Analysis III

Ben Green

Mathematical Institute, Oxford
E-mail address: ben.green@maths.ox.ac.uk

2000 *Mathematics Subject Classification.* Primary

Contents

Preface

The objective of this course is to present a rigorous theory of what it means to integrate a function $f : [a, b] \to \mathbb{R}$. For which functions f can we do this, and what properties does the integral have? Can we give rigorous and general versions of facts you learned in school, such as integration by parts, integration by substitution, and the fact that the integral of f' is just f?

We will present the theory of the *Riemann integral*, although the way we will develop it is much closer to what is known as the *Darboux integral.* The end product is the same (the Riemann integral and the Darboux integral are equivalent) but the Darboux development tends to be easier to understand and handle.

This is not the only way to define the integral. In fact, it has certain deficiencies when it comes to the interplay between integration and limits, for example. To handle these situations one needs the *Lebesgue integral*, which is discussed in a future course.

Students should be aware that every time we write "integrable" we mean "Riemann integrable". For example, later on we will exhibit a non-integrable function, but it turns out that this function is integrable in the sense of Lebesgue.

CHAPTER 1

Step functions and the Riemann integral

1.1. Step functions

We are going to define the (Riemann) integral of a function by approximating it using simple functions called step functions.

DEFINITION 1.1. Let $[a, b]$ be an interval. A function $\phi : [a, b] \to \mathbb{R}$ is called a step function if there is a finite sequence $a = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_n = b$ such that ϕ is constant on each open interval (x_{i-1}, x_i).

Remarks. We do not care about the values of f at the endpoints $x_0, x_1, \ldots, x_n$.

We call a sequence $a = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_n = b$ a partition $\mathcal{P}$, and we say that ϕ is a step function adapted to $\mathcal{P}$.

DEFINITION 1.2. A partition $\mathcal{P}'$ given by $a = x'_0 \leqslant \ldots \leqslant x'_{n'} \leqslant b$ is *refinement* of $\mathcal{P}$ if every x_i is an x'_j for some j.

LEMMA 1.1. *We have the following facts about partitions:*

(i) *Suppose that ϕ is a step function adapted to $\mathcal{P}$, and if $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then ϕ is also a step function adapted to $\mathcal{P}'$.*

(ii) *If $\mathcal{P}_1, \mathcal{P}_2$ are two partitions then there is a common refinement of both of them.*

(iii) *If ϕ_1, ϕ_2 are step functions then so are $\max(\phi_1, \phi_2)$, $\phi_1 + \phi_2$ and $\lambda\phi_i$ for any scalar λ.*

Proof. All completely straightforward; for (iii), suppose that ϕ_1 is adapted to $\mathcal{P}_1$ and that ϕ_2 is adapted to $\mathcal{P}_2$, and pass to a common refinement of $\mathcal{P}_1, \mathcal{P}_2$. □

If $X \subset \mathbb{R}$ is a set, the *indicator function* of X is the function $\mathbf{1}_X$ taking the value 1 for $x \in X$ and 0 elsewhere.

LEMMA 1.2. *A function $\phi : [a, b] \to \mathbb{R}$ is a step function if and only if it is a finite linear combination of indicator functions of intervals (open and closed).*

Proof. Suppose first that ϕ is a step function adapted to some partition $\mathcal{P}$, $a = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_n = b$. Then ϕ can be written as a weighted sum of the functions $\mathbf{1}_{(x_{i-1}, x_i)}$ (each an indicator function of an open interval) and the functions $\mathbf{1}_{\{x_i\}}$ (each an indicator function of a closed interval containing a single point).

Conversely, the indicator function of any interval is a step function, and hence so is any finite linear combination of these by Lemma 1.1. □

In particular, the step functions on $[a, b]$ form a vector space, which we occasionally denote by $\mathscr{L}_{\text{step}}[a, b]$.

1.2. I of a step function

It is obvious what the integral of a step function "should" be.

Definition 1.3. Let ϕ be a step function adapted to some partition $\mathcal{P}$, and suppose that $\phi(x) = c_i$ on the interval (x_{i-1}, x_i). Then we define

$$I(\phi) = \sum_{i=1}^{n} c_i(x_i - x_{i-1}).$$

We call this $I(\phi)$ rather than $\int_a^b \phi$, because we are going to define $\int_a^b f$ for a class of functions f much more general than step functions. It will then be a theorem that $I(\phi) = \int_a^b \phi$, rather than simply a definition.

Actually, there is a small subtlety to the definition. Our notation suggests that $I(\phi)$ depends only on ϕ, but its definition depended also on the partition $\mathcal{P}$. In fact, it does not matter which partition one chooses. If one is pedantic and writes

$$I(\phi; \mathcal{P}) = \sum_{i=1}^{n} c_i(x_i - x_{i-1})$$

then one may easily check that

$$I(\phi; \mathcal{P}) = I(\phi, \mathcal{P}')$$

for any refinement $\mathcal{P}'$ of $\mathcal{P}$. Now if ϕ is a step function adapted to both $\mathcal{P}_1$ and $\mathcal{P}_2$ then one may locate a common refinement $\mathcal{P}'$ and conclude that

$$I(\phi, \mathcal{P}_1) = I(\phi; \mathcal{P}') = I(\phi, \mathcal{P}_2).$$

Lemma 1.3. *The map* $I : \mathscr{L}_{\text{step}}[a, b] \to \mathbb{R}$ *is linear:* $I(\lambda\phi_1 + \mu\phi_2) = \lambda I(\phi_1) + \mu I(\phi_2)$.

Proof. This is obvious on passing to a common refinement of the partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ to which ϕ_1, ϕ_2 are adapted. □

1.3. Definition of the integral

Let $f : [a, b] \to \mathbb{R}$ be a bounded function. We say that a step function ϕ_- is a *minorant* for f if $f \geqslant \phi_-$ pointwise. We say that a step function ϕ^+ is a *majorant* for f if $f \leqslant \phi^+$ pointwise.

DEFINITION 1.4. A function f is integrable if

$$\sup_{\phi_-} I(\phi_-) = \inf_{\phi_+} I(\phi_+), \tag{1.1}$$

where the sup is over all minorants $\phi_- \leqslant f$, and the inf is over all majorants $\phi_+ \geqslant f$. These minorants and majorants are always assumed to be step functions. We define the integral $\int_a^b f$ to be the common value of the two quantities in (1.1).

We note that the sup and inf exist for any bounded function f. Indeed if $|f| \leqslant M$ then the constant function $\phi_- = -M$ is a minorant for f (so there is at least one) and evidently $I(\phi_-) \leqslant (b - a)M$ for all minorants. A similar proof applies to majorants.

