$$\begin{aligned}\int |f+g|^p &= \int |f+g|^{p-1}|f+g| \\ &\leqslant \int |f+g|^{p-1}|f| + \int |f+g|^{p-1}|g| \\ &\leqslant \left(\int |f+g|^p\right)^{1-p^{-1}} \left(\left(\int |f|^p\right)^{1/p} + \left(\int |g|^p\right)^{1/p}\right),\end{aligned}$$

from which we easily obtain Minkowski's inequality. ■

Exercises (6.1)

.1 Prove Hölder's inequality

$$\left|\sum_{n=1}^N x_n y_n\right| \leqslant \left(\sum_{n=1}^N |x_n|^p\right)^{1/p} \left(\sum_{n=1}^N |y_n|^q\right)^{1/q}$$

and Minkowski's inequality

$$\left(\sum_{n=1}^N |x_n+y_n|^p\right)^{1/p} \leqslant \left(\sum_{n=1}^N |x_n|^p\right)^{1/p} + \left(\sum_{n=1}^N |y_n|^p\right)^{1/p}$$

for finite sequences $x_1, \ldots, x_N$ and $y_1, \ldots, y_N$ of real numbers.

.2 A sequence (x_n) of real numbers is called ***p*-power summable** if the series $\sum_{n=1}^{\infty} |x_n|^p$ converges. Prove that if (x_n) is p-power summable and (y_n) is q-power summable, where p, q are conjugate exponents, then

(i) $\sum_{n=1}^{\infty} x_n y_n$ is absolutely convergent, and

(ii) Hölder's inequality holds in the form

$$\left|\sum_{n=1}^{\infty} x_n y_n\right| \leqslant \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p} \left(\sum_{n=1}^{\infty} |y_n|^q\right)^{1/q}.$$

Prove also that if (x_n) and (y_n) are both p-power summable, then so is $(x_n + y_n)$, and Minkowski's inequality

$$\left(\sum_{n=1}^{\infty} |x_n + y_n|^p\right)^{1/p} \leqslant \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p} + \left(\sum_{n=1}^{\infty} |y_n|^p\right)^{1/p}$$

holds.

.3 Let $p \geqslant 1$, and let l_p denote the set of all p-power summable sequences, taken with termwise addition and multiplication-by-scalars. Prove that

$$\left\|(x_n)_{n\geqslant 1}\right\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p}$$

defines a norm on l_p. (We define the normed space $l_p(\mathbb{C})$ of p-power summable sequences of complex numbers in the obvious analogous way.)

Let X be a measurable subset of $\mathbb{R}$, and $p \geqslant 1$. We define $L_p(X)$ to be the set of all functions f defined almost everywhere on $\mathbb{R}$ such that f is measurable, f vanishes almost everywhere on $\mathbb{R}\backslash X$, and $|f|^p$ is summable. Taken with the pointwise operations of addition and multiplication-by-scalars, $L_p(X)$ becomes a linear space. If we follow the usual practice of identifying two measurable functions that are equal almost everywhere, then

$$\|f\|_p = \left(\int |f|^p\right)^{1/p}$$

is a norm, called the L_p**-norm**, on $L_p(X)$: the only property of a norm that requires nontrivial verification is the triangle inequality, which in this case is Minkowski's inequality, disposed of in Proposition 31.

When $X = [a, b]$ is a compact interval, we write $L_p[a, b]$ rather than $L_p([a, b])$.

Exercises (6.2)

In these exercises, X is a measurable subset of $\mathbb{R}$.

.1 Let X be summable and $1 \leqslant r < s$. Prove the following.

(i) $L_s(X) \subset L_r(X)$. (Note that if $f \in L_s(X)$, then $|f|^r \in L_{s/r}(X)$.)

(ii) The linear mapping $f \rightsquigarrow f$ of $L_s(X)$ into $L_r(X)$ is bounded and has norm $\leqslant \mu(X)^{r^{-1}-s^{-1}}$.

.2 Let $1 \leqslant r \leqslant t \leqslant s < \infty$, $r \neq s$,

$$\alpha = \frac{t^{-1} - s^{-1}}{r^{-1} - s^{-1}}, \quad \beta = \frac{r^{-1} - t^{-1}}{r^{-1} - s^{-1}}.$$

and $f \in L_r(X) \cap L_s(X)$. Prove that $f \in L_t(X)$ and

$$\|f\|_t \leqslant \|f\|_r^\alpha \, \|f\|_s^\beta \, .$$

(Consider $|f|^{\alpha t} \, |f|^{\beta t}$.)

.3 Prove that the step functions that vanish outside X form a dense subspace of $L_p(X)$ for $p \geqslant 1$. (First consider the case where X is a compact interval.)

.4 **Clarkson's inequalities:** Let p, q be conjugate exponents, and let $f, g \in L_p(X)$. Prove that if $1 < p < 2$, then

$$2 \left(\|f\|_p^p + \|g\|_p^p \right)^{q-1} \geqslant \|f + g\|_p^q + \|f - g\|_p^q$$

and

$$\|f + g\|_p^p + \|f - g\|_p^p \geqslant 2 \left(\|f\|_p^q + \|g\|_p^q \right)^{p-1},$$

and that the reverse inequalities hold if $p \geqslant 2$.

Theorem 32 ***The Riesz–Fischer theorem:*** *$L_p(X)$ is a Banach space for all $p \geqslant 1$. More precisely, if (f_n) is a Cauchy sequence in $L_p(X)$, then there exist $f \in L_p(X)$ and a subsequence $(f_{n_k})_{k \geqslant 1}$ of (f_n) such that*

(i) $\lim_{n\to\infty} \|f - f_n\|_p = 0$, *and*

(ii) $f_{n_k} \to f$ *almost everywhere on X as $k \to \infty$.*

Proof. We illustrate the proof with the case $X = \mathbb{R}$ and $p > 1$. Given a Cauchy sequence (f_n) in $L_p(\mathbb{R})$, choose a subsequence $(f_{n_k})_{k \geqslant 1}$ such that

$$\|f_m - f_n\|_p \leqslant 2^{-k} \quad (m, n \geqslant n_k).$$

Then

$$\left\|f_{n_{k+1}} - f_{n_k}\right\|_p \leqslant 2^{-k}.$$

Writing $q = p/(p-1)$, we see from Lemma 30 that for each positive integer N, $\left|f_{n_{k+1}} - f_{n_k}\right|$ is summable over $[-N, N]$, and

$$\begin{aligned}\int \left|f_{n_{k+1}} - f_{n_k}\right| \chi_{[-N,N]} &\leqslant \left\|f_{n_{k+1}} - f_{n_k}\right\|_p \left(\int \chi_{[-N,N]}\right)^{1/q} \\ &\leqslant 2^{-k}(2N)^{1/q},\end{aligned}$$

so the series

$$\sum_{k=1}^{\infty} \int \left|f_{n_{k+1}} - f_{n_k}\right| \chi_{[-N,N]}$$

converges. It follows from Lebesgue's series theorem that there exists a null set E_N such that the series

$$\sum_{k=1}^{\infty} \left|f_{n_{k+1}}(x) - f_{n_k}(x)\right| \chi_{[-N,N]}(x)$$

converges for all $x \in \mathbb{R} \backslash E_N$, and the function $\sum\limits_{k=1}^{\infty} \left|f_{n_{k+1}} - f_{n_k}\right| \chi_{[-N,N]}$ is summable. Then

$$E = \bigcup_{N=1}^{\infty} E_N$$

is a null set, and

$$f(x) = \lim_{k \to \infty} f_{n_k}(x) = f_{n_1}(x) + \sum_{k=1}^{\infty} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right)$$

exists for all $x \in \mathbb{R} \backslash E$. The function f so defined is measurable, by Exercise $(5.2, 3)$. Since

$$\left\|f_{n_k}\right\|_p \leqslant \left\|f_{n_1}\right\|_p + \left\|f_{n_k} - f_{n_1}\right\|_p \leqslant \left\|f_{n_1}\right\|_p + \tfrac{1}{2}$$

for all k, we see from Fatou's lemma (Exercise $(2.6, 5)$) that $|f|^p$ is summable and hence that $f \in L_p(\mathbb{R})$. Moreover, if $n \geqslant n_i$, then by applying Fatou's lemma to the sequence $(|f_{n_k} - f_n|)_{k=i}^{\infty}$ we see that

$$\|f - f_n\|_p = \lim_{k \to \infty} \left\|f_{n_k} - f_n\right\|_p \leqslant 2^{-i}.$$

Hence $\lim_{n \to \infty} \|f - f_n\|_p = 0$. ■

Exercises (6.3)

.1 Prove the Riesz–Fischer theorem for a general measurable set $X \subset \mathbb{R}$. Prove it also in the case $p = 1$.

.2 Prove that the space l_p is complete for $p \geqslant 1$. (See Exercise $(6.1, 3)$.)

7 Double Integrals

We next illustrate multiple Lebesgue integration by considering double integrals

By an **interval in** $\mathbb{R}^2$ we mean the Cartesian product $I \times J$ of two intervals in $\mathbb{R}$. We say that $f : \mathbb{R}^2 \to \mathbb{R}$ is a**step function** if there exists a finite set of non-overlapping bounded intervals $I_k \times J_k$ $(k = 1, \cdots, n)$ such that

- $\triangleright$ $f(x) = 0$ for all x in the complement of $\bigcup_{k=1}^{n} I_k \times J_k$, and
- $\triangleright$ for each k, f has a constant value c_k on $I_k \times J_k$.

