Analysis III

Marc Lackenby[1]

[1]The original version of these notes was created by Ben Green; only very minor modifications have been made to them

Contents

Preface 1

Chapter 1. Step functions and the Riemann integral 3
- 1.1. Step functions 3
- 1.2. I of a step function 4
- 1.3. Definition of the integral 5
- 1.4. A lemma and some examples 6
- 1.5. Basic theorems about the integral 8

Chapter 2. Basic theorems about the integral 13
- 2.1. Continuous functions are integrable 13
- 2.2. Mean value theorems 14
- 2.3. Monotone functions are integrable 15

Chapter 3. Riemann sums 17

Chapter 4. Integration and differentiation 21
- 4.1. First fundamental theorem of calculus 21
- 4.2. Second fundamental theorem of calculus 22
- 4.3. Integration by parts 23
- 4.4. Substitution 24

Chapter 5. Limits and the integral 25
- 5.1. Interchanging the order of limits and integration 25
- 5.2. Interchanging the order of limits and differentiation 26
- 5.3. Power series and radius of convergence 27

Chapter 6. The exponential and logarithm functions 31
- 6.1. The exponential function 31
- 6.2. The logarithm function 33

Chapter 7. Improper integrals 35

Preface

These are the notes for Analysis III at Oxford. 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.

Please send any corrections to

`lackenby@maths.ox.ac.uk`.

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 a *refinement* of $\mathcal{P}$ if every x_i is an x'_j for some j.

LEMMA 1.3. *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 $\mathcal{P}_1, \mathcal{P}_2$. $\square$

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 1.4. *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 1.3. $\square$

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 1.5. 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 1.6. *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. $\square$

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 1.7. 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 a minorant and majorant, adapted to partitions $\mathcal{P}_-$ and $\mathcal{P}_+$ respectively. By passing to a common refinement we may assume that $\mathcal{P}_- = \mathcal{P}_+$. 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 minorant and majorants.

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.

Remark on dx. Integrals are often written using the dx notation. For example, $\int_0^1 x^2 dx$. This means the same as $\int_0^1 f$, where $f(x) = x^2$. We emphasise that in this course this is nothing more than a piece of notation. The dx tells us which variable f is a function of. This can sometimes be very useful to avoid confusion.

1.4. A lemma and some examples

The definition of the integral given in the previous section seems, at first sight, to be hard to verify in practice. It is defined in terms of *all* majorants ϕ_+ and minorants ϕ_- for the function f. How might we compute $\sup_{\phi_-} I(\phi_-)$ and $\inf_{\phi_+} I(\phi_+)$? The following very useful lemma provides a necessary and sufficient condition for a function f to be integrable. We will see that it can also be used to compute the integral in specific examples.

LEMMA 1.8. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Then the following are equivalent:*

(i) *f is integrable;*
(ii) *for every $\varepsilon > 0$, there is a majorant ϕ_+ and a minorant ϕ_- for f such that $I(\phi_+) - I(\phi_-) < \varepsilon$.*

Proof. Suppose first that f is integrable. Let $\varepsilon > 0$. Then by the approximation property for sup and inf, there is a minorant ϕ_- such that

$$I(\phi_-) > \sup_{\phi_-} I(\phi_-) - (\varepsilon/2)$$

and a majorant ϕ_+ such that

$$I(\phi_+) < \inf_{\phi_+} I(\phi_+) + (\varepsilon/2).$$

Since the sup and inf are assumed to be equal, we deduce that

$$I(\phi_+) - I(\phi_-) < \varepsilon.$$

Now suppose that (ii) holds. Let $\varepsilon > 0$ be arbitrary, and let ϕ_+ and ϕ_- be the majorant and minorant provided by (ii). Then

$$I(\phi_+) < I(\phi_-) + \varepsilon \leqslant \sup_{\phi_-} I(\phi_-) + \varepsilon.$$

So, taking the infimum over all majorants, we deduce that

$$\inf_{\phi_+} I(\phi_+) < \sup_{\phi_-} I(\phi_-) + \varepsilon.$$

Therefore, $\inf_{\phi_+} I(\phi_+)$ is squeezed between $\sup_{\phi_-} I(\phi_-)$ and $\sup_{\phi_-} I(\phi_-)+\varepsilon$. Since $\varepsilon > 0$ was arbitrary, we deduce that the inf and sup must be equal. In other words, f is integrable. $\square$

Once we know that f is integrable, then any majorant ϕ_+ and minorant ϕ_- as in (ii) gives an approximation to the integral, by (1.3). This is because $\int_a^b f$ lies between $I(\phi_-)$ and $I(\phi_+)$ which differ by less than ε.

The following example demonstrates how useful this is in practice.

