Chapter 2

Integration

At school, and in your methods courses last year and this, you will have been introduced to the integral. Typically, it is introduced as 'an inverse operation of differentiation'. In practice, to integrate a function $f : \mathbb{R} \to \mathbb{R}$, we try and find an $F : \mathbb{R} \to \mathbb{R}$ such that $F'(x) = f(x)$. We call F the antiderivative (or indefinite integral) of f, and write:

$$\int_a^b f(t)dt := F(b) - F(a).$$

For many of the functions that you are familiar with, this approach is fine. We have a catalogue of antiderivatives, and can follow this algorithm. When we want to make statements about *general* functions f, things become more difficult. We need a better definition of the integral, one which we can use even if we're not able to find an antiderivative. The basic idea that we shall pursue is the following:

$$\int_a^b f(t)dt := \text{'The area under the graph of } f \text{ between } a \text{ and } b\text{'}.$$

It will turn out that we can make sense of the area under the graph of f for a reasonably large class of functions. The approach we shall take will lead us to a definition of the integral known as the *Riemann* integral (strictly we shall use the Darboux integral, but the two are equivalent). It is not the only possible definition of an integral. If you take the course M3P19 Measure and Integration you will be introduced to the Lebesgue integral, which for many problems in analysis is more powerful.

The basic idea will be to split the region under the graph of f up and approximate the area by sufficiently narrow rectangles, as shown in Figures 2.1, 2.2. We will estimate the area from above by considering rectangles whose tops lie above the graph, and we will estimate the area from below by considering rectangles whose tops lie below the graph. For a 'reasonable' function, these two estimates should get closer and closer as the size of the rectangles shrinks to zero. Of course, since this is an analysis course, we will need to be precise about notions like 'closer and closer' and 'reasonable' in this context.

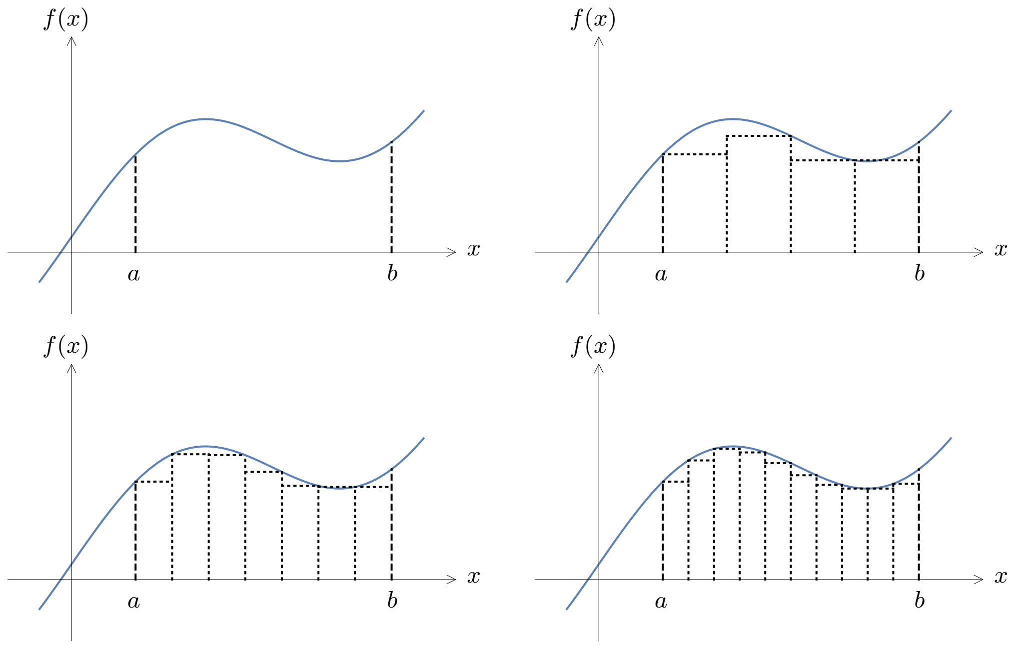

Figure 2.1 Estimating the area under a graph from below with rectangles

2.1 The Darboux integral on $\mathbb{R}$

The basic idea of our definition of the integral has been spelled out above, but we will need to spend a little bit of time setting up some machinery first in order to make our definition precise.