We define the **double integral** of such a function f to be

$$\int \int f = \sum_{k=1}^{n} c_k \left| I_k \right| \left| J_k \right|$$

where, we recall, $|I_k|$ denote the length of the interval I_k. Many fact, such as the following, about the double integral of a step function are established analogously to the corresponding facts for integrals of step functions on $\mathbb{R}$; in such cases we state the facts without proof.

Proposition 33 *Let f and g be two step functions on $\mathbb{R}^2$ such that $f(x) = g(x)$ for all but finitely many points x. Then $\int \int f = \int \int g$.*

Corollary 34 *The double integral of a step function f on $\mathbb{R}^2$ depends only on f and not on the choice of the finite set of non-overlapping intervals on each of which f is constant.*

Proposition 35 *Let f and g be step functions on $\mathbb{R}^2$, and let $\alpha, \beta \in \mathbb{R}$. Then $\alpha f + \beta g$ is a step function, and the double integral is a linear function:*

$$\int \int (\alpha f + \beta g) = \alpha \int \int f + \beta \int \int g.$$

Moreover, fg, $\max\{f, g\}$, $\min\{f, g\}$,

$$f^+ = \max\{f, 0\},$$

and

$$f^- = \max\{-f, 0\}$$

are step functions. Finally, if $f(x) \leqslant g(x)$ for all but finitely many (possibly no) points x, then $\int \int f \leqslant \int \int g$.

Proposition 36 *Let (f_n) be a sequence of step functions such that*

(i) $0 \leqslant f_{n+1} \leqslant f_n$ *for each n and*

(ii) $\lim_{n\to\infty} f_n(x) = 0$.

Then $\lim_{n\to\infty} \int\int f_n = 0$.

As in the one-variable case, a sequence (f_n) of step functions on $\mathbb{R}^2$ is called a **Riesz sequence of step functions** if $f_1 \leqslant f_2 \leqslant \cdots$ and the sequence $\left(\int\int f_n\right)_{n\geqslant 1}$ is bounded above. The sequence of integrals $\int\int f_n$ is then both increasing and bounded above, and so converges to a limit $l \in \mathbb{R}$.

A subset A of $\mathbb{R}^2$ is called a **null set** if there exists a Riesz sequence (f_n) of step functions such that $f_n(x) \uparrow \infty$ for each $x \in A$. It follows immediately that every subset of a null set is also a null set. A countable union of null sets is a null set; and a subset A of $\mathbb{R}^2$ is null if and only if it has **measure zero**, in the sense that for each $\varepsilon > 0$ there exists a sequence $(I_n \times J_n)_{n\geqslant 1}$ of intervals in $\mathbb{R}^2$ such that $A \subset \bigcup\limits_{n\geqslant 1} I_n \times J_n$ and $\sum\limits_{n=1}^{\infty} |I_n|\,|J_n| < \varepsilon$.

We say that a property $P(x)$ of real numbers holds **almost everywhere** if there exists a null set A such that $P(x)$ holds for all $x \in \mathbb{R}^2 \backslash A$; we then write "$P(x)$ a.e.". By the definition of "null set" a Riesz sequence of step functions converges almost everywhere.

We denote by $E^{\uparrow}\left(\mathbb{R}^2\right)$ the set of functions f that are limits almost everywhere of some Riesz sequence (f_n). The **double integral** of such a function f is defined to be $\int\int f = \lim_{n\to\infty} \int\int f_n$, and is independent of the Riesz sequence (f_n) converging almost everywhere to f.

Now let

$$L\left(\mathbb{R}^2\right) = \left\{f_1 - f_2 : f_1, f_2 \in E^{\uparrow}\left(\mathbb{R}^2\right)\right\}.$$

The elements of L are called **(Lebesgue) summable** functions. If $f_1, f_2 \in E^{\uparrow}$ and $f = f_1 - f_2$, define the (**Lebesgue) double integral** of f to be

$$\int\int f = \int\int f_1 - \int\int f_2.$$

As in the one-variable case, this is a good definition; $L\left(\mathbb{R}^2\right)$ is a linear space under pointwise algebraic operations, and $\int\int$ is a linear mapping of $L\left(\mathbb{R}^2\right)$ into $\mathbb{R}$; if f, g are summable functions defined almost everywhere on $\mathbb{R}^2$, then $\max\{f, g\}$, $\min\{f, g\}$, and $|f|$ are summable, and $\left|\int f\right| \leqslant \int |f|$; and the Lebesgue double integral is **translation invariant,** in the sense that if $f \in L\left(\mathbb{R}^2\right)$ and $y \in \mathbb{R}^2$, then the function f_y defined by $f_y(x) = f(x + y)$ is summable, and $\int\int f_y = \int\int f$.

We say that $f \in L\left(\mathbb{R}^2\right)$ is **summable over the subset** A **of** $\mathbb{R}^2$ if the function $f\chi_A$ belongs to $L\left(\mathbb{R}^2\right)$; we then write

$$\int\int_A f = \int\int f\chi_A.$$

If f is summable over the subsets A, B of $\mathbb{R}^2$, then it is summable over $A \cup B$ and $A \cap B$. Any element of $L\left(\mathbb{R}^2\right)$ is summable over any interval in $\mathbb{R}^2$.

With all the foregoing definitions at hand, we can modify the arguments used in the one-variable case to prove double-integral analogues of Lebesgue's series theorem, Beppo Levi's Theorem, Fatou's lemma, Lebesgue's dominated convergence theorem, ... We can also define measurability of functions defined almost everywhere on $\mathbb{R}^2$ and of subsets of $\mathbb{R}^2$, and then prove analogues of Propositions 25–28.

Our main aim in this section is to deal with the relation between double integration and repeated integration, familiar from second-year calculus courses. First, given a two-variable function f, we define a one-variable function $f_x : \mathbb{R} \to \mathbb{R}$ by $f_x(y) = f(x, y)$; likewise, if (f_n) is a sequence of two-variable functions, we define the corresponding one-variable functions $f_{n,x}$ by $f_{n,x}(y) = f_n(x, y)$.

Lemma 37 *If f is a step function on $\mathbb{R}^2$, then*

(i) *for each $x \in \mathbb{R}$, f_x is a step function;*

(ii) *$g(x) = \int f_x$ is a step function on $\mathbb{R}$; and*

(iii) *$\int\int f = \int g = \int \left(\int f_x\right)$.*

Proof. The general case easily follows from that in which f is a constant c on a product $I \times J$ of open intervals in $\mathbb{R}$, and is 0 elsewhere. Taking that case, for each $x \notin I$ we have $f_x = 0$ and therefore $\int f_x = 0$; while for each $x \in I$,

$$f_x(y) = \begin{cases} c & \text{if } y \in J \\ \\ 0 & \text{if } y \notin J \end{cases}$$

and so, by definition of the integral on $\mathbb{R}$, $\int f_x = c\,|J|$. Thus (i) holds, g is well defined, and

$$g(x) = \begin{cases} c\,|J| & \text{if } x \in I \\ \\ 0 & \text{if } x \notin I \end{cases}$$

Clearly, g is a step function on $\mathbb{R}$, and

$$\int g = c\,|J|\,|I| = \int\int f,$$

by definition of the double integral. ■

Lemma 38 *If A is a null set in $\mathbb{R}^2$, then there exists a null set B in $\mathbb{R}$ such that for each $x \in \mathbb{R} \backslash B$,*

$$A(x) = \{y \in \mathbb{R} : (x, y) \in A\}$$

is a null set in $\mathbb{R}$.

Proof. Choose a sequence $(f_n)_{n \geqslant 1}$ of step functions in $\mathbb{R}^2$ and a positive number M such that

- ▷ $f_1 \leqslant f_2 \leqslant \cdots$,
- ▷ $\int\int f_n \leqslant M$ for each n, and
- ▷ $f_n(x,y) \to \infty$ for all $(x,y) \in A$.

It follows from the first and third of these properties that for each $x \in \mathbb{R}$ the sequence $(f_{n,x})_{n\geqslant 1}$ is increasing, and $f_{n,x}(y) \to \infty$ for all $y \in A(x)$. On the other hand, by Lemma 37, $f_{n,x}$ is a step function on $\mathbb{R}$, the function

$$g_n : x \rightsquigarrow \int f_{n,x}$$

is a step function on $\mathbb{R}$, and $\int g_n = \int\int f_n \leqslant M$. Thus $(g_n)_{n\geqslant 1}$ is a Riesz sequence of step functions on $\mathbb{R}$. Hence there exists a null subset B of $\mathbb{R}$ such that $(g_n(x))_{n\geqslant 1}$ converges for all $x \in \mathbb{R}\backslash B$. In particular, for each such x there exists $M_x > 0$ such that $\int f_{n,x} \leqslant M_x$ for all n. Thus $(f_{n,x})_{n\geqslant 1}$ is a Riesz sequence of step functions on $\mathbb{R}$ and therefore converges almost everywhere in $\mathbb{R}$. Since this sequence diverges on $A(x)$, we conclude that $A(x)$ is a null set. ■

This brings us to the fundamental theorem of double integration, **Fubini's theorem.**

Theorem 39 *If f is summable over $\mathbb{R}^2$, then there exist a summable function g on $\mathbb{R}$ and a null set S in $\mathbb{R}$ such that*

(i) *for each $x \in \mathbb{R}\backslash S$, f_x is summable over $\mathbb{R}$ and $g(x) = \int f_x$;*

(ii) $\int g = \int\int f$.