We note moreover that, for any function f,

$$\sup_{\phi_-} I(\phi_-) \leqslant \inf_{\phi_+} I(\phi_+). \tag{1.2}$$

To see this, let $\phi_- \leqslant f \leqslant \phi_+$ be majorants, adapted to partitions $\mathcal{P}_-$ and $\mathcal{P}_+$ respectively. By passing to a common refinement we may assume that $\mathcal{P}_- = \mathcal{P}_+ = \mathcal{P}$. Then it is clear from the definition of $I(.)$ that $I(\phi_-) \leqslant I(\phi_+)$. Since ϕ_-, ϕ_+ were arbitrary, (1.2) follows.

It follows from (1.2) that if f is integrable then

$$I(\phi_-) \leqslant \int_a^b f \leqslant I(\phi_+) \tag{1.3}$$

whenever $\phi_- \leqslant f \leqslant \phi_+$ are a minorants/majorant.

Remark. If a function f is only defined on an open interval (a, b), then we say that it is integrable if an arbitrary extension of it to $[a, b]$ is. It follows immediately from the definition of step function (which does not care about the endpoints) that it does not matter which extension we choose.

1.4. Very basic theorems about the integral

In this section we assemble some basic facts about the integral. Their proofs are all essentially routine, but there are some labour-saving tricks to be exploited.

PROPOSITION 1.1. *Suppose that f is integrable on $[a, b]$. Then, for any c with $a < c < b$, f is Riemann integrable on $[a, c]$ and on $[c, b]$. Moreover $\int_a^b f = \int_c^b f + \int_a^c f$.*

Proof. Let M be a bound for f, thus $|f(x)| \leqslant M$ everywhere. In this proof it is convenient to assume that (i) all partitions of $[a,b]$ include the point c and that (ii) all minorants take the value $-M$ at c, and all majorants the value M. By refining partitions if necessary, this makes no difference to any computations involving $I(\phi_-), I(\phi_+)$.

Now observe that a minorant ϕ_- of f on $[a,b]$ is precisely the same thing as a minorant $\phi_-^{(1)}$ of f on $[a,c]$ juxtaposed with a minorant $\phi_-^{(2)}$ of f on $[c,b]$, and that $I(\phi_-) = I(\phi_-^{(1)}) + I(\phi_-^{(2)})$. A similar comment applies to majorants. Thus, since f is integrable,

$$\sup_{\phi_-} I(\phi_-) = \sup_{\phi_-^{(1)}} I(\phi_-^{(1)}) + \sup_{\phi_-^{(2)}} I(\phi_-^{(2)}) = \inf_{\phi_+^{(1)}} I(\phi_+^{(1)}) + \inf_{\phi_+^{(2)}} I(\phi_+^{(2)}) = \inf_{\phi_+} I(\phi_+). \tag{1.4}$$

Since $\sup_{\phi_-^{(i)}} I(\phi_-^{(i)}) \leqslant \inf_{\phi_+^{(i)}} I(\phi_+^{(i)})$ for $i = 1,2$, we are forced to conclude that equality holds: $\sup_{\phi_-^{(i)}} I(\phi_-^{(i)}) = \inf_{\phi_+^{(i)}} I(\phi_+^{(i)})$ for $i = 1,2$. Thus f is indeed integrable on $[a,c]$ and on $[c,b]$, and it follows from (1.4) that $\int_a^b f = \int_a^c f + \int_c^b f$. □

Corollary 1.1. *Suppose that $f : [a,b] \to \mathbb{R}$ is integrable, and that $[c,d] \subset [a,b]$. Then f is integrable on $[c,d]$.*

Proof. On the example sheet. □

Proposition 1.2. *If f, g are integrable on $[a,b]$ then so is $\lambda f + \mu g$ for any $\lambda, \mu \in \mathbb{R}$. Moreover $\int_a^b (\lambda f + \mu g) = \lambda \int_a^b f + \mu \int_a^b g$. That is, the integrable functions on $[a,b]$ form a vector space and the integral is a linear functional (linear map to $\mathbb{R}$) on it.*

Proof. Suppose that $\lambda > 0$. If $\phi_- \leqslant f \leqslant \phi_+$ are minorant/majorant for f, then $\lambda\phi_- \leqslant \lambda f \leqslant \lambda\phi_+$ are minorant and majorant for λf. Moreover $I(\lambda\phi_-) - I(\lambda\phi_+) = \lambda(I(\phi_+) - I(\phi_-))$ can be made arbitrarily small. Thus λf is integrable. Moreover $\inf_{\phi_+} I(\lambda\phi_+) = \lambda \inf_{\phi_+} I(\phi_+)$, $\inf_{\phi_-} I(\lambda\phi_-) = \lambda \sup_{\phi_-} I(\phi_-)$, and so $\int_a^b (\lambda f) = \lambda \int_a^b f$. If $\lambda < 0$ then we can proceed in a very similar manner. We leave this to the reader.

Now suppose that $\phi_- \leqslant f \leqslant \phi_+$ and $\psi_- \leqslant g \leqslant \psi_+$ are minorant/majorants for f, g. Then $\phi_- + \psi_- \leqslant f + g \leqslant \phi_+ + \psi_+$ are minorant/majorant for $f + g$ (note these are steps functions) and by Lemma 1.3 (linearity of I)

$$\inf_{\phi_+,\psi_+} I(\phi_+ + \psi_+) = \inf_{\phi_+} I(\phi_+) + \inf_{\psi_+} I(\psi_+) = \int_a^b f + \int_a^b g,$$

whilst

$$\sup_{\phi_-,\psi_-} I(\phi_- + \psi_-) = \sup_{\phi_-} I(\phi_-) + \sup_{\psi_-} I(\psi_-) = \int_a^b f + \int_a^b g.$$

It follows that indeed $f+g$ is integrable and $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.

That $\int_a^b (\lambda f + \mu g) = \lambda \int_a^b f + \mu \int_a^b g$ follows immediately by combining these two facts. □

PROPOSITION 1.3. *Suppose that f and g are integrable on $[a,b]$. Then* $\max(f,g)$ *and* $\min(f,g)$ *are both Riemann integrable, as is* $|f|$.

Proof. We have $\max(f,g) = g + \max(f-g,0)$, $\min(h,0) = -\max(-h,0)$ and $|h| = \max(h,0) - \min(h,0)$. Using these relations and Proposition 1.2, it is enough to prove that if f is integrable on (a,b), then so is $\max(f,0)$.