2.1.1 Partitions of an interval

Throughout this section, we will be working on a finite interval $I := [a, b] \subset \mathbb{R}$. A *partition*, $\mathcal{P}$, of the interval I is a finite sequence of numbers $\mathcal{P} := (x_0, \dots, x_k)$ satisfying:

$$a = x_0 < x_1 < \dots < x_{k-1} < x_k = b.$$

Given a partition, we can break the interval I up into segments $[x_j, x_{j+1}]$ which overlap only at their endpoints. We call these segments *subintervals*. We denote the length of the n^{th} subinterval by Δx_j, i.e.

$$\Delta x_j = x_{j+1} - x_j, \qquad j = 0, \dots, k-1.$$

Notice that we don't require the subintervals to all have the same length. The *mesh* of a partition $\mathcal{P} = (x_0 \dots, x_k)$ is the length of the longest subinterval:

$$\text{mesh } \mathcal{P} := \max_{j=0,\dots,k-1} \Delta x_j = \max_{j=0,\dots,k-1} (x_{j+1} - x_j).$$

Given two partitions, $\mathcal{P}$, $\mathcal{Q}$, we say that $\mathcal{Q}$ is a *refinement* of $\mathcal{P}$, or $\mathcal{Q}$ *refines* $\mathcal{P}$ if every point of $\mathcal{P}$ is also a point of $\mathcal{Q}$. In this case, we write $\mathcal{P} \preceq \mathcal{Q}$.

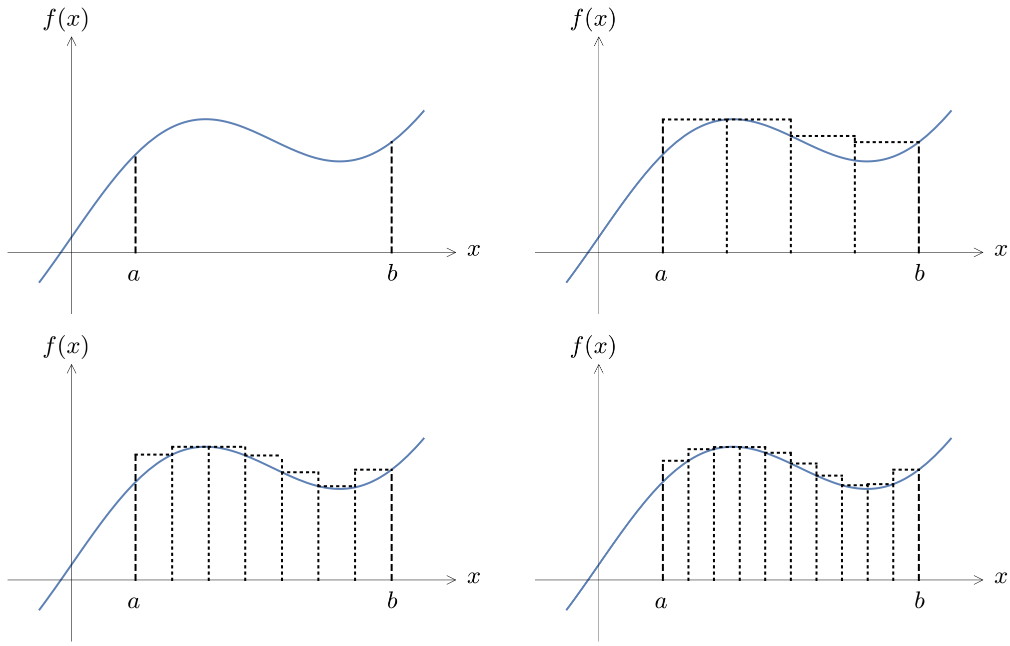

Figure 2.2 Estimating the area under a graph from above with rectangles

Lemma 2.1. *Given two partitions, $\mathcal{P}$, $\mathcal{Q}$, there is a unique minimal partition $\mathcal{R}$ such that $\mathcal{P} \preceq \mathcal{R}$* and *$\mathcal{Q} \preceq \mathcal{R}$. $\mathcal{R}$ is minimal in the sense that if $\mathcal{R}'$ satisfies $\mathcal{P} \preceq \mathcal{R}'$* and *$\mathcal{Q} \preceq \mathcal{R}'$, then $\mathcal{R} \preceq \mathcal{R}'$.*