Proof. It is enough to consider the case where $f \in E^{\uparrow}(\mathbb{R}^2)$. Choose a Riesz sequence (f_n) of step functions on $\mathbb{R}^2$ and a null set $A \subset \mathbb{R}^2$ such that $\lim_{n\to\infty} f_n(x,y) = f(x,y)$ for all $(x,y) \in \mathbb{R}^2\backslash A$, and $\lim_{n\to\infty}\int\int f_n = \int\int f$. By Lemma 37, for each n the mapping

$$g_n : x \rightsquigarrow \int f_{n,x}$$

exists as a step function on $\mathbb{R}$, and $\int g_n = \int\int f_n$. Hence $\lim_{n\to\infty}\int g_n = \int\int f$. Since $g_n \leqslant g_{n+1}$ and a convergent sequence is bounded, we see that (g_n) is a Riesz sequence of step functions on $\mathbb{R}$; so there exist $g \in E^{\uparrow}$ and a null set $S_1 \subset \mathbb{R}$ such that $\lim_{n\to\infty} g_n(x) = g(x)$ for each $x \in \mathbb{R}\backslash S_1$, and $\lim_{n\to\infty}\int g_n = \int g$.

Now, by Lemma 38, there exists a null set $S_2 \subset \mathbb{R}$ such that $A(x)$ is a null set for each $x \in \mathbb{R}\backslash S_2$. Then $S = S_1 \cup S_2$ is a null set in $\mathbb{R}$; $f_x(y) = \lim_{n\to\infty} f_n(x,y)$ for each $x \in \mathbb{R}\backslash S$ and all $y \in \mathbb{R}\backslash A(x)$; and $\left(\int f_n\right)_{n\geqslant 1}$ is a convergent, and hence bounded, sequence in $\mathbb{R}$. It follows that for each $x \in \mathbb{R}\backslash S$, $f_x \in E^{\uparrow}$ and

$$\int f_x = \lim_{n\to\infty}\int f_{n,x} = \lim_{n\to\infty} g_n(x) = g(x).$$

This proves (i). Moreover, we have

$$\int\left(\int f_x\right)=\int g=\lim_{n\to\infty}\int g_n=\int\int f,$$

which proves (ii). ■

Corollary 40 *If f is summable over $\mathbb{R}^2$, then there exist a summable function h on $\mathbb{R}$ and a null set T in $\mathbb{R}$ such that*

- ▷ *for each $y\in\mathbb{R}\backslash T$, the function $x\rightsquigarrow f(x,y)$ is summable over $\mathbb{R}$ and $h(y)=\int(x\rightsquigarrow f(x,y))$;*
- ▷ *$\int h=\int\int f$.*

Proof. Exercise. ■

Here is a partial converse of Fubini's theorem.

Theorem 41 *Let f be a nonnegative measurable function on $\mathbb{R}^2$. Suppose that there exist a summable function g on $\mathbb{R}$ and a null set $S\subset\mathbb{R}$ such that for each $x\in\mathbb{R}\backslash S$, the function f_x is summable and $g(x)=\int f_x$. Then f is summable and $\int\int f=\int g$.*

Proof. For each n define the step function ϕ_n by

$$\phi_n(x,y)=\begin{cases} n & \text{if } |x|\leqslant n \text{ and } |y|\leqslant n \\ 0 & \text{otherwise,}\end{cases}$$

and define

$$f_n=\min\{f,\phi_n\}.$$

Then f_n is measurable and dominated by the summable function ϕ_n (this is where we need $f\geqslant 0$), so f_n is summable. We prove that the sequence $\left(\int\int f_n\right)_{n\geqslant 1}$ is bounded. To this end, we apply Fubini's theorem to obtain, for each n, a summable function g_n on $\mathbb{R}$ and a null set $A_n\subset\mathbb{R}$ such that for each $x\in\mathbb{R}\backslash A_n$, $f_{n,x}$ is summable, $\int f_{n,x}=g_n(x)$, and $\int g_n=\int\int f_n$. Then

$$A=\bigcup_{n\geqslant 1}A_n$$

is a null set in $\mathbb{R}$, as is $S\cup A$. Moreover, for each n and each $x\in\mathbb{R}\backslash(S\cup A)$,

$$g_n(x)=\int f_{n,x}\leqslant\int f_x=g(x).$$

Hence $\int\int f_n=\int g_n\leqslant\int g$. Since $f_n\leqslant f_{n+1}$, we see that $(f_n)_{n\geqslant 1}$ is a Riesz sequence of summable functions on $\mathbb{R}^2$.

Now, $f_n(x,y) \to f(x,y)$ for all (x,y) in the domain of f, the complement of a set of measure zero. So, by Beppo Levi's Theorem, f is summable. Finally, by Fubini's theorem, there exist a summable function h on $\mathbb{R}$ and a null set $C \subset \mathbb{R}$ such that for each $x \in \mathbb{R}\backslash C$, f_x is summable and $h(x) = \int f_x$, and such that $\int h = \int\int f$; whence $h(x) = g(x)$ for all $x \in \mathbb{R}\backslash(S \cup C)$ and

$$\int g = \int h = \int\int f.$$

■

Exercises (7.1)

.1 Prove Corollary 40.

.2 Let f be a measurable function on $\mathbb{R}^2$ for which there exist a summable function g on $\mathbb{R}$ and a null set $A \subset \mathbb{R}$ such that $|f_x|$ is summable and $g(x) = \int |f_x|$ for each $x \in \mathbb{R}\backslash A$. Prove that f is summable.

.3 For each $\alpha \in \mathbb{R}$ the function $f_\alpha : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$f_\alpha(x,y) = \begin{cases} x(1-xy)^\alpha & \text{if } 0 \leqslant x < 1,\ 0 \leqslant y < 1 \\ \\ 0 & \text{otherwise.} \end{cases}$$

Find the values of α for which f_α is summable, and evaluate $\int f_\alpha$ for those values.

8 The Vitali covering theorem

The **outer measure** of a subset A of $\mathbb{R}$ is the quantity

$$\mu^*(A) = \inf\left\{\sum_{n\geqslant 1} |I_n| : (I_n)_{n\geqslant 1} \text{ is a cover of } A \text{ by bounded open intervals}\right\},$$

which we take as ∞ if the set on the right-hand side is unbounded.[4] If $\mu^*(A) \in \mathbb{R}$, we say that A **has finite outer measure**. Note that since, for any sequence $(I_n)_{n\geqslant 1}$ of bounded open intervals that covers A, the terms of the series $\sum_{n=1}^{\infty} |I_n|$ are all positive, the (possibly infinite) sum of the series does not depend on the order of those terms; this is a special case of Riemann's theorem on the rearrangement of absolutely convergent series ([3], page 34).

Note that A has outer measure zero if and only if for each $\varepsilon > 0$ there exists a sequence $(I_n)_{n\geqslant 1}$ of bounded open intervals such that $A \subset \bigcup_{n\geqslant 1} I_n$ and $\sum_{n=1}^{\infty} |I_n| < \varepsilon$ —in other words, if and only if A is a null set.

[4] It is possible to give a precise meaning to this use of ∞ as an "extended real number"; see pages 128–129 of [3].

Exercises (8.1)

.1 Show that for each $A \subset \mathbb{R}$, $\mu^*(A)$ is the infimum of $\sum_{n=1}^{\infty} |I_n|$ taken over all covers of A by sequences $(I_n)_{n \geqslant 1}$ of bounded, but not necessarily open, intervals.

.2 Prove that if a subset A of $\mathbb{R}$ has finite outer measure, then for each $\varepsilon > 0$ there exists a sequence (I_n) of disjoint bounded open intervals such that $A \subset \bigcup_{n \geqslant 1} I_n$ and $\sum_{n=1}^{\infty} |I_n| < \mu^*(A) + \varepsilon$. (Recall that a nonempty subset of $\mathbb{R}$ is open if and only if it is the union of a sequence $(J_n)_{n \geqslant 1}$ of pairwise disjoint sets, each of which is either empty or an open interval.)

.3 Show that $\mu^*(\varnothing) = 0$, and that if $A \subset B$, then $\mu^*(A) \leqslant \mu^*(B)$.

.4 Prove that for each $a \in \mathbb{R}$, $\mu^*(\{a\}) = 0$.

.5 Let A be a subset of $\mathbb{R}$, and $E \subset A$ a set of measure zero. Show that $\mu^*(A \backslash E) = \mu^*(A)$.

.6 Let A be a subset of a compact interval I. Prove that $\mu^*(A) + \mu^*(I \backslash A) \geqslant |I|$. (Perhaps surprisingly, we cannot replace inequality by equality in this result; this is a consequence of the existence of non-measurable sets.)

.7 Let (A_n) be a sequence of subsets of $\mathbb{R}$. Show that

$$\mu^*\left(\bigcup_{n \geqslant 1} A_n\right) \leqslant \sum_{n=1}^{\infty} \mu^*(A_n),$$

where the right-hand side is taken as ∞ if either any of its terms is ∞ or the series diverges. (If one of the sets A_n has infinite outer measure, then the inequality is trivial. If each A_n has finite outer measure, then for each positive integer n and each $\varepsilon > 0$ there exists a sequence $(I_{n,k})_{k \geqslant 1}$ of bounded open intervals such that $A_n \subset \bigcup_{k=1}^{\infty} I_{n,k}$ and $\sum_{k=1}^{\infty} |I_{n,k}| < \mu^*(A_n) + 2^{-n}\varepsilon$.)

Prove that if also the sets A_n are pairwise disjoint, then

$$\mu^*\left(\bigcup_{n \geqslant 1} A_n\right) = \sum_{n=1}^{\infty} \mu^*(A_n).$$

.10 Prove that a subset A of $\mathbb{R}$ has finite outer measure if and only if $l = \lim_{n \to \infty} \mu^*(A \cap [-n, n])$ exists, in which case $\mu^*(A) = l$.