Now the function $x \mapsto \max(x,0)$ is order-preserving (if $x \leqslant y$ then $\max(x,0) \leqslant \max(y,0)$) and non-expanding (we have $|\max(x,0) - \max(y,0)| \leqslant |x-y|$, as can be established by an easy case-check, according to the signs of x,y). It follows that if $\phi_- \leqslant f \leqslant \phi_+$ are minorant and majorant for f then $\max(\phi_-,0) \leqslant \max(f,0) \leqslant \max(\phi_+,0)$ are minorant and majorant for $\max(f,0)$ (it is obvious that they are both step functions) and, since f is integrable,

$$I(\max(\phi_+,0)) - I(\max(\phi_-,0)) \leqslant I(\phi_+) - I(\phi_-)$$

can be made arbitrarily small. □

PROPOSITION 1.4. *If f,g are both integrable on $[a,b]$ and if $f \leqslant g$ pointwise then $\int_a^b f \leqslant \int_a^b g$. In particular, $|\int_a^b f| \leqslant \int_a^b |f|$.*

Proof. The first part is immediately obvious from the fact that any minorant for f is also a minorant for g, and so

$$\int_a^b g = \sup_{\psi_-} I(\psi_-) \geqslant \sup_{\phi_-} I(\phi_-) = \int_a^b f,$$

where ψ_- ranges over minorants of g, and ϕ_- over minorants of f. For the second part, apply the first part to f and $|f|$, and also to $-f$ and $|f|$, obtaining $\pm \int_a^b f \leqslant \int_a^b |f|$. □

1.5. Not all functions are integrable

EXAMPLE 1.1. There is a bounded function $f : [0,1] \to \mathbb{R}$ which is not (Riemann) integrable.

Proof. Consider the function f such that $f(x) = 1$ if $x \in \mathbb{Q}$ and 0 if $x \notin \mathbb{Q}$. Since any open interval contains both rational points and points which are not rational, any step function majorising f must satisfy $\phi_+(x) \geqslant 1$ except possibly at the finitely many endpoints x_i, and hence $I(\phi_+) \geqslant 1$. Similarly any minorant ϕ_- satisfies $\phi_-(x) \leqslant 0$ except at finitely many points, and so $I(\phi_-) \leqslant 0$. This function f cannot possibly be integrable. □

Remark. Students will see in next year's course on Lebesgue integration that the Lebesgue integral of this function *does* exist (and equals 0).

1.6. Improper integrals – a brief discussion

If one attempts to assign a meaning to the integral of an unbounded function, or to the integral of a function over an unbounded domain, then one is trying to understand an *improper integral.* We will not attempt to systematically define what an improper integral is, but a few examples should make it clear what is meant in any given situation.

In discussing these examples we assume that the integrals of well-known functions are what you think they are – this will be justified rigorously later on, when we prove the fundamental theorem of calculus.

EXAMPLE 1.2. Consider the function $f(x) = \log x$. This is continuous on $(0, 1)$, but it is not integrable there since it is not bounded (it tends to $-\infty$ as $x \to \infty$). We have

$$\int_\varepsilon^1 \log x dx = [x \log x - x]_\varepsilon^1 = -1 - \varepsilon \log \varepsilon - \varepsilon.$$

Therefore

$$\lim_{\varepsilon \to 0^+} \int_\varepsilon^1 \log x dx = -1.$$

This will often be written as

$$\int_0^1 \log x dx = -1,$$

but strictly speaking this is not an integral as discussed in this course.

EXAMPLE 1.3. Consider the function $f(x) = 1/x^2$. We have

$$\int_1^K \frac{1}{x^2} dx = \big[-\frac{1}{x}\big]_1^K = 1 - \frac{1}{K}.$$

Therefore

$$\lim_{K \to \infty} \int_1^K \frac{1}{x^2} dx = 1.$$

This is invariably written

$$\int_1^\infty \frac{1}{x^2} dx = 1.$$

EXAMPLE 1.4. Define $f(x)$ to be $\log x$ if $0 < x \leqslant 1$, and $f(x) = \frac{1}{x^2}$ for $x \geqslant 1$. Then it makes sense to write

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

by which we mean

$$\lim_{K\to\infty,\varepsilon\to 0} \int_\varepsilon^K f(x)dx = 0.$$

This is a combination of the preceding two examples.

EXAMPLE 1.5. Define $f(x)$ to be $1/x$ for $0 < |x| \leqslant 1$, and $f(0) = 0$. Then f is unbounded as $x \to 0$. Excising the problematic region, one can look at

$$I_{\varepsilon,\varepsilon'} := \int_\varepsilon^1 f(x)dx + \int_{-1}^{-\varepsilon'} f(x)dx,$$

and one easily computes that

$$I_{\varepsilon,\varepsilon'} = \log \frac{\varepsilon'}{\varepsilon}.$$

This does not necessarily tend to a limit as $\varepsilon, \varepsilon' \to 0$ (for example, if $\varepsilon' = \varepsilon^2$ it does not tend to a limit). One will often hear the term *Cauchy principal value* (PV) for the limit $\lim_{\varepsilon\to 0} I_{\varepsilon,\varepsilon}$, which in this case equals 0. We won't discuss principal values any further in this course, and in this case it is *not* appropriate to write $\int_{-1}^1 \frac{1}{x}dx = 0$; one could possibly write PV $\int_{-1}^1 \frac{1}{x}dx = 0$.

EXAMPLE 1.6. Similarly to the last example, one should not write $\int_{-\infty}^\infty \sin x dx = 0$, even though $\lim_{K\to\infty} \int_{-K}^K \sin x dx = 0$ (because sin is an odd function). In this case, $\lim_{K,K'\to\infty} \int_{-K'}^K \sin x dx$ does not exist. One could maybe write

$$\mathrm{PV} \int_{-\infty}^\infty \sin x dx = 0,$$

but I would not be tempted to do so.

1.7. An example

Although we didn't mention it explicitly, it is obvious from the definition that every step function ϕ on $[a, b]$ is integrable and that $I(\phi) = \int_a^b \phi$ (take the minorant and majorant to be ϕ itself). In this section we pause to analyse a slightly less trivial example from first principles.

EXAMPLE 1.7. The function $f(x) = x$ is integrable on $[0, 1]$, and $\int_0^1 f(x)dx = \frac{1}{2}$.

Proof. We define explicit minorants and majorants. Let n be an integer to be specified later, and set $\phi_-(x) = \frac{i}{n}$ for $\frac{i}{n} \leqslant x \leqslant \frac{i+1}{n}$, $i = 0, 1, \dots, n-1$. Set $\phi_+(x) = \frac{j}{n}$ for $\frac{j-1}{n} \leqslant x \leqslant \frac{j}{n}$, $j = 1, \dots, n$. Then $\phi_- \leqslant f \leqslant \phi_+$ pointwise, so ϕ_-, ϕ_+

(being step functions) are minorant/majorant for f. We have

$$I(\phi_-) = \sum_{i=0}^{n-1} \frac{i}{n} \cdot \frac{1}{n} = \frac{1}{2}(1 - \frac{1}{n})$$

and

$$I(\phi_+) = \sum_{j=1}^{n} \frac{j}{n} \cdot \frac{1}{n} = \frac{1}{2}(1 + \frac{1}{n}).$$

It is clear that, by taking n sufficiently large, we can make both of these as close to $\frac{1}{2}$ as we like.