Proof. Given $\mathcal{P} := (x_0, \ldots, x_k)$ and $\mathcal{Q} := (y_0, \ldots, y_l)$, we simply take $\mathcal{R}$ to be the set $\{x_0, \ldots, x_k, y_0, \ldots, y_l\}$ with any duplicates deleted and re-ordered as appropriate to give $\mathcal{R} = (z_0, \ldots, z_m)$. This certainly gives a refinement of both $\mathcal{P}$ and $\mathcal{Q}$, and moreover it is minimal, as any other partition which refines both $\mathcal{P}$ and $\mathcal{Q}$ must contain all the points of $\mathcal{R}$ and hence must refine $\mathcal{R}$. □

We call $\mathcal{R}$ as constructed above the *common refinement* of $\mathcal{P}$ and $\mathcal{Q}$. Figure 2.3 shows a sketch of the common refinement of two partitions.

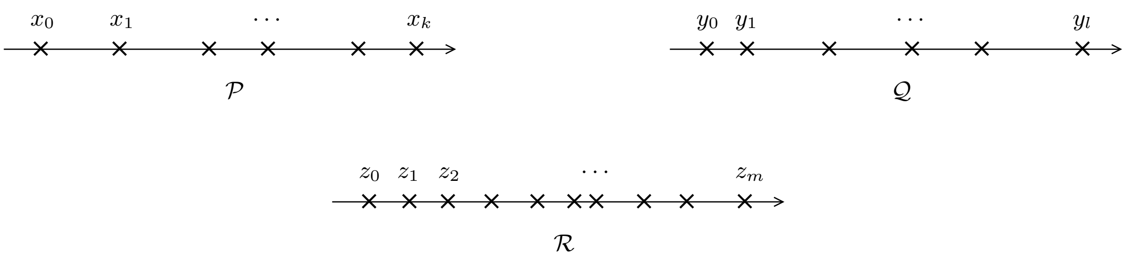

Figure 2.3 The common refinement $\mathcal{R}$ of two partitions $\mathcal{P}, \mathcal{Q}$.

2.1.2 Upper and lower Darboux integrals

Now, let us consider a bounded function $f : [a, b] \to \mathbb{R}$. It is helpful for visualising things to think of f as continuous to begin with, but we will not make that assumption. Given a partition $\mathcal{P} = (x_0, \dots, x_k)$ of the interval $[a, b]$, we introduce the numbers:

$$m_j := \inf_{x \in [x_j, x_{j+1}]} f(x),$$
$$M_j := \sup_{x \in [x_j, x_{j+1}]} f(x),$$

for $j = 0, \dots, k-1$, which are respectively the greatest lower and least upper bounds on f in the interval $[x_j, x_{j+1}]$. Since f is bounded, these are all finite and well-defined. Next, we introduce the *lower* and *upper Darboux sums* of f with respect to $\mathcal{P}$:

$$L(f, \mathcal{P}) := \sum_{j=0}^{k-1} m_j \Delta x_j$$
$$U(f, \mathcal{P}) := \sum_{j=0}^{k-1} M_j \Delta x_j.$$

Recall that $\Delta x_j = (x_{j+1} - x_j)$ is the length of the interval $[x_j, x_{j+1}]$, so each term in the sum represents the area of a rectangle with base Δx_j and height m_j (resp. M_j). The geometrical picture describing these sums is given in Figures 2.1, 2.2. In both cases, the sums represent the area of the rectangles, which in the case of the lower sum *under* estimates the area under the graph and in the case of the upper sum *over* estimates the area under the graph. In this discussion, we have assumed that the graph lies above the axis. In the case where it doesn't, we understand 'area' to be negative when it's below the axis.