EXAMPLE 1.9. 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 $\frac{i}{n} \leqslant x < \frac{i+1}{n}$, $i = 0, 1, \ldots, n-1$. Set $\phi_+(x) = \frac{j}{n}$ for $\frac{j-1}{n} \leqslant x < \frac{j}{n}$, $j = 1, \ldots, n$. 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

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

So, by Lemma 1.8, f is integrable. Moreover, the integral of f must lie between $\frac{1}{2}(1 - \frac{1}{n})$ and $\frac{1}{2}(1 + \frac{1}{n})$. Since n was arbitrary, the integral must be $\frac{1}{2}$. □

Earlier we defined $I(\phi)$ for a step function ϕ. We can now prove that this actually agrees with the integral of ϕ.

PROPOSITION 1.10. *Suppose that ϕ is a step function on $[a, b]$. Then ϕ is integrable, and $\int_a^b \phi = I(\phi)$.*

Proof. Take $\phi_- = \phi_+ = \phi$, and the result is immediate from Lemma 1.8 and (1.3). □

COROLLARY 1.11. *There is a non-negative integrable function f on $[a, b]$ which is not identically zero, but for which $\int_a^b f = 0$.*

Proof. Simply take f to be the zero function, modified at one point. □

But now we come to a "non-example".

EXAMPLE 1.12. 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).

1.5. 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 1.13. *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 $\int_a^b f = \int_a^c f + \int_c^b f$.*

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 $i = 1, 2$. (Here, we used the fact that if $x \leqslant x'$, $y \leqslant y'$ and $x + y = x' + y'$ then $x = x'$ and $y = y'$.) Thus f is indeed integrable on $[a, c]$ and on $[c, b]$, and it follows from (1.4) that $\int_a^b f = \int_a^c f + \int_c^b f$. $\square$

COROLLARY 1.14. *Suppose that $f : [a, b] \to \mathbb{R}$ is integrable, and that $[c, d] \subseteq [a, b]$. Then f is integrable on $[c, d]$.*

Proof. This is immediate. $\square$

PROPOSITION 1.15. *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 $\inf_{\phi_+} I(\lambda\phi_+) = \lambda \inf_{\phi_+} I(\phi_+)$, $\sup_{\phi_-} I(\lambda\phi_-) = \lambda \sup_{\phi_-} I(\phi_-)$, 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. If $\lambda = 0$, then λf is identically zero and hence is integrable by Proposition 1.10.

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 1.6 (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. $\square$

Corollary 1.16. *If f is integrable on $[a, b]$, and if $\tilde{f}$ differs from f in finitely many points, then $\tilde{f}$ is also integrable.*

Proof. The function $\tilde{f} - f$ is zero except at finitely many points. Suppose that these points are $x_1, \dots, x_{n-1}$. Then $\tilde{f} - f$ is a step function adapted to the partition $a = x_0 \leqslant x_1 \leqslant \dots \leqslant x_{n-1} \leqslant x_n = b$. By Proposition 1.10, $\tilde{f} - f$ is integrable, and hence so is $\tilde{f} = (\tilde{f} - f) + f$, by Proposition 1.15. $\square$

Proposition 1.17. *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 1.15, 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). Moreover,

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

Since f is integrable, this can be made arbitrarily small. $\square$

Proposition 1.18. *Suppose that f is integrable on $[a, b]$.*

(i) *We have $(b-a)\inf_{x\in[a,b]} f(x) \leqslant \int_a^b f \leqslant (b-a)\sup_{x\in[a,b]} f(x)$.*
(ii) *If g is another integrable function on $[a, b]$, and if $f \leqslant g$ pointwise, then $\int_a^b f \leqslant \int_a^b g$.*
(iii) $|\int_a^b f| \leqslant \int_a^b |f|$.

Proof. (i) is an immediate consequence of the fact that the constant function $\phi_-(x) = \inf_{x\in[a,b]} f(x)$ is a minorant for f on $[a, b]$, whilst $\phi_+(x) = \sup_{x\in[a,b]} f(x)$ is a majorant. Thus

$$(b-a)\inf_{x\in[a,b]} f(x) = I(\phi_-) \leqslant \sup_{\phi_-} I(\phi_-) = \int_a^b f,$$

and similarly for the upper bound.

(ii) Applying (i) to $g-f$ gives $\int_a^b (g-f) \geqslant 0$, from which the result is immediate from linearity of the integral.

(iii) Apply (ii) to f and $|f|$, and also to $-f$ and $|f|$, obtaining $\pm\int_a^b f \leqslant \int_a^b |f|$. $\square$

Finally, we look at products.