□

CHAPTER 2

Basic theorems about the integral

In this section we show that the integrable functions are in rich supply.

2.1. Continuous functions are integrable

Let $\mathcal{P}$ be a partition of $[a, b]$, $a = x_0 < x_1 < \cdots < x_n = b$. The *mesh* of $\mathcal{P}$ is defined to be $\max_i(x_i - x_{i-1})$. Thus if $\text{mesh}(\mathcal{P}) \leqslant \delta$ then every interval in the partition $\mathcal{P}$ has length at most δ. To give an example, if $[a, b] = [0, 1]$ and if $x_i = \frac{i}{N}$ then the mesh is $1/N$.

THEOREM 2.1. *A continuous function $f : [a, b] \to \mathbb{R}$ is integrable.*

Proof. Since f is continuous on a closed and bounded interval, f is also bounded. Suppose that M is a bound for f, so that $|f(x)| \leqslant M$ for all $x \in [a, b]$. We will also use the fact that a continuous function f is *uniformly* continuous. Let $\varepsilon > 0$, and let δ be so small that $|f(x) - f(y)| \leqslant \varepsilon$ whenever $|x - y| \leqslant \delta$. Let $\mathcal{P}$ be a partition with mesh $< \delta$. Let ϕ_+ be the step function whose value on (x_{i-1}, x_i) is $\sup_{x\in[x_{i-1},x_i]} f(x)$ and which takes the value M at the points x_i, and let ϕ_- be the step function whose value on (x_{i-1}, x_i) is $\inf_{x\in[x_{i-1},x_i]} f(x)$ and which takes the value $-M$ at the points x_i.

By construction, ϕ_+ is a majorant for f and ϕ_- is a minorant. Since a continuous function on a closed interval attains its bounds, there are $\xi_-, \xi_+ \in [x_{i-1}, x_i]$ such that $\sup_{x\in[x_{i-1},x_i]} f(x) = f(\xi_+)$ and $\inf_{x\in[x_{i-1},x_i]} f(x) = f(\xi_-)$.

For $x \in (x_{i-1}, x_i)$ we have $\phi_+(x) - \phi_-(x) \leqslant f(\xi_+) - f(\xi_-) \leqslant \varepsilon$. Therefore $\phi_+(x) - \phi_-(x) \leqslant \varepsilon$ for all except finitely many points in $[a, b]$, namely the points x_i.

It follows that $I(\phi_+) - I(\phi_-) \leqslant \varepsilon(b - a)$. Since ε was arbitrary, this concludes the proof. □

We can slightly strengthen this result, not insisting on continuity at the endpoints. This result would apply, for example, to the function $f(x) = \sin(1/x)$ on $(0, 1)$.

THEOREM 2.2. *A bounded continuous function $f : (a, b) \to \mathbb{R}$ is integrable.*

Proof. Suppose that $|f| \leqslant M$. Let $\varepsilon > 0$. Then f is continuous, and hence uniformly continuous, on $[a+\varepsilon, b-\varepsilon]$. Let δ be such that if $x, y \in [a+\varepsilon, b-\varepsilon]$ and $|x-y| \leqslant \delta$ then $|f(x) - f(y)| \leqslant \varepsilon$, and consider a partition $\mathcal{P}$ with $a = x_0$, $a + \varepsilon = x_1$, $b - \varepsilon = x_{n-1}$, $b = x_n$ and mesh $\leqslant \delta$.

Let ϕ_+ be the step function whose value on (x_{i-1}, x_i) is $\sup_{x \in [x_{i-1}, x_i]} f(x)$ when $i = 2, \dots, n-1$, and whose value on (x_0, x_1) and (x_{n-1}, x_n) is M.

Let ϕ_- be the step function whose value on (x_{i-1}, x_i) is $\inf_{x \in [x_{i-1}, x_i]} f(x)$ when $i = 2, \dots, n-1$, and whose value on (x_0, x_1) and (x_{n-1}, x_n) is $-M$.

Then $\phi_- \leqslant f \leqslant \phi_+$ pointwise. As in the proof of the previous theorem, we have $|\phi_+(x) - \phi_-(x)| \leqslant \varepsilon$ when $x \in (x_{i-1}, x_i)$, $i = 2, \dots, n-1$. On (x_0, x_1) and (x_{n-1}, x_n) we have the trivial bound $|\phi_+(x) - \phi_-(x)| \leqslant 2M$. Thus

$$I(\phi_+) - I(\phi-) \leqslant (b-a)\varepsilon + 2M \cdot 2\varepsilon,$$

which can be made arbitrarily small by taking ε arbitrarily small. □

Proposition 2.1 (Mean value theorem for integrals). *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous. Then there is some $c \in [a, b]$ such that*

$$\int_a^b f = (b-a) f(c).$$

Proof. Since f is continuous, it attains its maximum M and its minimum m. Moreover, the constant function $\phi_+ = M$ is a majorant for f, and the constant function ϕ_- is a minorant for f, so

$$m(b-a) = I(\phi_-) \leqslant \int_a^b f \leqslant I(\phi_+) = M(b-a),$$

which implies that

$$m \leqslant \frac{1}{b-a} \int_a^b f \leqslant M.$$

By the intermediate value theorem, f attains every value in $[m, M]$, and in particular there is some c such that

$$f(c) = \frac{1}{b-a} \int_a^b f.$$

□

2.2. Monotone functions are integrable

A function $f : [a, b] \to \mathbb{R}$ is said to be *monotone* if it is either non-decreasing (meaning $x \leqslant y$ implies $f(x) \leqslant f(y)$) or non-increasing (meaning $x \leqslant y$ implies $f(x) \geqslant f(y)$).

THEOREM 2.3. *A monotone function* $f : [a, b] \to \mathbb{R}$ *is integrable.*

Proof. By replacing f with $-f$ if necessary we may suppose that f is monotone non-decreasing, i.e. $f(x) \leqslant f(y)$ whenever $x \leqslant y$. Since $f(a) \leqslant f(x) \leqslant f(b)$, f is automatically bounded and in fact $|f(x)| \leqslant M$ where $M := \max(|f(a)|, |f(b)|)$.