Since $m_j \leq M_j$ for all $j = 0, k-1$, we certainly have:

$$L(f, \mathcal{P}) \leq U(f, \mathcal{P})$$

We expect that the area under the graph of f should lie somewhere between these two bounds. We also expect that a more refined partition should estimate the area more accurately. We can make this precise with the following theorem, which is really the crucial result of this section.

Theorem 2.2. *Let $f : [a, b] \to \mathbb{R}$ be bounded. Suppose that $\mathcal{P}$, $\mathcal{Q}$ are partitions of $[a, b]$ with $\mathcal{Q}$ a refinement of $\mathcal{P}$. Then:*

$$L(f, \mathcal{P}) \leq L(f, \mathcal{Q}) \leq U(f, \mathcal{Q}) \leq U(f, \mathcal{P}).$$

Proof. Suppose first that $\mathcal{P} = (x_0, \dots, x_k)$ and $\mathcal{Q} = (x_0, \dots, x_l, c, x_{l+1}, \dots x_k)$, so that $\mathcal{P}$ and $\mathcal{Q}$ differ only in that a single subinterval $[x_l, x_{l+1}]$ has been split into $[x_l, c]$ and $[c, x_{l+1}]$. From basic properties of the infimum, we have:

$$\inf_{x \in [x_l, c]} f(x) \geq \inf_{x \in [x_l, x_{l+1}]} f(x), \quad \text{and} \quad \inf_{x \in [c, x_{l+1}]} f(x) \geq \inf_{x \in [x_l, x_{l+1}]} f(x).$$

Now, if we take the difference of the lower sums associated to $\mathcal{P}$ and $\mathcal{Q}$, then the only terms which will not cancel will be those involving the subinterval $[x_l, x_{l+1}]$:

$$\begin{aligned} L(f,\mathcal{Q}) - L(f,\mathcal{P}) &= (c - x_l) \inf_{x\in[x_l,c]} f(x) + (x_{l+1} - c) \inf_{x\in[c,x_{l+1}]} f(x) \\ &\quad - (x_{l+1} - x_l) \inf_{x\in[x_l,x_{l+1}]} f(x) \\ &\geq [(c - x_l) + (x_{l+1} - c) - (x_{l+1} - x_l)] \inf_{x\in[x_l,x_{l+1}]} f(x), \\ &\geq 0. \end{aligned}$$

Thus we have:

$$L(f,\mathcal{P}) \leq L(f,\mathcal{Q}).$$

Now, note that by Exercise 4.2 *a*), we have $U(f,\mathcal{P}) = -L(-f,\mathcal{P})$, so that:

$$U(f,\mathcal{Q}) = -L(-f,\mathcal{Q}) \leq -L(-f,\mathcal{P}) = U(f,\mathcal{P}),$$

so that we have established

$$L(f,\mathcal{P}) \leq L(f,\mathcal{Q}) \leq U(f,\mathcal{Q}) \leq U(f,\mathcal{P}).$$

in the special case where $\mathcal{P} = (x_0, \ldots, x_k)$ and $\mathcal{Q} = (x_0, \ldots, x_l, c, x_{l+1}, \ldots x_k)$. For the general result, we observe that if $\mathcal{P} \preceq \mathcal{Q}$ then we can obtain $\mathcal{Q}$ by inserting a finite number of points into $\mathcal{P}$, and we can apply our special case result at each stage to obtain the general result. □

Corollary 2.3. *Let $f : [a,b] \to \mathbb{R}$ be bounded. Suppose $\mathcal{P}, \mathcal{P}'$ are any two partitions of $[a,b]$. Then:*

$$L(f,\mathcal{P}) \leq U(f,\mathcal{P}').$$

Proof. Let $\mathcal{Q}$ be the common refinement of $\mathcal{P}, \mathcal{P}'$. Then by Theorem 2.2 we have:

$$L(f,\mathcal{P}) \leq L(f,\mathcal{Q}) \leq U(f,\mathcal{Q}) \leq U(f,\mathcal{P}'),$$