.11 Prove that μ^* is translation invariant—that is, $\mu^*(A + t) = \mu^*(A)$ for each $A \subset \mathbb{R}$ and each $t \in \mathbb{R}$, where $A + t = \{x + t : x \in A\}$.

Proposition 42 *The outer measure of any interval in $\mathbb{R}$ equals the length of the interval.*

Proof. Consider, to begin with, a bounded closed interval $[a, b]$. For each $\varepsilon > 0$ we have $[a, b] \subset (a - \varepsilon, b + \varepsilon)$ and therefore

$$\mu^*([a, b]) \leqslant |(a - \varepsilon, b + \varepsilon)| = b - a + 2\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we conclude that $\mu^*([a, b]) \leqslant b - a$. To prove the reverse inequality, let (I_n) be any sequence of bounded open intervals that covers $[a, b]$. Applying the Heine–Borel–Lebesgue Theorem (1.4.6), and re-indexing the terms I_n (which we can do without loss of generality), we may assume that for some N,

$$[a, b] \subset I_1 \cup I_2 \cup \cdots \cup I_N.$$

There exists an interval I_{k_1}, where $1 \leqslant k_1 \leqslant N$, that contains a; let this interval be (a_1, b_1). Either $b < b_1$, in which case we stop the procedure, or else $b_1 \leqslant b$. In the latter case, $b_1 \in [a, b] \backslash (a_1, b_1)$; so there exists an interval I_{k_2}, where $1 \leqslant k_2 \leqslant N$ and $k_2 \neq k_1$, that contains b_1; call this interval (a_2, b_2). Repeating this argument, we obtain intervals $(a_1, b_1), (a_2, b_2), \ldots$ in the collection $\{I_1, \ldots, I_N\}$ such that for each i, $a_i < b_{i-1} < b_i$. This procedure must terminate with the construction of (a_j, b_j) for some $j \leqslant N$. Then $b \in (a_j, b_j)$, so

$$\begin{aligned}
\sum_{n=1}^{N} |I_n| &\geqslant \sum_{i=1}^{j} (b_i - a_i) \\
&= b_j - (a_j - b_{j-1}) - (a_{j-1} - b_{j-2}) \\
&\quad - \cdots - (a_2 - b_1) - a_1 \\
&> b_j - a_1.
\end{aligned}$$

It follows that $\sum\limits_{n=1}^{\infty} |I_n| > b - a$ and therefore, since (I_n) was any sequence of bounded open intervals covering $[a, b]$, that $\mu^*([a, b]) \geqslant b - a$. Coupled with the reverse inequality already established, this proves that $\mu^*([a, b]) = b - a$. The proof for other types of interval is left as the next exercise. ■

Exercises (8.2)

.1 Complete the proof of Proposition 42 in the remaining cases.

.2 Let $\{I_1, \ldots, I_N\}$ be a finite set of bounded open intervals covering $\mathbb{Q} \cap [0, 1]$. Prove that $\sum\limits_{n=1}^{N} |I_n| \geqslant 1$. (Given $\varepsilon > 0$, extend each I_n, if necessary, to ensure that it has rational endpoints and that the total length of the intervals is increased by at most ε. Then argue as in the proof of Proposition 42.)

.3 Let X be a subset of $\mathbb{R}$ with finite outer measure. Prove that for each $\varepsilon > 0$ there exists an open set $A \supset X$ with finite outer measure, such that $\mu^*(A) < \mu^*(X) + \varepsilon$. (Use Exercise (8.1, 1).) Show that if X is also bounded, then we can choose A to be bounded.

.4 Let A, B be subsets of $\mathbb{R}$ with finite outer measure. Prove that $\mu^*(A \cup B) \leqslant \mu^*(A) + \mu^*(B)$.

Let X be a subset of $\mathbb{R}$, and $\mathcal{V}$ a family of non-degenerate intervals. We say that $\mathcal{V}$ is a **Vitali covering** of X if for each $\varepsilon > 0$ and each $x \in X$ there exists $I \in \mathcal{V}$ such that $x \in I$ and $|I| < \varepsilon$.The fundamental result about such coverings is the **Vitali covering theorem.**

Theorem 43 *Let $\mathcal{V}$ be a Vitali covering of a set $X \subset R$ with finite outer measure. Then for each $\varepsilon > 0$ there exists a finite set $\{I_1, \ldots, I_N\}$ of pairwise disjoint intervals in $\mathcal{V}$ such that*

$$\mu^*\left(X \backslash \bigcup_{n=1}^{N} I_n\right) < \varepsilon.$$

We postpone the proof of this very useful theorem until we have dealt with some auxiliary exercises.

Exercises (8.3)

.1 Let $\mathcal{V}$ be a Vitali covering of a subset X of $\mathbb{R}$, x a point of X, and A an open subset of $\mathbb{R}$ containing X. Show that for each $\varepsilon > 0$ there exists $I \in \mathcal{V}$ such that $x \in I$, $I \subset A$, and $|I| < \varepsilon$.

.2 Let $I_1, \ldots, I_N$ be finitely many closed intervals belonging to a Vitali covering $\mathcal{V}$ of a subset X of $\mathbb{R}$ with finite outer measure, and let $x \in X \backslash \bigcup_{n=1}^{N} I_n$. Show that for each $\varepsilon > 0$ there exists $I \in \mathcal{V}$ such that $x \in I$, $|I| < \varepsilon$, and I is disjoint from $\bigcup_{n=1}^{N} I_n$.

We are now able to give the proof of the Vitali covering theorem.

Proof. If necessary replacing the intervals in I by their closures, we may assume that $\mathcal{V}$ consists of closed intervals. Referring to Exercise (8.2, 2), choose an open set $A \supset X$ with finite outer measure. In view of Exercise (8.3,1), we may assume without loss of generality that

$$I \subset A \text{ for each } I \in \mathcal{V}. \tag{2}$$

Choosing any interval I_1 in the covering $\mathcal{V}$, we construct, inductively, an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ with $\lambda_1 = 0$, and a sequence $(I_n)_{n \geqslant 1}$ of pairwise disjoint intervals in $\mathcal{V}$, with the following properties for each $n \geqslant 2$:

(i) If $\lambda_n = 0$, then

$$|I_n| > \frac{1}{2} \sup \left\{ |I| : I \in \mathcal{V} \text{ and } I \cap \bigcup_{k=1}^{n-1} I_k = \varnothing \right\};$$

(ii) If $\lambda_n = 1$, then $I_n = I_{n-1}$ and $X \subset \bigcup_{k=1}^{n-1} I_k$.

Assume that we have constructed $\lambda_1, \ldots, \lambda_n$ and $I_1, \ldots, I_n$ with the applicable properties. If $\lambda_n = 1$, then we set $\lambda_{n+1} = 1$ and $I_{n+1} = I_n$. If, as we now assume, $\lambda_n = 0$, then either $X \subset \bigcup_{k=1}^{n} I_k$ and so $\mu^* \left(X \backslash \bigcup_{k=1}^{n} I_k \right) = 0$, in which case we set $\lambda_{n+1} = 1$ and $I_{n+1} = I_n$; or else $X \not\subset \bigcup_{k=1}^{n} I_k$. In the latter event, Exercise $(8.3, 2)$ shows that the set

$$S := \left\{ |I| : I \in \mathcal{V},\ I \cap \bigcup_{k=1}^{n} I_k = \varnothing \right\}$$

is nonempty. Since, by (2), S is bounded above by $\mu^*(A)$, it follows that $\sup S$ exists; moreover, as each $I \in \mathcal{V}$ is non-degenerate, $\sup S > 0$. To complete our inductive construction, we now choose $I_{n+1} \in \mathcal{V}$ such that $I_{n+1} \cap \bigcup_{k=1}^{n} I_k = \varnothing$ and $|I_{n+1}| > \frac{1}{2} \sup S$, and we set $\lambda_{n+1} = 0$.

If $\lambda_n = 1$ for some n, then we are through. So we may assume that $\lambda_n = 0$ for all n, and hence that the construction produces an infinite sequence $(I_n)_{n \geqslant 1}$ of pairwise disjoint elements of $\mathcal{V}$. Since the partial sums of the series $\sum_{n=1}^{\infty} |I_n|$ are bounded by $\mu^*(A)$, the series converges. Given $\varepsilon > 0$, we can therefore find N such that

$$\sum_{n=N+1}^{\infty} |I_n| < \frac{\varepsilon}{5}.$$

For each $n > N$ let x_n be the midpoint of I_n, and let J_n be the closed interval with midpoint x_n and length $5\,|I_n|$. It suffices to prove that

$$X \backslash \bigcup_{n=1}^{N} I_n \subset \bigcup_{n=N+1}^{\infty} J_n. \tag{3}$$

For then

$$\mu^* \left(X \backslash \bigcup_{n=1}^{N} I_n \right) \leqslant \sum_{n=N+1}^{\infty} |J_n| = 5 \sum_{n=N+1}^{\infty} |I_n| < \varepsilon.$$

To prove (3), consider any $x \in X \backslash \bigcup_{n=1}^{N} I_n$. By Exercise (8.3,2), there exists $J \in \mathcal{V}$ such that $x \in J$ and $J \cap \bigcup_{n=1}^{N} I_n = \varnothing$. We claim that $J \cap I_m$ is nonempty for some $m > N$. If this were not the case, then for each m we would have $J \cap \bigcup_{n=1}^{m} I_n = \varnothing$ and therefore

$$\begin{aligned} |J| &\leqslant \sup \left\{ |I| : I \in \mathcal{V},\ I \cap \bigcup_{k=1}^{m} I_k = \varnothing \right\} \\ &< 2 \left| I_{m+1} \right| \rightarrow 0 \text{ as } m \rightarrow \infty. \end{aligned}$$

(For the last step, recall that $\sum_{m=1}^{\infty} |I_m|$ is convergent.) It would then follow that $|J| = 0$, which is absurd as $\mathcal{V}$ contains only non-degenerate intervals. Thus

$$\nu = \min\{m > N : J \cap I_m \neq \varnothing\}$$

is well defined, $J \cap \bigcup_{n=1}^{\nu-1} I_n = \varnothing$, and therefore

$$|J| \leqslant \sup \left\{ |I| : I \in \mathcal{V},\ I \cap \bigcup_{k=1}^{\nu-1} I_k = \varnothing \right\} < 2 \left| I_\nu \right|.$$

Since $x \in J$ and $J \cap I_\nu \neq \varnothing$, we see that

$$|x - x_\nu| \leqslant |J| + \tfrac{1}{2} |I_\nu| < 2 |I_\nu| + \tfrac{1}{2} |I_\nu| = \tfrac{5}{2} |I_\nu|.$$

Hence $x \in J_\nu$. This establishes (3) and completes the proof. ∎

In the remainder of this section we apply the Vitali covering theorem in the proofs of some fundamental results in the theory of differentiation and integration.