Let $\varepsilon > 0$, and take a partition with mesh $< \varepsilon^2/4M^2$. Define step functions adapted to this partition as follows. On (x_{i-1}, x_i), define $\phi_+(x) = f(x_i)$ and $\phi_-(x) = f(x_{i-1})$. Defining $\phi_-(x_i) = -M$ and $\phi_+(x_i) = M$. Then ϕ_+ is a majorant for f and ϕ_- is a minorant.

Say that i is *good* if $f(x_i) - f(x_{i-1}) \leqslant \varepsilon$, or equivalently if $\phi_+(x) - \phi_-(x) \leqslant \varepsilon$ for $x \in (x_{i-1}, x_i)$, and bad otherwise. There cannot be more than $2M/\varepsilon$ bad intervals, since f increases by more than ε over each such interval yet is bounded everywhere by M.

The contribution to $I(\phi_+) - I(\phi-)$ from the bad intervals is thus at most $2M \cdot \frac{2M}{\varepsilon} \cdot \frac{\varepsilon^2}{4M^2} = \varepsilon$, since we *always* have $|\phi_+ - \phi_-| \leqslant 2M$.

The contribution from the good intervals is at most $\varepsilon(b - a)$.

Thus $I(\phi_+) - I(\phi_-) \leqslant (b - a + 1)\varepsilon$, which can be made arbitrarily small by making ε small enough. □

CHAPTER 3

Riemann sums

The way in which we have been developing the integral is closely related to the approach taken by Darboux. In this chapter we discuss what is essentially Riemann's original way of defining the integral, and show that it is equivalent. This is of more than merely historical interest: the equivalence of the definitions has several useful consequences.

If $\mathcal{P}$ is a partition and $f : [a,b] \to \mathbb{R}$ is a function then by a *Riemann sum* adapted to $\mathcal{P}$ we mean an expression of the form

$$\Sigma(f;\mathcal{P},\vec{\xi}) = \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1}),$$

where $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ and $\xi_j \in [x_{j-1}, x_j]$.

Proposition 3.1. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Fix a sequence of partitions $\mathcal{P}^{(i)}$. For each i, let $\Sigma(f,\mathcal{P}^{(i)},\vec{\xi}^{(i)})$ be a Riemann sum adapted to $\mathcal{P}^{(i)}$. Suppose that there is some constant c such that, no matter how $\vec{\xi}^{(i)}$ is chosen, $\Sigma(f,\mathcal{P}^{(i)},\vec{\xi}^{(i)}) \to c$. Then f is integrable and $c = \int_a^b f$.*

Proof. Let $\varepsilon > 0$, and let $M = \sup_{x\in[a,b]} |f(x)|$ be a bound for f. Suppose that $\mathcal{P}^{(i)}$ is $a = x_0^{(i)} \leqslant \ldots \leqslant x_{n_i}^{(i)} = b$. For each j, choose some point $\xi_j^{(i)} \in [x_{j-1}^{(i)}, x_j^{(i)}]$ such that $f(\xi_j^{(i)}) \geqslant \sup_{x\in[x_{j-1}^{(i)},x_j^{(i)}]} f(x) - \varepsilon$. (Note that f does not necessarily attain its supremum on this interval.) Let ϕ_+^i be a step function taking the value $f(\xi_j^{(i)}) + \varepsilon$ on $(x_{j-1}^{(i)}, x_j^{(i)})$, and with $\phi_+^{(i)}(x_j^{(i)}) = M$. Then $\phi_+^{(i)}$ is a majorant for f. It is easy to see that

$$I(\phi_+^{(i)}) \leqslant \varepsilon(b-a) + \Sigma(f;\mathcal{P}^{(i)},\vec{\xi}^{(i)}).$$

Taking i large enough that $\Sigma(f;\mathcal{P}^{(i)},\vec{\xi}^{(i)}) \leqslant c + \varepsilon$, we therefore have

$$I(\phi_+^{(i)}) \leqslant \varepsilon(b-a) + c + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, it follows that

$$\inf_{\phi_+} I(\phi_+) \leqslant c.$$

By an identical argument,

$$\sup_{\phi_-} I(\phi_-) \geqslant c.$$

Therefore

$$c \leqslant \sup_{\phi_-} I(\phi_-) \leqslant \inf_{\phi_+} I(\phi_+) \leqslant c,$$

and so all these quantities equal c. □

This suggests that we could use such Riemann sums to *define* the integral, perhaps by taking some natural choice for the sequences of partitions $\mathcal{P}^{(i)}$ such as $x_j^{(i)} = a + \frac{j}{i}(b-a)$ (the partition into i equal parts). However, Proposition 3.1 does not imply that this definition is equivalent to the one we have been using, since we have not shown that the Riemann sums converge if f is integrable. In fact, this requires an extra hypothesis. Recall that the *mesh* $\mathrm{mesh}(\mathcal{P})$ of a partition is the length of the longest subinterval in $\mathcal{P}$.

PROPOSITION 3.2. *Let $\mathcal{P}^{(i)}$, $i = 1, 2, \dots$ be a sequence of partitions satisfying* $\mathrm{mesh}(\mathcal{P}^{(i)}) \to 0$. *Suppose that f is integrable. Then $\lim_{i\to\infty} \Sigma(f, \mathcal{P}^{(i)}, \vec{\xi}^{(i)}) = \int_a^b f$, no matter what choice of $\vec{\xi}^{(i)}$ we make.*

Proof. Let M be a bound for f, that is to say $|f(x)| \leqslant M$ for all x. Let $\varepsilon > 0$. Then there are a minorant and majorant ϕ_-, ϕ_+ for f with

$$I(\phi_+) - \varepsilon \leqslant \int_a^b f \leqslant I(\phi-) + \varepsilon. \tag{3.1}$$

By refining the partitions underlying ϕ_-, ϕ_+ if necessary, we may assume that both of these functions are adapted to some partition $\mathcal{P}$, $a = x_0 \leqslant x_1 \leqslant \dots \leqslant x_n \leqslant b$. By replacing ϕ_+ with $\min(M, \phi_+)$ and ϕ_- with $\max(-M, \phi_-)$ we may assume that $-M \leqslant \phi_- \leqslant \phi_+ \leqslant M$ pointwise.

Let $\delta := \varepsilon/nM$. We claim that if $\mathcal{P}' : a = x'_0 \leqslant x'_1 \leqslant \dots \leqslant x'_{n'} = b$ is a partition of mesh at most δ, and if $\Sigma(f, \mathcal{P}, \vec{\xi}')$ is any Riemann sum associated to this partition, then

$$-3\varepsilon + \int_a^b f \leqslant \Sigma(f, \mathcal{P}', \vec{\xi}') \leqslant 3\varepsilon + \int_a^b f. \tag{3.2}$$

A moment's reflection convinces one that this claim implies the proposition, since $\varepsilon > 0$ was arbitrary.