which gives the result. □

Exercise 4.1. Let $A \subset \mathbb{R}^n$, suppose $f, g : A \to \mathbb{R}$ are bounded, and let $\lambda \geq 0$. Show that:

a) $\sup_{x\in A]} -f(x) = -\inf_{x\in A} f(x)$,

b) $\inf_{x\in A} -f(x) = -\sup_{x\in A} f(x)$,

c) $\sup_{x\in A} \lambda f(x) = \lambda \sup_{x\in A} f(x)$,

d) $\sup_{x\in A} (f(x) + g(x)) \leq \sup_{x\in A} f(x) + \sup_{x\in A} g(x)$,

e) $\inf_{x\in A} (f(x) + g(x)) \geq \inf_{x\in A} f(x) + \inf_{x\in A} g(x)$,

f) $\inf_{x\in A} \lambda f(x) = \lambda \inf_{x\in A} f(x)$,

g) $\left|\sup_{x\in A} f(x)\right| \leq \sup_{x\in A} |f(x)|$,

h) $\left|\inf_{x\in A} f(x)\right| \leq \sup_{x\in A} |f(x)|$.

Exercise 4.2. Let $[a, b] \subset \mathbb{R}$ be any finite interval and $\mathcal{P}$ be any partition of $[a, b]$. Suppose $f, g : [a, b] \to \mathbb{R}$ are bounded functions, and that $\lambda \geq 0$. Show that:

a) $U(-f, \mathcal{P}) = -L(f, \mathcal{P})$,

b) $L(-f, \mathcal{P}) = -U(f, \mathcal{P})$,

c) $U(\lambda f, \mathcal{P}) = \lambda U(f, \mathcal{P})$,

d) $L(\lambda f, \mathcal{P}) = \lambda L(f, \mathcal{P})$,

e) $U(f + g, \mathcal{P}) \leq U(f, \mathcal{P}) + U(g, \mathcal{P})$,

f) $L(f + g, \mathcal{P}) \geq L(f, \mathcal{P}) + L(g, \mathcal{P})$.

Exercise 4.3. Suppose that $\mathcal{P} \preceq \mathcal{Q}$. Show that:

$$0 \leq U(f, \mathcal{Q}) - L(f, \mathcal{Q}) \leq U(f, \mathcal{P}) - L(f, \mathcal{P})$$

2.1.3 The Darboux Integral

We are now ready to define the integral:

Definition 2.1. The *lower* and *upper Darboux integrals* are defined as follows:

$$\underline{\int_a^b} f(x)dx := \sup_{\mathcal{P}} L(f, \mathcal{P}) \qquad \overline{\int_a^b} f(x)dx := \inf_{\mathcal{P}} U(f, \mathcal{P})$$

where the supremum and infimum are taken over all partitions of $[a, b]$. By Corollary 2.3 we have that:

$$\underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx.$$

We say that a bounded function $f : [a, b] \to \mathbb{R}$ is *Darboux integrable*, or simply *integrable*,[1] if it is bounded, and the upper and lower Darboux integrals are equal. In this case, we write:

$$\int_a^b f(x)dx := \underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx.$$

Notice that the lower and upper Darboux integrals are always well defined for a bounded function f.

As a notational convention, x appears in the integral as a 'dummy variable'. We can replace x with any other letter and the meaning is unchanged:

$$\int_a^b f(x)dx = \int_a^b f(t)dt = \int_a^b f(\alpha)d\alpha, \quad \text{etc.}$$

Example 2.1. Suppose $f : [a, b] \to \mathbb{R}$ is the constant function $f(x) = c$. Then we find, for any partition $\mathcal{P}$ that $m_j = M_j = c$ and so:

$$L(f, \mathcal{P}) = U(f, \mathcal{P}) = \sum_{j=0}^{k-1} c\Delta x_j = c\sum_{j=0}^{k-1} (x_{j+1} - x_j) = c(a - b).$$

[1] If you continue with your study of analysis you will discover that there are other inequivalent definitions of integrability, in which case one has to distinguish what type of integrability one refers to.

as the sum is a telescoping sum. Thus f is integrable and:

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

Example 2.2. Suppose $f : [-1, 1] \to \mathbb{R}$ is given by:

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

For $l \in \mathbb{N}$, let us consider the partition of $[-1, 1]$ given by $\mathcal{P}_l = (-1, -2^{-l}, 2^{-l}, 1)$. For the intervals I= $[-1, -2^{-l}], [2^{-l}, 1]$ we have that:

$$\sup_{x\in I} f(x) = \inf_{x\in I} f(x) = 0,$$

so these intervals do not contribute to either the upper or lower sums. For the interval $I = [-2^{-l}, 2^{-l}]$, we have:

$$\inf_{x\in I} f(x) = 0, \quad \sup_{x\in I} f(x) = 1, \quad \Delta x = 2^{-l+1},$$

so that:

$$L(f, \mathcal{P}_l) = 0, \qquad U(f, \mathcal{P}_l) = 2^{-l+1}.$$

We conclude that[2]:

$$\overline{\int_{-1}^{1}} f(x)dx = \inf_{\mathcal{P}} U(f, \mathcal{P}) \leq 0, \qquad \underline{\int_{-1}^{1}} f(x)dx = \sup_{\mathcal{P}} L(f, \mathcal{P}) \geq 0,$$

from which, since:

$$\underline{\int_{-1}^{1}} f(x)dx \leq \overline{\int_{a}^{b}} f(x)dx,$$

we deduce that f is integrable, and

$$\int_{-1}^{1} f(x)dx = 0.$$

Example 2.3. Consider the Dirichlet function $f : [0, 1] \to \mathbb{R}$:

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q}, \\ 0 & x \notin \mathbb{Q}. \end{cases}$$

If $I \subset [0, 1]$ is any interval of positive length, then $\sup_{x\in I} f(x) = 1$ and $\inf_{x\in I} f(x) = 0$, so that for any partition $\mathcal{P}$ we have $M_j = 1$, $m_j = 0$. We conclude that:

$$L(f, \mathcal{P}) = 0, \qquad U(f, \mathcal{P}) = 1.$$

Thus, we see that this function is *not* integrable.

[2] Note carefully that we cannot immediately conclude that both are zero.

An important property of the integral is that it is linear:

Theorem 2.4 (Linearity of the integral)**.** *Suppose that $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ are both integrable, and that $\lambda \in \mathbb{R}$. Then $f + \lambda g$ is integrable and we have:*

$$\int_a^b (f(x) + \lambda g(x))\, dx = \int_a^b f(x)dx + \lambda \int_a^b g(x)dx.x$$

Proof. We shall proceed in two stages: first we will show that if g is integrable, then so is $-g$. Then we will establish that $f + \lambda g$ is integrable for $\lambda \geq 0$. The full result will follow from combining these two facts.

1. Note that for any partition, $L(-g, \mathcal{P}) = -U(g, \mathcal{P})$, which implies:

$$\underline{\int_a^b} - g(x)dx = \sup_{\mathcal{P}} L(-g, \mathcal{P}) = \sup_{\mathcal{P}} -U(g, \mathcal{P}) = -\inf_{\mathcal{P}} U(g, \mathcal{P}) = -\int_a^b g(x)dx$$

   Similarly,

$$\overline{\int_a^b} - g(x)dx = \inf_{\mathcal{P}} U(-g, \mathcal{P}) = \inf_{\mathcal{P}} -L(g, \mathcal{P}) = -\sup_{\mathcal{P}} U(g, \mathcal{P}) = -\int_a^b g(x)dx,$$

   so that $-g$ is integrable and:

$$\int_a^b -g(x)dx = -\int_a^b g(x)dx.$$

2. Now, suppose $\lambda \geq 0$, then for any partition $\mathcal{P}$, we have by Exercise 4.2 that:

$$L(f, \mathcal{P}) + \lambda L(g, \mathcal{P}) \leq L(f + \lambda g, \mathcal{P}),$$

   From the definition of the supremum we of course have:

$$L(f + \lambda g, \mathcal{P}) \leq \sup_{\mathcal{Q}} L(f + \lambda g, \mathcal{Q}),$$

   so that:

$$L(f, \mathcal{P}) + \lambda L(g, \mathcal{P}) \leq \sup_{\mathcal{Q}} L(f + \lambda g, \mathcal{Q}).$$