Proposition 1.19. *Suppose that $f, g : [a, b] \to \mathbb{R}$ are two integrable functions. Then their product fg is integrable.*

Proof. Write $f = f_+ - f_-$, where $f_+ = \max(f, 0)$ and $f_- = -\min(f, 0)$, and similarly for g. Then $fg = f_+g_+ - f_-g_+ - f_+g_- + f_-g_-$, and so it suffices to prove the statement for non-negative functions such as $f_\pm, g_\pm$. Suppose, then, that $f, g \geqslant 0$. Let $\varepsilon > 0$, and let $\phi_- \leqslant f \leqslant \phi_+$, $\psi_- \leqslant g \leqslant \psi_+$ be minorants and majorants for f, g with $I(\phi_+) - I(\phi_-), I(\psi_+) - I(\psi_-) \leqslant \varepsilon$. Replacing ϕ_- with $\max(\phi_-, 0)$ if necessary (and similarly for ψ_-), we may assume that $\phi_-, \psi_- \geqslant 0$ pointwise. Replacing ϕ_+ with $\min(\phi_+, M)$, where $M = \max\{\sup_{[a,b]} f, \sup_{[a,b]} g\}$ (and similarly for ψ_+) we may assume that $\phi_+, \psi_+ \leqslant M$ pointwise. By refining partitions if necessary, we may assume that all of these step functions are adapted to the same partition $\mathcal{P}$. Now observe that $\phi_-\psi_-, \phi_+\psi_+$ are both step functions and that $\phi_-\psi_- \leqslant fg \leqslant \phi_+\psi_+$ pointwise. Moreover, if $0 \leqslant u, v, u', v' \leqslant M$ and $u \leqslant u'$, $v \leqslant v'$ then we have

$$u'v' - uv = (u'-u)v' + (v'-v)u \leqslant M(u' - u + v' - v). \tag{1.5}$$

Applying this on each interval of the partition $\mathcal{P}$, with $u = \phi_-$, $u' = \phi_+$, $v = \psi_-$, $v' = \psi_+$, we have

$$I(\phi_+\psi_+) - I(\phi_-\psi_-) \leqslant M(I(\phi_+) - I(\phi_-) + I(\psi_+) - I(\psi_-)) \leqslant 2\varepsilon M.$$

Since $\varepsilon > 0$ was arbitrary, the result follows. $\square$

Remark. Here is a sketch of an alternative proof, which is arguably a little slicker, or at least easier notationally. Note the identity $fg = \frac{1}{4}(f+g)^2 - \frac{1}{4}(f-g)^2$. Thus it suffices to show that if f is integrable then so is f^2. Replacing f by $|f|$, we may assume that $f \geqslant 0$ pointwise. Then proceed as above but with $f = g$, $\phi_- = \psi_-$, $\phi_+ = \psi_+$. In place of (1.5) one may instead use $(u')^2 - u^2 \leqslant 2M(u' - u)$.

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 $\mathrm{mesh}(\mathcal{P}) \leqslant \delta$ 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. *Continuous functions $f : [a, b] \to \mathbb{R}$ are integrable.*

Proof. Since f is continuous on a closed and bounded interval, f is also bounded. 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 $f(x_i)$ 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 $f(x_i)$ 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 bounded 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 $x \in [a, b]$.

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

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. *Bounded continuous functions $f : (a, b) \to \mathbb{R}$ are 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. □

In the first chapter, we gave a simple example of a nonnegative function f which has zero integral, but is not identically zero. The following simple lemma shows that this cannot happen in the world of continuous functions.

Lemma 2.3. *Suppose that $f : [a,b] \to \mathbb{R}$ is a continuous function with $f \geqslant 0$ pointwise and $\int_a^b f = 0$. Then $f(x) = 0$ for $x \in [a,b]$.*

Proof. Suppose not. Then there is some point $x \in [a,b]$ with $f(x) > 0$, let us say $f(x) = \varepsilon$. Since f is continuous, there is some $\delta > 0$ such that if $|x-y| \leqslant \delta$ then $|f(x)-f(y)| \leqslant \varepsilon/2$, and hence $|f(x)| \geqslant \varepsilon/2$. The set of all $y \in [a,b]$ with $|x-y| \leqslant \delta$ is a subinterval $I \subset [a,b]$ with length ℓ at least $\min(b-a,\delta)$, and so

$$\int f \geqslant \int_I f \geqslant \frac{\varepsilon}{2}\min(b-a,\delta) > 0.$$

□

2.2. Mean value theorems

The integrals of continuous functions satisfy various "mean value theorems". Here is a simple instance of such a result.

Proposition 2.4. *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. By Proposition 1.18 (i),

$$m(b-a) \leqslant \int_a^b f \leqslant M(b-a),$$

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

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

□

The following slightly more complicated result, which generalises the above, may be established in essentially the same way.