We give the proof of the upper bound in the claim (3.2), the argument for the lower bound being essentially identical. Written out in full, the Riemann sum is

$$\Sigma(f; \mathcal{P}', \vec{\xi}') = \sum_{j=1}^{n'} f(\xi'_j)(x'_j - x'_{j-1}).$$

Since ϕ_+ is a majorant for f, we evidently have

$$\Sigma(f; \mathcal{P}', \vec{\xi}') \leqslant \sum_{j=1}^{n'} \phi_+(\xi'_j)(x'_j - x'_{j-1}).$$

Say that j is *good* if the interval (x'_{j-1}, x'_j) is wholly contained in one of the subintervals of $\mathcal{P}$. Say that j is *bad* if this is not the case. In this latter scenario, one or more of the partitioning points x_k of $\mathcal{P}$ lies in the interior (x'_{j-1}, x'_j). This means that the total length of the bad intervals is at most $n\delta$. Since $\phi_+ \leqslant M$ pointwise, it follows that

$$\Sigma(f; \mathcal{P}', \vec{\xi}') \leqslant n\delta M + \sum_{j \text{ good}} \phi_+(\xi'_j)(x'_j - x'_{j-1}). \tag{3.3}$$

However, since ϕ_+ is constant and equal to ξ'_j on (x'_{j-1}, x'_j) when j is good, and certainly at least $-M$ on the bad intervals, which have total length at most $n\delta$, we have

$$I(\phi_+) \geqslant \sum_{j \text{ good}} \phi_+(\xi'_j)(x'_j - x'_{j-1}) - n\delta M. \tag{3.4}$$

Comparing (3.3), (3.4) and using (3.1) gives

$$\Sigma(f; \mathcal{P}', \vec{\xi}') \leqslant 2n\delta M + I(\phi_+) = 2\varepsilon + I(\phi_+) \leqslant 3\varepsilon + \int_a^b f,$$

which is precisely the upper bound in the claim (3.2).

□

Proposition 3.1 and 3.2 together allow us to give an alternative definition of the integral. This is basically Riemann's original definition.

Proposition 3.3. *Let $f : [a, b] \to \mathbb{R}$ be a function. Let $\mathcal{P}^{(i)}$, $i = 1, 2, \dots$ be a sequence of partitions with* $\mathrm{mesh}(\mathcal{P}^{(i)}) \to 0$. *Then f is integrable if and only if* $\lim_{i\to\infty} \Sigma(f, \mathcal{P}^{(i)}, \vec{\xi^{(i)}})$ *is equal to some constant c, independently of the choice of* $\vec{\xi^{(i)}}$. *If this is so, then $\int_a^b f = c$.*

Finally, we caution that it is important that the limit must exist for *any* choice of $\vec{\xi^{(i)}}$. Suppose, for example, that $[a, b] = [0, 1]$ and that $\mathcal{P}^{(i)}$ is the partition into i equal parts, thus $x_j^{(i)} = \frac{j}{i}$ for $j = 1, \dots, i$. Take $\xi_j^{(i)} = \frac{j}{i}$; then the Riemann sum $\Sigma(f, \mathcal{P}^{(i)}.\vec{\xi^{(i)}})$ is equal to

$$S_i(f) := \frac{1}{i} \sum_{j=1}^{i} f(\frac{j}{i}).$$

By Proposition 3.2, if f is integrable then

$$S_i(f) \to \int_a^b f.$$

However, the converse is not true. Consider, for example, the function f introduced in the first chapter, with $f(x) = 1$ for $x \in \mathbb{Q}$ and $f(x) = 0$ otherwise. This function is not integrable, as we established in that chapter. However,

$$S_i(f) = 1 \qquad \text{for all } i.$$

CHAPTER 4

Integration and differentiation

It is a well-known fact, which goes by the name of "the fundamental theorem of calculus" that "integration and differentiation are inverse to one another and that if $f = F'$ then $\int_a^b f = F(b) - F(a)$". Our objective in this chapter is to prove rigorous versions of this fact. We will prove two statements, sometimes known as the first and second fundamental theorems of calculus respectively, though there does not seem to be complete consensus on this matter.

4.1. First fundamental theorem of calculus

The first thing to note is that the statement just given is not true without some additional assumptions. Consider, for instance, the function $F : \mathbb{R} \to \mathbb{R}$ defined by $F(0) = 0$ and $F(x) = x^2 \sin \frac{1}{x^2}$ for $x \neq 0$. Then it is a standard exercise to show that F is differentable everywhere, with $f = F'$ given by $f(0) = 0$ and $f(x) = 2x \sin(1/x^2) - \frac{2}{x} \cos(1/x^2)$. In particular, f is unbounded on any interval containing 0, and so it has no majorants and is not integrable according to our definition.

An even worse example (the Volterra function) can be constructed with f bounded, but still not integrable. This construction is rather elaborate and we will not give it here.

These constructions show that a hypothesis of integrability should be built into any statement of the fundamental theorem of calculus.

Theorem 4.1 (First fundamental theorem). *Suppose that f is integrable on (a, b). Define a new function $F : [a, b] \to \mathbb{R}$ by*

$$F(x) := \int_a^x f(t)dt.$$

Then F is continuous. Moreover, if f is continuous at $c \in (a, b)$ then F is differentiable at c and $F'(c) = f(c)$.

Proof. The fact that F is continuous follows immediately from the fact that f is bounded (which it must be, as it is integrable), say by M. Then

$$|F(c+h) - F(c)| = |\int_c^{c+h} f(t)dt| \leqslant \int_c^{c+h} |f(t)| \leqslant Mh.$$

In fact, this argument directly establishes that F is uniformly continuous (and in fact uniformly Lipschitz).

Now we turn to the second part. Suppose that $c \in (a, b)$ and that $h > 0$ is sufficiently small that $c + h < b$. We have

$$F(c+h) - F(c) = \int_c^{c+h} f(t)dt.$$

Since f is continuous at c, there is a function $\varepsilon(h)$, $\varepsilon(h) \to 0$ as $h \to 0$, such that we have $|f(t) - f(c)| \leqslant \varepsilon(h)$ for all $t \in [c, c+h]$. Therefore

$$|F(c+h) - F(c) - hf(c)| = |\int_c^{c+h} (f(t) - f(c))| \leqslant \varepsilon(h)h. \tag{4.1}$$

Essentially the same argument works for $h < 0$ (in fact, exactly the same argument works if we interpret $\int_c^{c+h} f(t)dt$ in the natural way as $-\int_{c+h}^{c} f(t)dt$). Statement (4.1) is exactly the definition of F being differentiable at c with derivative f. □

We note that F is not necessarily differentiable assuming only that f is Riemann-integrable. For example if we take the function f defined by $f(t) = 0$ for $t \leqslant \frac{1}{2}$ and $f(t) = 1$ for $t > \frac{1}{2}$ then f is integrable on $[0, 1]$, and the function $F(x) = \int_0^x f(t)dt$ is given by $F(x) = 0$ for $x \leqslant \frac{1}{2}$ and $F(x) = x - \frac{1}{2}$ for $\frac{1}{2} \leqslant x \leqslant 1$. Evidently, F fails to be differentiable at $\frac{1}{2}$.