   Now, suppose that $\mathcal{P}$, $\mathcal{P}'$ are any two partitions and $\mathcal{R}$ is their common refinement. Then

$$L(f, \mathcal{P}) + \lambda L(g, \mathcal{P}') \leq L(f, \mathcal{R}) + \lambda L(g, \mathcal{R}).$$

   Putting these estimates together gives:

$$L(f, \mathcal{P}) + \lambda L(g, \mathcal{P}') \leq \sup_{\mathcal{Q}} L(f + \lambda g, \mathcal{Q}),$$

   for any pair of partitions $\mathcal{P}$, $\mathcal{P}'$. Taking the supremum over $\mathcal{P}, \mathcal{P}'$ we have:

$$\sup_{\mathcal{P}} L(f, \mathcal{P}) + \lambda \sup_{\mathcal{P}'} L(g, \mathcal{P}') \leq \sup_{\mathcal{Q}} L(f + \lambda g, \mathcal{Q}),$$

or equivalently:

$$\int_a^b f(x)dx + \lambda \int_a^b g(x)dx \leq \underline{\int_a^b} [f(x) + \lambda g(x)]\, dx.$$

Applying this estimate to $-f$ and $-g$ we have:

$$\sup_{\mathcal{P}} L(-f, \mathcal{P}) + \lambda \sup_{\mathcal{P}'} L(-g, \mathcal{P}') \leq \sup_{\mathcal{Q}} L(-\left(f + \lambda g\right), \mathcal{Q})$$

which implies:

$$\sup_{\mathcal{P}} -U(f, \mathcal{P}) + \lambda \sup_{\mathcal{P}'} -U(g, \mathcal{P}') \leq \sup_{\mathcal{Q}} -U(f + \lambda g, \mathcal{Q})$$

and thus:

$$-\inf_{\mathcal{P}} U(f, \mathcal{P}) - \lambda \inf_{\mathcal{P}'} U(g, \mathcal{P}') \leq -\inf_{\mathcal{Q}} U(f + \lambda g, \mathcal{Q})$$

so that:

$$\overline{\int_a^b} [f(x) + \lambda g(x)]\, dx \leq \int_a^b f(x)dx + \lambda \int_a^b g(x)dx.$$

Since by definition:

$$\underline{\int_a^b} [f(x) + \lambda g(x)]\, dx \leq \overline{\int_a^b} [f(x) + \lambda g(x)]\, dx,$$

we conclude that

$$\underline{\int_a^b} [f(x) + \lambda g(x)]\, dx = \overline{\int_a^b} [f(x) + \lambda g(x)]\, dx = \int_a^b f(x)dx + \lambda \int_a^b g(x)dx,$$

which is the result we require. Combining these two cases gives the general result.

□

Before we establish some further results concerning integrable functions, we need a useful lemma.

Lemma 2.5. *A function $f : [a, b] \to \mathbb{R}$ is integrable if and only if for any $\epsilon > 0$ there exists a partition $\mathcal{P}$ such that:*

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon. \tag{2.1}$$

Proof. Note that if $\mathcal{P}$ is any partition, from the definition of the supremum and infimum we have:

$$\sup_{\mathcal{Q}} L(f, \mathcal{Q}) \geq L(f, \mathcal{P}), \qquad \inf_{\mathcal{Q}} U(f, \mathcal{Q}) \leq U(f, \mathcal{P}).$$

Combining these two inequalities, we have that for any $\mathcal{P}$:

$$0 \leq \inf_{\mathcal{Q}} U(f, \mathcal{Q}) - \sup_{\mathcal{Q}} L(f, \mathcal{Q}) < U(f, \mathcal{P}) - L(f, \mathcal{P})$$

Now suppose that for every $\epsilon > 0$, there exists a partition $\mathcal{P}$ such that (2.1) holds. Then for any ϵ we have:

$$0 \leq \inf_{\mathcal{Q}} U(f, \mathcal{Q}) - \sup_{\mathcal{Q}} L(f, \mathcal{Q}) < \epsilon,$$

which implies (since ϵ is arbitrary):

$$0 = \inf_{\mathcal{Q}} U(f, \mathcal{Q}) - \sup_{\mathcal{Q}} L(f, \mathcal{Q}) = \overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx$$

so that f is integrable.