PROPOSITION 2.5. *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous, and that $w : [a, b] \to \mathbb{R}$ is a nonnegative integrable function. Then there is some $c \in [a, b]$ such that*

$$\int_a^b fw = f(c)\int_a^b w.$$

Proof. First one should remark that fw is indeed integrable, this being a consequence of Proposition 1.19. As in the proof of Proposition 2.4, write M, m for the maximum and minimum of f respectively. Then $mw \leqslant fw \leqslant Mw$ pointwise, and so

$$m\int_a^b w \leqslant \int_a^b fw \leqslant M\int_a^b w.$$

If $\int_a^b w = 0$ then the result follows immediately; otherwise, we may divide through to get

$$m \leqslant \frac{\int_a^b fw}{\int_a^b w} \leqslant M.$$

Since both m and M are values attained by f, the result now follows from the intermediate value theorem. □

Remark. Just to be clear, Proposition 2.4 is the case $w = 1$ of Proposition 2.5.

2.3. Monotone functions are integrable

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

THEOREM 2.6. *Monotone functions $f : [a, b] \to \mathbb{R}$ are integrable.*

Proof. By replacing f with $-f$ if necessary we may suppose that f is monotone increasing, i.e. $f(x) \leqslant f(y)$ whenever $x \leqslant y$. Since $f(a) \leqslant f(x) \leqslant f(b)$, f is automatically bounded.

Let n be a positive integer, and consider the partition of $[a, b]$ into n equal parts. Thus $\mathcal{P}$ is $a = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_n = b$, with $x_i = a + \frac{i}{n}(b-a)$. On (x_{i-1}, x_i), define $\phi_+(x) = f(x_i)$ and $\phi_-(x) = f(x_{i-1})$. Define $\phi_-(x_i) = f(x_i)$ and $\phi_+(x_i) = f(x_i)$. Then ϕ_+ is a majorant for f and ϕ_- is a minorant. We have

$$\begin{aligned} I(\phi_+) - I(\phi_-) &= \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))(x_i - x_{i-1}) \\ &= \frac{b-a}{n} \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \\ &= \frac{1}{n}(b-a)(f(b) - f(a)). \end{aligned}$$

Taking n large, this can be made as small as desired. $\square$

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 $\vec{\xi} = (\xi_1, \dots, \xi_n)$ 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 $\varepsilon > 0$. Let i be chosen so that $\Sigma(f; \mathcal{P}^{(i)}, \vec{\xi}^{(i)}) \leqslant c + \varepsilon$, no matter which $\vec{\xi}^{(i)}$ is chosen. Write $\mathcal{P} = \mathcal{P}^{(i)}$, and suppose that $\mathcal{P}$ is $a = x_0 \leqslant \dots \leqslant x_n = b$. For each j, choose some point $\xi_j \in [x_{j-1}, x_j]$ such that $f(\xi_j) \geqslant \sup_{x \in [x_{j-1}, x_j]} f(x) - \varepsilon$. (Note that f does not necessarily attain its supremum on this interval.) Let ϕ_+ be a step function taking the value $f(\xi_j) + \varepsilon$ on (x_{j-1}, x_j), and with $\phi_+(x_j) = f(x_j)$. Then ϕ_+ is a majorant for f. It is easy to see that

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

We therefore have

$$I(\phi_+) \leqslant \varepsilon(b - a) + c + \varepsilon.$$

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. Throughout the proof, write $M := \sup_{x\in[a,b]} |f(x)|$. Let $\mathcal{P} : a = x_0 \leqslant x_1 \leqslant \cdots \leqslant x_n = b$ be a partition. In this proof it is convenient to introduce the notion of the *optimal* majorant $\phi_+^{\mathcal{P}}$ for f relative to $\mathcal{P}$ (and similarly minorant). This is the majorant defined by

$$\phi_+^{\mathcal{P}} := \begin{cases} \sup_{x\in(x_{i-1},x_i)} f(x) & \text{on } (x_{i-1}, x_i) \\ f(x_i) & \text{at the points } x_i. \end{cases}$$

It is easy to see that if ϕ_+ is any majorant for f adapted to $\mathcal{P}$, then $I(\phi_+^{\mathcal{P}}) \leqslant I(\phi_+)$. Similarly, $I(\phi_-^{\mathcal{P}}) \geqslant I(\phi_-)$, and so

$$I(\phi_+^{\mathcal{P}}) - I(\phi_-^{\mathcal{P}}) \leqslant I(\phi_+) - I(\phi_-).$$

Let $\varepsilon > 0$. Since f is integrable it follows from what we just said that there is a partition $\mathcal{P} : a = x_0 \leqslant x_1 \leqslant \cdots \leqslant x_n = b$ such that $I(\phi_+^{\mathcal{P}}) - I(\phi_-^{\mathcal{P}}) < \varepsilon$. In particular, since $I(\phi_-) \leqslant \int_a^b f$ for any minorant ϕ_-,

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

Set $\delta := \varepsilon/nM$. Let $\mathcal{P}' : a = x'_0 \leqslant x'_1 \leqslant \ldots \leqslant x'_{n'} = b$ be any partition with $\mathrm{mesh}(\mathcal{P}') \leqslant \delta$, and consider an arbitrary Riemann sum

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

This is equal to $I(\psi)$, where the step function ψ is defined to be $f(\xi'_j)$ on (x'_{j-1}, x'_j) and $f(x'_j)$ at the x'_j.