Let I be an interval in $\mathbb{R}$. We say that a mapping $f : I \rightarrow \mathbb{R}$ is **absolutely continuous** if for each $\varepsilon > 0$ there exists $\delta > 0$ such that if $([a_k, b_k])_{k=1}^{n}$ is a finite family of non-overlapping compact subintervals of I such that $\sum_{k=1}^{n} (b_k - a_k) < \delta$, then $\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon$. On the other hand, if $a, b \in I$ and $a < b$, then we define the **variation** of f on the interval $[a, b]$ to be

$$T_f(a, b) = \sup \left\{ \sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)| : a = x_0 \leqslant x_1 \leqslant \cdots \leqslant x_n = b \right\},$$

if this quantity exists as a real number; in that case we say that f **has bounded variation** on $[a, b]$.

Exercises (8.4)

.1 Prove that an absolutely continuous function on I is both uniformly continuous and bounded.

.2 Let f, g be absolutely continuous functions on I. Prove that the functions $f+g$, $f-g$, λf (where $\lambda \in \mathbb{R}$), and fg are absolutely continuous, and that if $\inf\{|f(x)| : x \in I\} > 0$, then $1/f$ is absolutely continuous.

.3 Prove that if f is differentiable, with bounded derivative, on an interval I, then f is absolutely continuous.

.4 Let f have bounded variation on $[a, b]$. Prove that f is bounded on I; that

$$T_f(a, b) = T_f(a, x) + T_f(x, b)$$

for each $x \in [a, b]$; and that $T_f(a, \cdot)$ is an increasing function on I.

.5 Define $f : [0, 1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x^2} & \text{if } 0 < x \leqslant 1 \\ \\ 0 & \text{if } x = 0. \end{cases}$$

Prove that f is differentiable at each point of $[0, 1]$ but does not have bounded variation on that interval.

.6 Prove that $f : [a, b] \to \mathbb{R}$ has bounded variation if and only if it can be expressed as the difference of two increasing functions. (For "only if" note the last part of Exercise (8.4, 4).)

.7 Let f be absolutely continuous on a compact interval $I = [a, b]$. Prove that f has bounded variation in I, that the variation function $T_f(a, \cdot)$ is absolutely continuous on I, and that f is the difference of two absolutely continuous, increasing functions on I.

A simple corollary of the mean value theorem, one that suffices for many applications, states that if f is continuous on $[a, b]$ and $|f'(x)| \leqslant M$ for all $x \in (a, b)$, then $|f(b) - f(a)| \leqslant M(b - a)$. Our first application of the Vitali covering theorem generalises this corollary, and can be regarded an extension of the mean value theorem itself.

Proposition 44 *Let f be an absolutely continuous mapping of a compact interval $I = [a, b]$ into $\mathbb{R}$, and F a differentiable-almost-everywhere increasing mapping of I into $\mathbb{R}$ such that $|f'(x)| \leqslant F'(x)$ almost everywhere on I. Then*

$$|f(b) - f(a)| \leqslant F(b) - F(a). \tag{4}$$

Proof. Let $E \subset I$ be a set of measure zero such that $|f'(x)| \leqslant F'(x)$ for each $x \in X = I \backslash E$. We may assume without loss of generality that $a, b \in E$. Given $\varepsilon > 0$, choose $\delta > 0$ as in the definition of absolute continuity. For each $x \in X$ there exist arbitrarily small $r > 0$ such that $[x, x+r] \subset (a, b)$,

$$\begin{aligned} |f(x+r) - f(x) - f'(x)r| &< \varepsilon r, \\ |F(x+r) - F(x) - F'(x)r| &< \varepsilon r, \end{aligned}$$

and therefore

$$\begin{aligned} |f(x+r) - f(x)| &\leqslant |f'(x)|\, r + \varepsilon r \\ &\leqslant F'(x)r + \varepsilon r \\ &\leqslant F(x+r) - F(x) + 2\varepsilon r. \end{aligned}$$

The sets of the form $[x, x+r]$, for such $r > 0$, form a Vitali covering of X. By the Vitali covering theorem, there exists a finite, pairwise disjoint collection $([x_k, x_k + r_k])_{k=1}^N$ of sets of this type such that

$$\mu^* \left(X \backslash \bigcup_{k=1}^N [x_k, x_k + r_k] \right) < \delta.$$

We may assume that $x_k + r_k < x_{k+1}$ for $1 \leqslant k \leqslant N-1$. Thus

$$x_1 - a + \sum_{k=1}^{N-1} (x_{k+1} - x_k - r_k) + b - x_N - r_N < \delta,$$

and therefore

$$|f(x_1) - f(a)| + \sum_{k=1}^{N-1} |f(x_{k+1}) - f(x_k + r_k)| + |f(b) - f(x_N + r_N)| < \varepsilon.$$

It follows that

$$\begin{aligned} |f(b) - f(a)| \leqslant |f(x_1) - f(a)| &+ \sum_{k=1}^{N-1} |f(x_{k+1}) - f(x_k + r_k)| \\ &+ |f(b) - f(x_N + r_N)| + \sum_{k=1}^{N} |f(x_k + r_k) - f(x_k)| \end{aligned}$$

$$
\begin{aligned}
&< \varepsilon + \sum_{k=1}^{N} (F(x_k + r_k) - F(x_k) + 2\varepsilon r_k) \\
&\leqslant \varepsilon + F(x_1) - F(a) + \sum_{k=1}^{N-1} (F(x_{k+1}) - F(x_k + r_k)) \\
&\qquad\qquad + \sum_{k=1}^{N} (F(x_k + r_k) - F(x_k) + 2\varepsilon r_k) \\
&\quad + (F(b) - F(x_N + r_N)) \\
&= \varepsilon + F(b) - F(a) + 2\varepsilon \sum_{k=1}^{N} r_k \\
&< F(b) - F(a) + \varepsilon(1 + 2b - 2a).
\end{aligned}
$$

Since $\varepsilon > 0$ is arbitrary, we conclude that (4) holds. ■

Exercises (8.5)

.1 Let f be absolutely continuous on $I = [a, b]$, and suppose that for some constant M, $|f'| \leqslant M$ almost everywhere on I. Prove that $|f(b) - f(a)| \leqslant M(b - a)$.

.2 Let $f : [a, b] \to \mathbb{R}$ be an absolutely continuous function such that $f'(x) = 0$ almost everywhere on $I = [a, b]$. Give two proofs that f is a constant function. (For one proof use the Vitali covering theorem.)

.3 Let f, F be continuous on $I = [a, b]$, and suppose there exists a countable subset D of I such that $|f'(x)| \leqslant F'(x)$ for all $x \in I \backslash D$. Show that $|f(b) - f(a)| \leqslant F(b) - F(a)$. (We may assume that D is countably infinite. Let $d_1, d_2, \ldots$ be a one-one mapping of $\mathbb{N}^+$ onto D. Given $\varepsilon > 0$, let X be the set of all points $x \in I$ such that

$$
|f(\xi) - f(a)| \leqslant F(\xi) - F(a) + \varepsilon \left(\xi - a + \sum_{\{n : d_n < \xi\}} 2^{-n} \right)
$$

for all $\xi \in [a, x)$, and let $s = \sup X$. Assume that $s < b$, and derive a contradiction.)

.4 Let f be continuous on $I = [a, b]$, and suppose there exists a countable subset D of I such that $f'(x) = 0$ for all $x \in I \backslash D$. Prove that f is constant on I.

.5 Let C be the Cantor set. Show that $[0, 1] \backslash C$ is a countable union of non-overlapping open intervals $(J_n)_{n \geqslant 1}$ whose lengths sum to 1, and that C has measure zero.

For each $x = \sum\limits_{n=1}^{\infty} a_n 3^{-n} \in C$ define $F(x) = \sum\limits_{n=1}^{\infty} a_n 2^{-n-1}$. Show that

(i) if x has two ternary expansions, then they produce the same value for $F(x)$, so that F is a function on C;

(ii) F is a strictly increasing, continuous mapping of C onto $[0,1]$;

(iii) C is uncountable; and

(iv) F extends to an increasing continuous mapping that is constant on each J_n, equals 0 throughout $(-\infty, 0]$, and equals 1 throughout $[1, \infty)$.

Prove that for each $\delta > 0$ there exist finitely many points

$$a_1 < 0 < b_1 < a_2 < \cdots < b_{N-1} < a_N < 1 < b_N$$

of $[-1, 2]$ such that $C \subset \bigcup_{n=1}^{N} [a_n, b_n]$,

$$\sum_{n=1}^{N} (F(b_n) - F(a_n)) = 1,$$

and $\sum_{n=1}^{N} (b_n - a_n) < \delta$. (Thus F is increasing and continuous, but not absolutely continuous, on $[-1, 2]$.)