4.2. Second fundamental theorem of calculus

We turn now to the "second form" of the fundamental theorem.

THEOREM 4.2 (Second fundamental theorem). *Suppose that $F : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Suppose furthermore that its derivative F' is integrable on (a, b). Then*

$$\int_a^b F'(t)dt = F(b) - F(a).$$

Proof. Let $\mathcal{P}$ be a partition, $a = x_0 < x_1 < \cdots < x_n = b$. We claim that some Riemann sum $\Sigma(f; \mathcal{P}, \xi)$ is equal to $F(b) - F(a)$. By Proposition 3.2 (the harder direction of the equivalence between integrability and limits of Riemann sums), the second fundamental theorem follows immediately from this.

The claim is an almost immediate consequence of the mean value theorem. By that theorem, we may choose $\xi_i \in (x_{i-1}, x_i)$ so that $F'(\xi_i)(x_i - x_{i-1}) = F(x_i) - F(x_{i-1})$. Summing from $i = 1$ to n gives

$$\Sigma(F'; \mathcal{P}, \xi) = \sum_{i=1}^{n} (F(x_i) - F(x_{i-1})) = F(b) - F(a).$$

□

4.3. Integration by parts

Everyone knows that integration by parts says that

$$\int_a^b fg' = f(b)g(b) - f(a)g(a) - \int_a^b f'g.$$

We are now in a position to prove a rigorous version of this.

Before doing so, we establish a lemma which we could have proven at the start of the course, in Section 1.4.

LEMMA 4.1. *Suppose that $f, g : (a, b) \to \mathbb{R}$ are integrable. Then so is their product fg.*

Proof. Since $fg = \frac{1}{4}(f+g)^2 - \frac{1}{4}(f-g)^2$, it is enough to show that f^2 is integrable if f is. For any scalar $\lambda \in \mathbb{R}$ we have $f^2 = (f+\lambda)^2 - 2\lambda f - \lambda^2$, and so it suffices to show that $(f+\lambda)^2$ is integrable. Taking λ sufficiently large, we may therefore assume $f \geqslant 0$.

Suppose that $0 \leqslant f(x) \leqslant M$. Let $\phi_- \leqslant f \leqslant \phi_+$ be minorant/majorant for f with $I(\phi_+) - I(\phi_-) \leqslant \varepsilon$. Replacing ϕ_+ with $\min(\phi_+, M)$ and ϕ_- with $\max(\phi_-, 0)$, we may assume that $\phi_+ \leqslant M$ and $\phi_- \geqslant 0$ pointwise.

Now we have $\phi_-^2 \leqslant f^2 \leqslant \phi_+^2$ pointwise, and so ϕ_-^2, ϕ_+^2, being step functions, are minorant/majorant for f. Moreover

$$\begin{aligned} I(\phi_+^2) - I(\phi_-^2) &= I(\phi_+^2 - \phi_-^2) \\ &= I((\phi_+ + \phi_-)(\phi_+ - \phi_-)) \\ &\leqslant 2MI(\phi_+ - \phi_-) \leqslant 2M\varepsilon, \end{aligned}$$

since $\phi_+ + \phi_- \leqslant 2M$ everywhere. Since ε was arbitrary, the result follows. □

Remark. In fact it is true that if f is integrable and ψ is continuous then $\psi \circ f$ is integrable, where $\psi \circ f(t) = \psi(f(t))$ is the composition of f with ψ. The proof of the preceding lemma established this in the special case $\psi(x) = x^2$. We leave the general case as an exercise for the interested student.

PROPOSITION 4.1. *Suppose that $f, g : [a, b] \to \mathbb{R}$ are continuous functions, differentiable on (a, b). Suppose that the derivatives f', g' are integrable on (a, b). Then fg' and $f'g$ are integrable on (a, b), and*

$$\int_a^b fg' = f(b)g(b) - f(a)g(a) - \int_a^b f'g.$$

Proof. We use the second form of the fundamental theorem of calculus, applied to the function $F = fg$. We know from basic differential calculus that F is differentiable and $F' = f'g + fg'$. By Lemma 4.1 and the assumption that f', g' are integrable, F' is integrable on (a, b). Applying the fundamental theorem gives

$$\int_a^b F'(t)dt = F(b) - F(a),$$

which is obviously equivalent to the stated claim. □

4.4. Substitution

PROPOSITION 4.2 (Substitution rule). *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous and that $\phi : [c, d] \to [a, b]$ is continuous on $[c, d]$, has $\phi(c) = a$ and $\phi(d) = b$, and maps (c, d) to (a, b). Suppose moreover that ϕ is differentiable on (c, d) and that its derivative ϕ' is integrable on this interval. Then*

$$\int_a^b f(x)dx = \int_c^d f(\phi(t))\phi'(t)dt.$$

Proof. Let us first remark that $(f \circ \phi)(t) = f(\phi(t))$ is continuous and hence integrable on $[c, d]$. It therefore follows from Lemma 4.1 that $f(\phi(t))\phi'(t)$ is integrable as a function of t, so the statement does at least make sense.

Since f is continuous on $[a, b]$, it is integrable. The first fundamental theorem of calculus implies that its antiderivative

$$F(x) := \int_a^x f(t)dt$$

is continuous on $[a, b]$, differentiable on (a, b) and that $F' = f$.

Write $(F \circ \phi)(t) := F(\phi(t))$. By the chain rule and the fact that $\phi((c, d)) \subset (a, b)$, $F \circ \phi$ is differentiable on (c, d), and

$$(F \circ \phi)' = (F' \circ \phi)\phi' = (f \circ \phi)\phi'.$$

By the remarks at the start of the proof, it follows that $(F \circ \phi)'$ is an integrable function. By the second form of the fundamental theorem,

$$\begin{aligned}\int_c^d f(\phi(t))\phi'(t)dt &= \int_c^d (F \circ \phi)' \\ &= (F \circ \phi)(d) - (F \circ \phi)(c) \\ &= F(b) - F(a) \\ &= F(b) = \int_a^b f(x)dx.\end{aligned}$$

□

CHAPTER 5

Limits and the integral

5.1. Interchanging the order of limits and integration

Suppose we have a sequence of functions f_n converging to a limit function f. If this convergence is merely pointwise, integration need not preserve the limit.