Let us compare ψ and the optimal majorant $\phi_+^{\mathcal{P}}$.

Say that j is *good* if $[x'_{j-1}, x'_j] \subset (x_{i-1}, x_i)$ for some i. If j is good then, for $t \in (x'_{j-1}, x'_j)$,

$$\psi(t) = f(\xi'_j) \leqslant \sup_{x \in [x'_{j-1}, x'_j]} f(x) \leqslant \sup_{x \in (x_{i-1}, x_i)} f(x) = \phi_+^{\mathcal{P}}(t). \tag{3.2}$$

If j is bad (i.e. not good) then we cannot assert such a bound, but we do have the trivial bound

$$\psi(t) \leqslant \phi_+^{\mathcal{P}}(t) + 2M \tag{3.3}$$

for all j.

Now if j is bad then we have $x_i \in [x'_{j-1}, x'_j]$ for some i. No x_i can belong to more than two intervals $[x'_{j-1}, x'_j]$, so there cannot be more than $2n$ bad values of j. Therefore the total length of the corresponding intervals (x'_{j-1}, x'_j) is at most $2\delta n = 2\varepsilon/M$.

It therefore follows, using (3.2) on the good intervals and (3.3) on the bad, that

$$\Sigma(f; \mathcal{P}', \vec{\xi'}) = I(\psi) \leqslant I(\phi_+^{\mathcal{P}}) + 2M \cdot \frac{2\varepsilon}{M} = I(\phi_+^{\mathcal{P}}) + 4\varepsilon. \tag{3.4}$$

Combining this with (3.1) yields

$$\Sigma(f; \mathcal{P}', \vec{\xi'}) \leqslant \int_a^b f + 5\varepsilon.$$

There is a similar lower bound, proven in an analogous manner.

Since ε was arbitrary, this concludes the proof. □

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 $\mathcal{P}^{(i)}$, $i = 1, 2, \ldots$ 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 $j = 1, \ldots, i$. 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". 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. The first theorem deals with integration followed by differentiation. In the second theorem, we differentiate, then integrate.

So far, we have considered integrals of the form $\int_a^b f$. But we now want to vary the interval over which we integrate, as follows. We define the function

$$F(x) = \int_a^x f$$

for $x \in [a, b]$. Under suitable assumptions, we will show that F is differentiable with derivative f.

4.1. First fundamental theorem of calculus

The first thing to notice is that it is just not true that integration and differentiation are inverses without some additional assumptions.

Example 4.1. If f is not continuous, then F can be differentiable but it need not be the case that $F' = f$. For example, let $f : [0, 1] \to \mathbb{R}$ be the function that takes value 1 at $x = \frac{1}{2}$ but that is 0 elsewhere. Then F is identically zero. Hence, F is differentiable and F' is the zero function. This shows that in general, when you integrate and then differentiate, you might not get the original function back.

Example 4.2. 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 $\frac{1}{2}$.

However, the first fundamental theorem of calculus asserts that the function F *is* differentiable and $F' = f$, as long as f is continuous.

THEOREM 4.3 (First fundamental theorem). *Suppose that f is integrable on (a,b). Define a new function $F : [a,b] \to \mathbb{R}$ by*

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

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

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

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

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

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

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

Divide through by h:

$$\left| \frac{F(c+h) - F(c)}{h} - f(c) \right| \leqslant \varepsilon. \tag{4.1}$$

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

4.2. Second fundamental theorem of calculus

We turn now to the "second form" of the fundamental theorem, which deals with differentiation, followed by integration. Here, we cannot get such a strong result as the first fundamental theorem.

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 of F' should be built into any statement of the second fundamental theorem of calculus.

THEOREM 4.4 (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*

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

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

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

$\square$

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.

PROPOSITION 4.5. *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 Proposition 1.19 and the assumption that f', g' are integrable, F' is integrable on (a, b). Applying the fundamental theorem gives

$$\int_a^b F' = F(b) - F(a),$$

which is obviously equivalent to the stated claim. □

4.4. Substitution

PROPOSITION 4.6 (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*

$$\int_a^b f = \int_c^d (f \circ \phi)\phi'.$$

Remark. It may help to see the statement written out in full:

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

Proof. Let us first remark that $f \circ \phi$ is continuous and hence integrable on $[c, d]$. It therefore follows from Proposition 1.19 that $(f \circ \phi)\phi'$ is integrable on $[c, d]$, 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

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

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

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,

$$\begin{aligned}\int_c^d (f \circ \phi)\phi' &= \int_c^d (F \circ \phi)' \\ &= (F \circ \phi)(d) - (F \circ \phi)(c) \\ &= F(b) - F(a) \\ &= F(b) = \int_a^b f.\end{aligned}$$

□

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 5.2. *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 5.3. *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 5.2, 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 5.4. 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 5.5. *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 5.2 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 5.5 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.