Finally, show that $F'(x) = 0$ for all $x \in [0,1] \backslash C$, but $F(1) > F(0)$.

The last two exercises deserve further comment. Consider a continuous function F on $[0,1]$ whose derivative exists and vanishes throughout $[0,1] \backslash E$. If E is countable, then Exercise (8.5, 4) shows that F is constant. On the other hand, Exercise (8.5, 5) shows that if E is uncountable and of measure zero, then F need not be constant; but if, in that case, F is absolutely continuous, then it follows from Exercise(8.5, 2) that it is constant.

Although the derivative of a function f may not exist at a point $x \in \mathbb{R}$, one or more of the following quantities—the **Dini derivates** of f at x—may:

$$\begin{aligned}
D^{+} f(x) &= \lim_{h \to 0^{+}} \sup \frac{f(x+h) - f(x)}{h}, \\
D_{+} f(x) &= \lim_{h \to 0^{+}} \inf \frac{f(x+h) - f(x)}{h}, \\
D^{-} f(x) &= \lim_{h \to 0^{-}} \sup \frac{f(x+h) - f(x)}{h}, \\
D_{-} f(x) &= \lim_{h \to 0^{-}} \inf \frac{f(x+h) - f(x)}{h}.
\end{aligned}$$

We consider $D^{+} f(x)$ to be undefined if

— either there is no $h > 0$ such that f is defined throughout the interval $[x, x+h]$

— or else $(f(x+h)-f(x))/h$ remains unbounded as $h \to 0^+$.

Similar comments apply to the other derivates of f.

Exercises (8.6)

.1 Prove that $D^+f(x) \geqslant D_+f(x)$ and $D^-f(x) \geqslant D_-f(x)$ whenever the quantities concerned make sense.

.2 Prove that f is differentiable on the right (respectively, left) at x if and only if $D^+f(x) = D_+f(x)$ (respectively, $D^-f(x) = D_-f(x)$).

.3 Let f be a mapping of $\mathbb{R}$ into $\mathbb{R}$, and define $g(x) = -f(-x)$. Prove that for each $x \in \mathbb{R}$, $D^+g(x) = D^-f(-x)$ and $D_-g(x) = D_+f(-x)$.

.4 Let $f : [a,b] \to \mathbb{R}$ be continuous, and suppose that one of the four derivates of f is nonnegative throughout (a,b). Prove that f is an increasing function on $[a,b]$. (Show that $x \rightsquigarrow f(x) + \varepsilon x$ is increasing for each $\varepsilon > 0$.)

.5 Consider a function $f : [a,b] \to \mathbb{R}$, and real numbers r, s with $r > s$. Define

$$E = \left\{x \in (a,b) : D^+f(x) > r > s > D_-f(x)\right\}.$$

Let X be an open set such that $E \subset X$ and $\mu^*(X) < \mu^*(E) + \varepsilon$ (see Exercise (8.1, 2)). Prove that the intervals of the form $(x-h, x)$ such that $x \in E$, $h > 0$, $[x-h, x] \subset X$, and $f(x) - f(x-h) < sh$ form a Vitali covering of E. Hence prove that for each $\varepsilon > 0$ there exist finitely many points $x_1, \ldots, x_m$ of E, and finitely many positive numbers $h_1, \ldots, h_m$, such that the intervals $J_i = (x_i - h_i, x_i)$ $(1 \leqslant i \leqslant m)$ form a pairwise disjoint collection,

$$\mu^*\left(\bigcup_{i=1}^m J_i\right) > \mu^*(E) - \varepsilon$$

and

$$\sum_{i=1}^m \left(f(x_i) - f(x_i - h_i)\right) < s\left(\mu^*(E) + \varepsilon\right).$$

Again applying the Vitali covering theorem, prove that there exist finitely many points $y_1, \ldots, y_n$ of $E \cap \bigcup_{i=1}^m J_i$, and finitely many positive numbers $h'_1, \ldots, h'_n$, such that

$$y_k + h'_k < y_{k+1} \quad (1 \leqslant k \leqslant n-1),$$

for each k there exists i such that $(y_k, y_k + h'_k) \subset J_i$, and

$$\sum_{k=1}^n \left(f(y_k + h'_k) - f(y_k)\right) > r\left(\mu^*(E) - 2\varepsilon\right).$$

Our next theorem shows, in particular, that the differentiability of the function F can be dropped from the hypotheses of Proposition 44.

Theorem 45 *An increasing function $F : \mathbb{R} \to \mathbb{R}$ is differentiable (and has nonnegative derivative) almost everywhere. Moreover,*

(i) *F' is measurable, and*

(ii) *whenever $a < b$, F' is summable over $[a,b]$ and satisfies*

$$\int_a^b F' \leqslant F(b) - F(a).$$

Proof. To prove that F is differentiable a.e., it suffices to show that the sets

$$\begin{aligned} S &= \left\{x \in \mathbb{R} : D^+F(x) \text{ is undefined}\right\}, \\ T &= \left\{x \in \mathbb{R} : D^+F(x) > D_-F(x)\right\} \end{aligned}$$

have measure zero. For, applying this and Exercise $(8.6, 3)$ to the increasing function $x \rightsquigarrow -F(-x)$, we then see that $D^-F(x) \leqslant D_+F(x)$ almost everywhere; whence, by Exercise $(8.6, 1)$

$$D^+F(x) \leqslant D_-F(x) \leqslant D^-F(x) \leqslant D_+F(x) \leqslant D^+F(x) \in \mathbb{R}$$

almost everywhere. (Note that as F is increasing, $D_+F(x)$ and $D_-F(x)$ are everywhere defined and nonnegative.) Thus the four Dini derivates of F are equal almost everywhere. Reference to Exercise $(8.6, 2)$ then completes the proof.

Leaving S to the next set of exercises, we now show that T has measure zero. Since T is the union of a countable family of sets of the form

$$E = \left\{x \in (a,b) : D^+F(x) > r > s > D_-F(x)\right\},$$

where $a < b$ and r, s are rational numbers with $r > s$, it is enough to prove that such a set E has measure zero. Accordingly, fix $\varepsilon > 0$ and use Exercise $(8.6, 5)$ to obtain

(i) finitely many points $x_1, \ldots, x_m$ of (a,b), and finitely many positive numbers $h_1, \ldots, h_m$, such that the intervals $J_i = (x_i - h_i, x_i)$ $(1 \leqslant i \leqslant m)$ form a pairwise disjoint collection,

$$\mu^*\left(\bigcup_{i=1}^m J_i\right) > \mu^*(E) - \varepsilon,$$

and

$$\sum_{i=1}^m \left(F(x_i) - F(x_i - h_i)\right) < s\left(\mu^*(E) + \varepsilon\right);$$

(ii) finitely many points $y_1, \ldots, y_n$ of $E \cap \bigcup_{i=1}^{m} J_i$, and finitely many positive numbers $h'_1, \ldots, h'_n$, such that

$$y_k + h'_k < y_{k+1} \quad (1 \leqslant k \leqslant n-1)\,, \tag{5}$$

for each k there exists i with $(y_k, y_k + h'_k) \subset J_i$, and

$$\sum_{k=1}^{n} (F(y_k + h'_k) - F(y_k)) > r\,(\mu^*(E) - 2\varepsilon)\,.$$

For each i with $1 \leqslant i \leqslant m$ let

$$S_i = \{k : (y_k, y_k + h'_k) \subset J_i\}\,.$$

Since F is increasing, it follows from (5) that

$$\sum_{k \in S_i} (F(y_k + h'_k) - F(y_k)) \leqslant F(x_i) - F(x_i - h_i).$$

Thus, as the intervals J_i are disjoint,

$$\sum_{i=1}^{m} (F(x_i) - F(x_i - h_i)) \geqslant \sum_{k=1}^{n} (F(y_k + h'_k) - F(y_k))\,,$$

so that

$$s\,(\mu^*(E) + \varepsilon) > r\,(\mu^*(E) - 2\varepsilon)\,.$$

Since $\varepsilon > 0$ is arbitrary, it follows that $s\mu^*(E) \geqslant r\mu^*(E)$. But $r > s$, so we must have $\mu^*(E) = 0$.