EXAMPLE 5.1. There is a sequence of integrable functions $f_n : [0,1] \to \mathbb{R}$ (in fact, step functions) such that $f_n(x) \to 0$ pointwise for all $x \in [0,1]$ but $\int f_n = 1$ for all n. Thus $\lim_{n\to\infty} \int_0^1 f_n = 1$, whilst $\int_0^1 \lim_{n\to\infty} f_n = 0$, and so interchange of integration and limit is not valid in this case.

Proof. Define $f_n(x)$ to be equal to n for $0 < x < \frac{1}{n}$ and 0 elsewhere. □

However, if $f_n \to f$ uniformly then the situation is much better.

THEOREM 5.1. *Suppose that $f_n : [a,b] \to \mathbb{R}$ are integrable, and that $f_n \to f$ uniformly on $[a,b]$. Then f is also integrable, and*

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b f = \int_a^b \lim_{n\to\infty} f_n.$$

Proof. Let $\varepsilon > 0$. Since $f_n \to f$ uniformly, there is some choice of n such that we have $|f_n(x) - f(x)| \leqslant \varepsilon$ for all $x \in [a,b]$.

Now f_n is integrable, and so there is a majorant ϕ_+ and a minorant ϕ_- for f_n with $I(\phi_+) - I(\phi_-) \leqslant \varepsilon$.

Define $\tilde{\phi}_+ := \phi_+ + \varepsilon$ and $\tilde{\phi}_- := \phi_- - \varepsilon$. Then $\tilde{\phi}_-, \tilde{\phi}_+$ are minorant/majorant for f. Moreover

$$\begin{aligned} I(\tilde{\phi}_+) - I(\tilde{\phi}_-) &\leqslant 2\varepsilon(b-a) + I(\phi_+) - I(\phi_-) \\ &\leqslant 2\varepsilon(b-a) + \varepsilon. \end{aligned}$$

Since ε was arbitrary, this shows that f is integrable. Now

$$|\int_a^b f_n - \int_a^b f| \leqslant \int_a^b |f_n - f| \leqslant (b-a) \sup_{x\in[a,b]} |f_n(x) - f(x)|.$$

Since $f_n \to f$ uniformly, it follows that

$$\lim_{n\to\infty} |\int_a^b f_n - \int_a^b f| = 0,$$

and hence that

$$\lim_{n\to\infty}\int_a^b f_n = \int_a^b f = \int_a^b \lim_{n\to\infty} f_n.$$

This concludes the proof. □

An immediate corollary of this is that we may integrate series term-by-term under suitable conditions.

Corollary 5.1. *Suppose that $\phi_i : [a,b] \to \mathbb{R}$, $i = 1, 2, \dots$ are integrable functions and that $|\phi_i(x)| \leqslant M_i$ for all $x \in [a,b]$, where $\sum_{i=1}^\infty M_i < \infty$. Then the sum $\sum_i \phi_i$ is integrable and*

$$\int_a^b \sum_i \phi_i = \sum_i \int_a^b \phi_i.$$

Proof. This is immediate from the Weierstrass M-test and Theorem 5.1, applied with $f_n = \sum_{i=1}^n \phi_i$. □

5.2. Interchanging the order of limits and differentiation

The behaviour of limits with respect to differentiation is much worse than the behaviour with respect to integration.

Example 5.2. There is a sequence of functions $f_n : [0,1] \to \mathbb{R}$, each continuously differentiable on $(0,1)$, such that $f_n \to 0$ uniformly but such that f_n' does not converge at every point.

Proof. Take $f_n(x) = \frac{1}{n}\sin(n^2 x)$. Then $f_n'(x) = -n\cos(n^2 x)$. Taking $x = \frac{\pi}{4}$, we see that if n is a multiple of 4 then $f_n'(x) = -n$, which certainly does not converge. □

If, however, we assume that the derivatives f_n' converge uniformly then we do have a useful result.

Proposition 5.1. *Suppose that $f_n : [a,b] \to \mathbb{R}$, $n = 1, 2, \dots$ is a sequence of functions with the property that f_n is continuously differentiable on (a,b), that f_n converges pointwise to some function f on $[a,b]$, and that f_n' converges uniformly to some bounded function g on (a,b). Then f is differentiable and $f' = g$. In particular, $\lim_{n\to\infty} f_n' = (\lim_{n\to\infty} f_n)'$.*

Proof. First note that, since the f_n' are continuous and $f_n' \to g$ uniformly, g is continuous. Since we are also assuming g is bounded, it follows from Theorem 2.2 that g is integrable.

We may therefore apply the first form of the fundamental theorem of calculus to g. Since g is continuous, the theorem says that if we define a function $F : [a,b] \to \mathbb{R}$

by

$$F(x) := \int_a^x g(t)dt$$

then F is differentiable with $F' = g$. By the second form of the fundamental theorem of calculus applied to f_n, we have

$$\int_a^x f_n'(t)dt = f_n(x) - f_n(a).$$

Taking limits as $n \to \infty$ and using the fact that $f_n \to f$ pointwise, we obtain

$$\lim_{n\to\infty} \int_a^x f_n'(t)dt = f(x) - f(a).$$

However, since $f_n' \to g$ uniformly, it follows from Theorem 5.1 that

$$\lim_{n\to\infty} \int_a^x f_n'(t)dt = \int_a^x g(t)dt.$$

Thus

$$F(x) = \int_a^x g(t)dt = f(x) - f(a).$$

It follows immediately that f is differentiable and that its derivative is the same as that of F, namely g. □

Remark. Note that the statement of Proposition 5.1 involves only differentiation. However, the proof involves a considerable amount of the theory of integration. This is a theme that is seen throughout mathematical analysis. For example, the nice behaviour of complex differentiable functions (which you will see in course A2 next year) is a consequence of Cauchy's *integral* formula.

As with our result about integrals, we can record a "series variant" of Proposition 5.1.

COROLLARY 5.2. *Suppose we have a sequence of continuous functions $\phi_i : [a, b] \to \mathbb{R}$, continuously differentiable on (a, b), with $\sum_i \phi_i$ converging pointwise. Suppose that $|\phi_i'(x)| \leqslant M_i$ for all $x \in (a, b)$, where $\sum_i M_i < \infty$. Then $\sum \phi_i$ is differentiable and*

$$(\sum_i \phi_i)' = \sum_i \phi_i'.$$

Proof. Apply Proposition 5.1 with $f_n := \sum_{i=1}^n \phi_i$. By the Weierstrass M-test, $f_n' = \sum_{i=1}^n \phi_i'$ does converge pointwise to some bounded function, which we may call g. □