5.3. Power series and radius of convergence

In this section we link to some results you will have seen in Analysis II. The proofs you saw there were slightly unpleasant. The use of integration is the "correct" way to prove these statements.

Let us begin by recording a "series variant" of Proposition 5.5.

COROLLARY 5.6. *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 5.5 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. □

Now suppose we have a sequence $(a_i)_{i=0}^{\infty}$ of real numbers. Then the expression $\sum_{i=0}^{\infty} a_i x^i$ is called a (formal) power series. The word "formal" means that we are not actually evaluating the sum over i; indeed, this may well not be possible for a given choice of the a_i and x. In fact, a formal power series is just the same thing as a sequence $(a_i)_{i=0}^{\infty}$, only written a different way; it is a very similar concept to that of a generating function.

DEFINITION 5.7. Given a formal power series $\sum_i a_i x^i$, we define its radius of convergence R to be the supremum of all $|x|$ for which the sum $\sum_{i=0}^{\infty} |a_i x^i|$ converges. If this sum converges for all x, we write $R = \infty$.

Theorem 5.9 below was called the "differentiation theorem for power series" in Analysis II. We isolate a simple lemma from the proof.

LEMMA 5.8. *Suppose that* $0 \leqslant \lambda < 1$. *Then* $\sum_{i=0}^{\infty} \lambda^i$ *and* $\sum_{i=1}^{\infty} i\lambda^{i-1}$ *both converge.*

Proof. By the well-known geometric series formula we have

$$\sum_{i=0}^{n-1} \lambda^i = \frac{1-\lambda^n}{1-\lambda}.$$

Letting $n \to \infty$ gives the first statement immediately, the value of the sum being $\frac{1}{1-\lambda}$.

For the second statement, we differentiate the geometric series formula. This gives

$$\sum_{i=1}^{n-1} i\lambda^{i-1} = \frac{1+(n-1)\lambda^n - n\lambda^{n-1}}{(1-\lambda)^2},$$

which tends to $\frac{1}{(1-\lambda)^2}$ as $n \to \infty$. $\square$

**Remark.* In the last step of this lemma we used the fact that $\lim_{n\to\infty} n\lambda^n = 0$. Since we are developing theory which will allow us to define the exponential function and explore its basic properties, we should check that we know how to prove such a statement without using the exponential function, or else our arguments would be circular. One method is to use the binomial theorem, ignoring all but the third term, to get

$$\frac{1}{\lambda^n} = \left(1 + (\frac{1}{\lambda} - 1)\right)^n \geqslant \binom{n}{2}(\frac{1}{\lambda} - 1)^2.$$

Therefore

$$n\lambda^n \leqslant \frac{n}{\binom{n}{2}(\frac{1}{\lambda} - 1)^2},$$

which fairly obviously tends to 0 as $n \to \infty$.

THEOREM 5.9. *Suppose a formal power series* $\sum_{i=0}^{\infty} a_i x^i$ *has radius of convergence* R. *Then the series converges for* $|x| < R$, *giving a well-defined function*

$f(x) = \sum_{i=0}^{\infty} a_i x^i$. *Moreover, f is differentiable on $(-R, R)$, and its derivative is given by term-by-term differentiation, that is to say $f'(x) = \sum_{i=1}^{\infty} i a_i x^{i-1}$. Moreover, the radius of convergence for this power series for f' is at least R.*

Proof. If $R = 0$, there is nothing to prove. Suppose that $R > 0$. Let R_1 satisfy $0 < R_1 < R$. We apply Corollary 5.6 with $\phi_i(x) = a_i x^i$ and $[a, b] = [-R_1, R_1]$. We need to check that the hypotheses of that result are satisfied. By definition of the radius of convergence, there is some R_0 satisfying $R_1 < R_0 \leqslant R$ for which $\sum_i |a_i R_0^i|$ converges, and in particular $|a_i R_0^i|$ is bounded, uniformly in i: let us say that $|a_i R_0^i| \leqslant K$. Then if $x \in [a, b]$ we have

$$|\phi_i(x)| \leqslant K(\frac{R_1}{R_0})^i$$

and

$$|\phi_i'(x)| \leqslant \frac{K}{R_0} i (\frac{R_1}{R_0})^{i-1}. \tag{5.1}$$

The first condition of Corollary 5.6, that is to say pointwise convergence of $\sum_i \phi_i(x)$, is now immediate from the first part of Lemma 5.8. Taking $M_i := \frac{K}{R_0} i (\frac{R_1}{R_0})^{i-1}$, we obtain the other condition of Corollary 5.6 from the second part of Lemma 5.8.

It now follows from Corollary 5.6 that f is differentiable on $(-R_1, R_1)$, and that is derivative is given by term-by-term differentiation of the power series for f. Since $R_1 < R$ was arbitrary, we may assert the same on $(-R, R)$.