For each n write

$$F_n\,(x) = F\left(x + \frac{1}{n}\right) \quad (x \in \mathbb{R})$$

and

$$f_n = n\,(F_n - F)\,.$$

Note that, being an increasing function, F is Riemann integrable, and therefore summable, over any proper compact interval $[a, b]$ in $\mathbb{R}$; whence F, F_n and f_n are summable over $[a, b]$. Let $a \leqslant \xi < b$, and choose a positive integer N such that $\xi + \frac{1}{N} < b$. For all $n \geqslant N$ we have

$$\begin{aligned}\int_a^\xi f_n &= n\int_a^\xi F_n - n\int_a^\xi F \\ &= n\int_{a+\frac{1}{n}}^{\xi+\frac{1}{n}} F - n\int_a^\xi F \\ &= n\int_\xi^{\xi+\frac{1}{n}} F - n\int_a^{a+\frac{1}{n}} F,\end{aligned}$$

where we have used Exercise $(2.3, 4)$ to produce the second-last line. Now, F is increasing and $\xi + \frac{1}{n} < b$, so

$$n \int_{\xi}^{\xi+\frac{1}{n}} F \leqslant n \int_{\xi}^{\xi+\frac{1}{n}} F(b) = F(b).$$

On the other hand,

$$n \int_{a}^{a+\frac{1}{n}} F \geqslant n \int_{a}^{a+\frac{1}{n}} F(a) = F(a).$$

It follows that

$$\int_{a}^{\xi} f_n \leqslant F(b) - F(a) \qquad (n \geqslant N).$$

Next we note that since $f_n(x) \to F'(x)$ almost everywhere, F' is measurable. Also, $f_n \geqslant 0$, so Fatou's lemma (Exercise $(2.6, 5)$) can be applied, to show that F' is summable over $[a, \xi]$ and

$$\int_{a}^{\xi} F' \leqslant \liminf \int_{a}^{\xi} f_n \leqslant F(b) - F(a).$$

Since

$$\left(F'\chi_{\left[a, b-\frac{1}{k}\right]}\right)(x) \to \left(F'\chi_{[a,b]}\right)(x) \text{ as } k \to \infty$$

for all $x \in [a, b]$, and, as is easily confirmed, $\left(F'\chi_{\left[a, b-\frac{1}{k}\right]}\right)_{k \geqslant 1}$ is a Riesz sequence of summable functions, we now see from Beppo Levi's Theorem that

$$\int_{a}^{b} F' = \lim_{n\to\infty} \int F'\chi_{\left[a, b-\frac{1}{n}\right]} = \lim_{n\to\infty} \int_{a}^{b-\frac{1}{n}} F' \leqslant F(b) - F(a),$$

as we wanted. ■

Fubini's series theorem, our next result, provides a good application of Theorem 45.

Theorem 46 *Let $(F_n)_{n\geqslant 1}$ be a sequence of increasing continuous functions on $\mathbb{R}$ such that $F(x) = \sum_{n=1}^{\infty} F_n(x)$ converges for all $x \in \mathbb{R}$. Then almost everywhere, F is differentiable, $\sum_{n=1}^{\infty} F_n'(x)$ converges, and $F'(x) = \sum_{n=1}^{\infty} F_n'(x)$.*

Proof. Fix real numbers a, b with $a < b$. It suffices to prove that $F'(x) = \sum_{n=1}^{\infty} F_n'(x)$ almost everywhere on $I = [a, b]$: for then we can apply the result to the intervals $[-n, n]$ as n increases through $\mathbb{N}^+$. If necessary replacing F_n by $F_n - F_n(a)$, we may assume that $F_n(a) = 0$. Write

$$s_n(x) = F_1(x) + \cdots + F_n(x) \quad (x \in I)$$

and note that $F - s_n = \sum_{k=n+1}^{\infty} F_k$ is increasing and nonnegative on I. By Theorem 45, s_n is differentiable on $I \backslash A_n$ for some set A_n of measure zero; likewise, F (which is clearly increasing) is differentiable on $I \backslash A_0$ for some set A_0 of measure zero. Then

$$A = \bigcup_{n=0}^{\infty} A_n$$

has measure zero. Since both $F - s_{n+1}$ and $s_{n+1} - s_n$ are increasing functions, for each $x \in I \backslash A$ we have

$$s_n'(x) \leqslant s_{n+1}'(x) \leqslant F'(x). \tag{6}$$

It follows from the monotone sequence principle that $\sum\limits_{n=1}^{\infty} F_n'(x)$ converges to a sum $\leqslant F'(x)$.

Now choose an increasing sequence $(n_k)_{k \geqslant 1}$ of positive integers such that for each k,

$$0 \leqslant F(b) - s_{n_k}(b) \leqslant 2^{-k}.$$

Since $F - s_{n_k}$ is an increasing function, for each $x \in I$ we obtain the inequalities

$$0 \leqslant F(x) - s_{n_k}(x) \leqslant 2^{-k}.$$

Hence $\sum\limits_{k=1}^{\infty} (F(x) - s_{n_k}(x))$ converges, by comparison with $\sum\limits_{k=1}^{\infty} 2^{-k}$. Applying the first part of the proof with F_k replaced by $F - s_{n_k}$, we now see that, almost everywhere on I, $\sum\limits_{k=1}^{\infty} \left(F'(x) - s_{n_k}'(x)\right)$ converges and therefore

$$\lim_{k \to \infty} \left(F'(x) - s_{n_k}'(x)\right) = 0.$$

It follows from (6) that

$$F'(x) = \lim_{n \to \infty} s_n(x) = \sum_{n=1}^{\infty} F_n'(x)$$

almost everywhere on I. ■

Fubini's series theorem is a crucial tool in a development of the Lebesgue integral as an antiderivative, a development due to Riesz. for details of this, see [9] or [3].

Exercises (8.7)

.1 Let f be an increasing function on $[a, b]$, and for each positive integer n define

$$S_n = \left\{x \in (a, b) : D^+ f(x) > n\right\}.$$

Prove that

$$\mu^*(I\backslash S_n) < n^{-1}\left(f(b) - f(a)\right)$$

and hence that the set of those $x \in (a,b)$ at which $D^+f(x)$ is undefined has measure zero. (Use the Vitali covering theorem to show that there exist finitely many points $x_1, x_2, \ldots, x_m$ of (a,b), and positive numbers $h_1, h_2, \ldots, h_m$, such that $x_k + h_k < x_{k+1}$ and $f(x_k + h_k) - f(x_k) > nh_k$.)

.2 Let E be a bounded subset of $\mathbb{R}$ that has measure zero, and let a be a lower bound for E. For each positive integer n choose a bounded open set $A_n \supset E$ such that $\mu^*(A_n) < 2^{-n}$, and define

$$f_n(x) = \begin{cases} 0 & \text{if } x < a \\ \\ \mu^*(A_n \cap [a,x]) & \text{if } x \geqslant a. \end{cases}$$

Show that

(i) $f = \sum_{n=1}^{\infty} f_n$ is an increasing continuous function on $\mathbb{R}$;

(ii) $D^+f(x)$ is undefined for each $x \in E$.

.3 Let f be a bounded function that is continuous almost everywhere on a compact interval I. Let M be a bound for $|f|$ on I, let $E \subset I = [a,b]$ be a set of measure zero such that f is continuous on $X = I\backslash E$, and let $\varepsilon > 0$. We may assume that $a, b \in E$. For each $x \in X$ there exist arbitrarily small $r > 0$ such that $[x, x+r] \subset I$ and

$$|f(x') - f(x'')| < \frac{\varepsilon}{2(b-a)} \quad (x \leqslant x' \leqslant x'' \leqslant x + r).$$

The sets $[x, x+r]$ of this type form a Vitali cover of X. With the aid of the Vitali covering theorem, construct a partition P of I such that $U(P,f) - L(P,f) < \varepsilon$. This provides an alternative proof of one half of Theorem 23.

9 The Fundamental Theorem of Calculus

Our aim in this final section is to prove the following version of the **fundamental theorem of calculus.**

Theorem 47 *Let f be summable, and define $F : \mathbb{R} \to \mathbb{R}$ by*

$$F(x) = \int_{-\infty}^{x} f.$$

Then $F'(x) = f(x)$ almost everywhere.

In order to prove this theorem, we first prove it in the case where f is bounded.

Lemma 48 *Let f and F be as in the statement of* Theorem 47*, and suppose that f is bounded on $\mathbb{R}$. Then $F'(x) = f(x)$ almost everywhere.*

Proof. Let c be a bound for $|f|$ on $\mathbb{R}$. We may assume that $f \geqslant 0$. Then F is an increasing function on $\mathbb{R}$. It follows from Theorem 45 that F is differentiable almost everywhere. For each positive integer n and each $x \in \mathbb{R}$ set

$$F_n(x) = F\left(x + \frac{1}{n}\right)$$

and

$$f_n(x) = n\left(F_n\left(x\right) - F(x)\right) = n \int_x^{x+\frac{1}{n}} f.$$

Note that

$$f_n(x) \leqslant n \int_x^{x+\frac{1}{n}} c = c$$

for all x in the domain of f. Now, by Exercise $(2.7, 5)$, F is continuous; so both F_n and f_n are continuous and therefore, by Exercise $(5.1, 3)$, measurable. For all a, b with $a < b$, since $f_n(x) \to F'(x)$ almost everywhere on $[a, b]$, we see from Lebesgue's dominated convergence theorem that $F'\chi_{[a,b]}$ is summable and that for each $x \in [a, b]$,

$$\begin{aligned}
\int_a^x F' &= \lim_{n\to\infty} \int_a^x f_n \\
&= \lim_{n\to\infty} n \int_a^x \left(t \rightsquigarrow F\left(t + \frac{1}{n}\right) - F(t)\right) \\
&= \lim_{n\to\infty} n \left(\int_{-\infty}^{x+\frac{1}{n}} F - \int_{-\infty}^{x} F\right) \\
&= \lim_{n\to\infty} n \int_x^{x+\frac{1}{n}} F \\
&= F(x),
\end{aligned}$$

where the second last step uses translation invariance (Exercise $(2.3, 4)$) and the final line follows from Exercise $(4.1, 3)$. Hence

$$\int_a^x \left(F' - f\right) = 0 \quad \left(x \in [a, b]\right).$$

It follows from Exercise $(2.7, 6)$ that $F'(x) = f(x)$ almost everywhere on $[a, b]$. A by-now-standard argument shows that $F'(x) = f(x)$ almost everywhere, and completes the proof. ∎

We now give the proof of the fundamental theorem of calculus.