Finally, it follows from (5.1) and Lemma 5.8 that the radius of convergence of the power series for f' is at least R_1. Since $R_1 < R$ was arbitrary, the radius of convergence of this power series is at least R, as claimed. □

By applying this theorem repeatedly, it follows that under the same assumptions f is infinitely differentiable on $(-R, R)$, with all of its derivatives being given by term-by-term differentiation.

CHAPTER 6

The exponential and logarithm functions

Every mathematician knows the exponential and logarithm functions and their basic properties. However, most of us rarely bother to think about which of these properties are in fact *definitions*, and which are *theorems*. In this chapter (which is not, strictly, on the schedules) we discuss these two functions and the relation between them from first principles. This provides some good exercise in the material we have developed so far, as well as useful material for examples. Similar treatments can be given for the sine and cosine functions (which, of course, are closely related to the exponential function via Euler's relation $e^{i\theta} = \cos\theta + i\sin\theta$): see the third example sheet.

6.1. The exponential function

We begin with a simple lemma (the solution to the simplest possible differential equation).

LEMMA 6.1. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function with $f' = f$ identically and $f(0) = 0$. Then f is identically zero.*

Proof. Since f is continuous, it attains its maximum value on $[0, \frac{1}{2}]$ at some point x. Suppose that $x > 0$. By the mean value theorem,

$$f(x) = f(x) - f(0) = xf'(\xi) = xf(\xi)$$

for some point $\xi \in (0, x)$. Therefore

$$f(x) \leqslant xf(x) \leqslant \frac{1}{2}f(x),$$

which implies that $f(x) \leqslant 0$. That is, $f \leqslant 0$ on $[0, \frac{1}{2}]$. Applying the same argument to $-f$ gives $f \geqslant 0$ on $[0, \frac{1}{2}]$, and so $f = 0$ identically on $[0, \frac{1}{2}]$.

We may now apply the same argument to $g(x) = f(x - \frac{1}{2})$, which satisfies $g' = g$ and $g(0) = 0$. We conclude that g is identically zero on $[0, \frac{1}{2}]$, and hence that f is identically zero on $[\frac{1}{2}, 1]$ and hence on $[0, 1]$. Continuing in this manner eventually shows that f is identically zero on the whole of $\mathbb{R}$. $\square$

THEOREM 6.2 (The exponential function). *For $x \in \mathbb{R}$, define*

$$e(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \tag{6.1}$$

Then

(i) *The series converges for all x, and $e(x)$ is a differentiable function satisfying $e' = e$.*
(ii) *We have $e(x) > 0$ for all $x \in \mathbb{R}$.*
(iii) *We have $e(x+y) = e(x)e(y)$ for all $x, y \in \mathbb{R}$.*

Proof. (i) Clearly we may work on bounded intervals $[-X, X]$, and let $X \to \infty$ at the end. Fix X from now on.

We will apply Corollary 5.6, taking $\phi_k(x) = x^k/k!$. Since the ϕ_k are continuously differentiable and $\phi_k'(x) = x^{k-1}/(k-1)!$, the result will follow if we can show that $|\phi_k(x)| \leqslant M_k$ for $x \in [-X, X]$, where $\sum_k M_k$ converges. (In particular it will then follow from the Weierstrass M-test that $\sum_k \phi_k(x)$ converges, which is one of the hypotheses of Corollary 5.6).

To establish this, we can use very crude bounds. Observe that $k! \geqslant (k/2)^{k/2}$ for all k, since the product for $k!$ contains at least $k/2$ terms of size $k/2$ or greater. Therefore

$$\frac{x^k}{k!} \leqslant (\frac{2x^2}{k})^{k/2}.$$

If $|x| \leqslant X$ and $k \geqslant 8X^2$, this is bounded above by 2^{-k}, which clearly converges when summed over k. (One could also use the "ratio test" here, though personally I wouldn't bother when a direct proof is so short.)

(ii) Suppose that $e(a) = 0$. Consider the function $f(x) = e(x+a)$; then $f(0) = 0$ and $f' = f$. By Lemma 6.1, f is identically zero and hence so is e. But this is a contradiction, as e is clearly not identically zero (for example $e(0) = 1$).

Thus e never vanishes. Since it is continuous, and positive somewhere, the intermediate value theorem implies that it is positive everywhere.

(iii) Consider the function $\tilde{e}(x) = \frac{e(x+y)}{e(y)}$. As just established, $e(y) \neq 0$ and so for every fixed y this is a continuous function of x. Moreover by the chain rule we have $\tilde{e}'(x) = \tilde{e}(x)$, and by direct substitution we have $\tilde{e}(0) = e(0) = 1$.

Therefore the function $f := e - \tilde{e}$ satisfies the hypotheses of Lemma 6.1. It follows that $\tilde{e}(x) = e(x)$, which is what we were required to prove. □

Of course, $e(x)$ is just the standard exponential function, which we customarily denote by e^x. In view of the properties just established, this notation is sensible. The point we wish to emphasise is that (6.1) is the *definition*, whilst properties (i), (ii) and (iii) (though they are so familiar as to feel almost like axioms) require proof.

6.2. The logarithm function

As everyone knows, the logarithm is the inverse of the exponential function. It is a function from $(0, \infty)$ to $\mathbb{R}$. Whilst one can define it like this, it is in many ways more natural to define it as an integral.

THEOREM 6.3. *For $x > 0$, define*

$$L(x) = \int_1^x \frac{dy}{y}. \tag{6.2}$$

Then

(i) *L is differentiable with derivative $\frac{1}{x}$ at each $x > 0$;*
(ii) *$L(e^t) = t$ for all $t \in \mathbb{R}$.*

Remarks. As with the exponential function, (6.2) is a definition, and the other statements are theorems, even though they are incredibly well-known. Of course, the function L is more usually written log.

To ensure the definition makes sense when $x < 1$, we define $\int_b^a f$ to be $-\int_a^b f$ when $a < b$. (We could have developed the theory of the integral $\int_a^b$ when $b < a$ more generally, but this is rather dull and routine.)

Proof. (i) This is *almost* immediate from the first fundamental theorem of calculus except that we need to convince ourselves that it still applies when $x \leqslant 1$. This may be done as follows. Let $c > 0$ and write

$$\int_1^x \frac{dy}{y} = \int_c^x \frac{dy}{y} - \int_c^1 \frac{dy}{y}.$$

It is easy to check that this holds for any $c > 0$. Then we may apply the fundamental theorem of calculus to get that $L'(x) = \frac{1}{x}$ for any $x > c$. Since c was arbitrary, the result follows.

(ii) We use the substitution rule, Proposition 4.6, taking $f(y) = \frac{1}{y}$ and $\phi(t) = e^t$. Note that $f(\phi(t))\phi'(t) = 1$, since $\phi' = \phi$. We therefore have

$$\int_1^{e^x} \frac{dt}{t} = \int_0^x (f \circ \phi)\phi' = x.$$

We leave the reader to check that the conditions required in the substitution lemma are valid (this is easy). $\square$

CHAPTER 7

Improper integrals

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.

EXAMPLE 7.1. 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 $x \to 0$). However, it *is* integrable on any interval $[\varepsilon, 1]$, $\varepsilon > 0$.

By the second fundamental theorem of calculus (and the fact that if $F(x) = x \log x - x$ then $F'(x) = \log x$) we have

$$\int_\varepsilon^1 \log x dx = [x \log x - x]_\varepsilon^1 = -1 - \varepsilon \log \varepsilon - \varepsilon. \tag{7.1}$$

We claim that

$$\lim_{\varepsilon \to 0^+} \varepsilon \log \varepsilon = 0. \tag{7.2}$$

Once this is shown, it follows from (7.1) that

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

This will often be written as

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

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

Let us give a proof of (7.2). We will show how this follows straight from the definition of $\log x$ given in Section 6.2. In Example Sheet 4, Q2, you will find another proof using the fact that log is inverse to exp.

To prove (7.2), recall that

$$\log \varepsilon = -\int_\varepsilon^1 \frac{dx}{x}$$

for $\varepsilon < 1$. We divide the range of integration into the ranges $[\varepsilon, \sqrt{\varepsilon}]$ and $[\sqrt{\varepsilon}, 1]$. On the first range we have $1/x \leqslant 1/\varepsilon$ and so

$$|\int_{\varepsilon}^{\sqrt{\varepsilon}} \frac{dx}{x}| \leqslant \frac{1}{\sqrt{\varepsilon}}.$$

On the second range we have $1/x \leqslant 1/\sqrt{\varepsilon}$ and so

$$|\int_{\sqrt{\varepsilon}}^{1} \frac{dx}{x}| \leqslant \frac{1}{\sqrt{\varepsilon}}.$$

It follows that

$$|\log \varepsilon| \leqslant \frac{2}{\sqrt{\varepsilon}},$$

from which (7.2) follows immediately.

EXAMPLE 7.2. Consider the function $f(x) = 1/x^2$. We would like to discuss the integral of this "on $[1, \infty)$", but this is not permitted by the way we have defined the integral, which requires a bounded interval. However, on any bounded interval $[1, K]$ 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 7.3. 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 7.4. Define $f(x)$ to be $1/x$ for $0 < |x| \leqslant 1$, and $f(0) = 0$. Then f is unbounded as $x \to 0$, and so we cannot define the integral $\int_{-1}^{1} f$. We can nonetheless try to make some kind of sense of this quantity.

Excising the problematic region around 0, 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 7.5. 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.