Proof. Under the hypotheses of Theorem 47, we may assume without loss of generality that $f \geqslant 0$. For each positive integer n and each $x \in \mathbb{R}$ define

$$F_n(x) = \int_{-\infty}^{x} (f - \min\{f, n\}).$$

Since $f - \min\{f, n\} \geqslant 0$, we see that F_n is an increasing function on $\mathbb{R}$ and so, by Theorem 45, has a nonnegative derivative almost everywhere. On the other hand, since $0 \leqslant \min\{f(x), n\} \leqslant n$ almost everywhere on $\mathbb{R}$, the preceding lemma shows that

$$\frac{\mathrm{d}}{\mathrm{d}x} \int_{\infty}^{x} \min\{f, n\} = \min\{f(x), n\} \quad \text{a.e.}$$

Hence

$$F'(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_n + \frac{\mathrm{d}}{\mathrm{d}x} \int_{-\infty}^{x} \min\{f, n\} \geqslant \min\{f(x), n\} \quad \text{a.e.}$$

Letting $n \to \infty$, and recalling that a countable union of null sets is null, we conclude that $F'(x) \geqslant f(x)$ almost everywhere. In view of this and Theorem 45, we see that for each positive integer k,

$$\int_{-k}^{k} F' \geqslant \int_{-k}^{k} f = F(k) - F(-k) \geqslant \int_{-k}^{k} F'.$$

Thus

$$\int (F' - f)\, \chi_{[-k,k]} = 0.$$

Since $(F' - f)\, \chi_{[-k,k]]} \geqslant 0$ almost everywhere, it follows that $(F' - f)\, \chi_{[-k,k]} = 0$, and therefore $F' \chi_{[-k,k]} = f \chi_{[-k,k]}$, almost everywhere. Since the union of a sequence of null sets is null, we conclude that $F' = f$ almost everywhere. ■

Under what conditions is a given function $F : [a, b] \to \mathbb{R}$ an indefinite integral—that is, when does there exist a summable function f such that $F(x) = \int_a^x f$ a.e. on $[a, b]$?

Proposition 49 *Let f be a summable function, and let $a < b$. Then the indefinite integral*

$$F(x) = \int_a^x f$$

defines an absolutely continuous function on $[a, b]$.

Proof. Let $\varepsilon > 0$. By Exercise (5.3, 4), there exists $\delta > 0$ such that if A is an integrable set with $\mu(A) < \delta$, then $\int_A |f| < \varepsilon$. Let $([a_k, b_k])_{k=1}^{n}$ be a finite collection of non-overlapping compact subintervals of $[a,b]$ such that $\sum\limits_{k=1}^{n}(b_k-a_k)<\delta$, and let $A=\bigcup\limits_{k=1}^{n}[a_k,b_k]$. Then $\mu(A)<\delta$, so

$$\begin{aligned}\sum_{k=1}^{n}|F(b_k)-F(a_k)| &= \sum_{k=1}^{n}\left|\int_{a_k}^{b_k} f\right| \\ &\leqslant \sum_{k=1}^{n}\int_{a_k}^{b_k}|f| = \int_A |f| < \varepsilon.\end{aligned}$$

■

Conversely, we have

Proposition 50 *Let F be an absolutely continuous mapping on a compact interval $I=[a,b]$. Then there exists a summable function f such that*

$$F(x)-F(a)=\int_a^x f \qquad (x\in I).$$

Proof. Being absolutely continuous, F has bounded variation, by Exercise (8.4, 7); so, by Exercise (8.4, 6), there exist increasing functions F_1,F_2 such that $F=F_1-F_2$. Referring to Theorem 45, we see that F is differentiable almost everywhere, F' is summable over $[a,b]$, $|F'(x)|\leqslant F_1'(x)+F_2'(x)$ almost everywhere;,and

$$\int_a^b |F'|\leqslant F_1(b)+F_2(b)-F_1(a)-F_2(a).$$

Let

$$G(x)=\int_a^x F' \qquad (a\leqslant x\leqslant b).$$

Then G is absolutely continuous, by Proposition 49, as is the function $F-G$. By the fundamental theorem of calculus, $F'(x)-G'(x)=0$ almost everywhere on $[a,b]$. It follows from Exercise (8.5, 2) that $F-G$ is constant on $[a,b]$; whence

$$F(x)-F(a)=G(x)-G(a)=\int_a^x F',$$

which reduces to the desired property of F if we take $f=F'$. ■

The function $F:[0,1]\to\mathbb{R}$ discussed in Exercise (8.5, 5) is shown there to be neither absolutely continuous nor (inevitably, in view of the preceding theorem) an indefinite integral. Moreover, according to that exercise, $F'(x)=0$ almost everywhere on $[0,1]$ and

$$\int_0^1 F'=0<F(1)-F(0).$$

Appendix: The axiom of choice

In the early years of this century it was recognised that the following principle, the **axiom of choice**, was necessary for the proofs of several important theorems in mathematics.

AC *If $\mathcal{F}$ is a nonempty family of pairwise disjoint nonempty sets, then there exists a set that intersects each member of $\mathcal{F}$ in exactly one element.*

In particular, Zermelo used this axiom explicitly in his proof that every set S can be well ordered—that is, there is a total partial order $\geqslant$ on S with respect to which every nonempty subset of X has a least element [13]. It was shown by Gödel [5] in 1939 that the axiom of choice is consistent with the axioms of Zermelo–Fraenkel set theory (ZF), in the sense that the axiom can be added to ZF without leading to a contradiction, and by Cohen [4] in 1963 that the negation of the axiom of choice is also consistent with ZF. Thus the axiom of choice is *independent* of ZF: it can be neither proved nor disproved without adding some extra principles to ZF.

The axiom of choice is commonly used in an equivalent form (the one we used in the construction of a non-measurable set):

AC′ *If A and B are nonempty sets, $S \subset A \times B$, and for each $x \in A$ there exists $y \in B$ such that $(x, y) \in S$, then there exists a function $f : A \to B$ —called a* choice function *for S —such that $(x, f(x)) \in S$ for each $x \in A$.*

To prove the equivalence of these two forms of the axiom of choice, first assume that the original version AC of the axiom holds, and consider nonempty sets A, B and a subset S of $A \times B$ such that for each $x \in A$ there exists $y \in B$ with $(x, y) \in S$. For each $x \in A$ let

$$F_x = \{x\} \times \{y \in B : (x, y) \in S\}.$$

Then $\mathcal{F} = (F_x)_{x \in A}$ is a nonempty family of pairwise disjoint sets, so, by AC, there exists a set C that has exactly one element in common with each F_x. We now define the required choice function $f : A \to B$ by setting

$$(x, f(x)) = \text{ the unique element of } C \cap F_x$$

for each $x \in A$.

Now assume that the alternative form AC′ of the axiom of choice holds, and consider a nonempty family $\mathcal{F}$ of pairwise disjoint nonempty sets. Taking

$$\begin{aligned} A &= \mathcal{F}, \\ B &= \bigcup_{X \in \mathcal{F}} X, \\ S &= \{(X, x) : X \in \mathcal{F},\, x \in X\} \end{aligned}$$

in AC$'$, we obtain a function

$$f : \mathcal{F} \to \bigcup_{X \in \mathcal{F}} X$$

such that $f(X) \in X$ for each $X \in \mathcal{F}$. The range of f is then a set that has exactly one element in common with each member of $\mathcal{F}$.

There are two other choice principles that are widely used in analysis. The first of these, the **principle of countable choice,** is the case $A = \mathbb{N}$ of AC$'$. The second is the **principle of dependent choice:**

> *If $a \in A$, $S \subset A \times A$, and for each $x \in A$ there exists $y \in A$ such that $(x, y) \in S$, then there exists a sequence $(a_n)_{n \geqslant 1}$ in A such that $a_1 = a$ and $(a_n, a_{n+1}) \in S$ for each n.*

It is a good exercise to show that the axiom of choice entails the principle of dependent choice, and that the principle of dependent choice entails the principle of countable choice. Since the last two principles can be derived as consequences of the axioms of ZF, they are definitely weaker than the axiom of choice.

For a fuller discussion of axioms of choice and related matters, see the article by Jech on pages 345–370 of [2].

References

[1] E. Asplund and L. Bungart, *A First Course in Integration*, Holt, Rinehart & Winston, New York, 1966.

[2] J. Barwise, *Handbook of Mathematical Logic,* North-Holland, Amsterdam, 1977.

[3] D.S. Bridges, *Foundations of Real and Abstract Analysis*, Graduate Texts in Mathematics **174**, Springer-Verlag, Heidelberg, 1998.

[4] P.J. Cohen, *Set Theory and the Continuum Hypothesis,* W.A. Benjamin, Inc., New York, 1966.

[5] K. Gödel, *The Consistency of the axiom of choice and the Generalized Continuum Hypothesis with the Axioms of Set Theory,* Annals of Mathematics Studies **3**, Princeton University Press, Princeton, NJ, 1940.

[6] P.R. Halmos, *Measure Theory*, Graduate Texts in Mathematics **18**, Springer-Verlag, Heidelberg, 1974.

[7] H. Lebesgue, 'Intégrale, longueur, aire', Annali di Mat. (3) **7**, 231–259, 1902.

[8] H. Lebesgue, *Leçons sur l'intégration et la recherche des fonctions primitives,* Gauthier-Villars, Paris, 1904.

[9] F. Riesz, 'Sur l'intégrale de Lebesgue comme l'opération inverse de la dérivation', Ann. Scuola Norm. Sup. Pisa (2)**5**, 191–212, 1936.

[10] F. Riesz and B. Sz.-Nagý, *Functional Analysis*, Frederick Ungar, New York, 1955; reprinted in paperback by Dover Publications Inc., New York, 1990.

[11] H.L. Royden, *Real Analysis* (3rd edn), Macmillan Publ. Co., New York, 1988.

[12] Walter Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1970.

[13] E. Zermelo, 'Beweis, dass jede Menge wohlgeordnet werden kann', Math. Annalen **59**, 514–516, 1904.