Conversely, suppose that f is integrable, so that $\inf_{\mathcal{Q}} U(f, \mathcal{Q}) = \sup_{\mathcal{Q}} L(f, \mathcal{Q}) = I$. Then from the properties of the supremum and infimum we can find $\mathcal{P}, \mathcal{P}'$ such that:

$$U(f, \mathcal{P}) < I + \frac{\epsilon}{2}, \qquad L(f, \mathcal{P}') > I - \frac{\epsilon}{2}$$

Taking $\mathcal{R}$ to be the common refinement of $\mathcal{P}, \mathcal{P}'$, we deduce:

$$U(f, \mathcal{R}) < I + \frac{\epsilon}{2}, \qquad L(f, \mathcal{R}) > I - \frac{\epsilon}{2},$$

which implies:

$$U(f, \mathcal{R}) - L(f, \mathcal{R}) < \epsilon.$$

□

Note that by Exercise 4.3, if $\mathcal{P}$ satisfies (2.1), then so does any partition $\mathcal{Q}$ with $\mathcal{P} \preceq \mathcal{Q}$. This formulation is often more useful to actually work with, as we shall now see, when we consider additivity of the integral. It is important to know that if we break up an interval into two or more subintervals, then the integral behaves as we would expect.

Theorem 2.6 (Additivity of the integral)**.** *Suppose $f : [a, b] \to \mathbb{R}$, and let $c \in (a, b)$. Then f is integrable if and only if $f|_{[a,c]}$ and $f|_{[c,b]}$ are integrable, and:*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

Proof. 1. First suppose $f : [a, b] \to \mathbb{R}$ is integrable. Fix $\epsilon > 0$. By Lemma 2.5 there exists a partition of $[a, b]$, $\mathcal{P}$ such that:

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon.$$

By refining $\mathcal{P}$ by adding the point c if necessary, we can assume that c is an endpoint of $\mathcal{P}$, so that:

$$\mathcal{P} = (x_0, \ldots, x_l, c, x_{l+1}, \ldots, x_k).$$

We have, by Exercise 4.5

$$L(f, \mathcal{P}) = L(f|_{[a,c]}, \mathcal{P}_L) + L(f|_{[c,b]}, \mathcal{P}_R),$$
$$U(f, \mathcal{P}) = U(f|_{[a,c]}, \mathcal{P}_L) + U(f|_{[c,b]}, \mathcal{P}_R),$$

where

$$\mathcal{P}_L = (x_0, \ldots, x_l, c), \qquad \mathcal{P}_R = (c, x_{l+1}, \ldots, x_k).$$

We obtain:

$$\begin{aligned}&\Big[U(f|_{[a,c]}, \mathcal{P}_L) - L(f|_{[a,c]}, \mathcal{P}_L)\Big] + \Big[U(f|_{[c,b]}, \mathcal{P}_R) - L(f|_{[c,b]}, \mathcal{P}_R)\Big] \\ &\quad = U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon.\end{aligned}$$

Now, both terms in square brackets are non-negative, so we conclude that:

$$U(f|_{[a,c]}, \mathcal{P}_L) - L(f|_{[a,c]}, \mathcal{P}_L) < \epsilon, \qquad U(f|_{[c,b]}, \mathcal{P}_R) - L(f|_{[c,b]}, \mathcal{P}_R) < \epsilon,$$

which implies that indeed $f|_{[a,c]}$ and $f|_{[c,b]}$ are integrable.

2. Now conversely, suppose $f|_{[a,c]}$ and $f|_{[c,b]}$ are integrable. Fix $\epsilon > 0$. Then there exist partitions $\mathcal{P}_\mathcal{L}$ of $[a,c]$ and $\mathcal{P}_\mathcal{R}$ of $[c,b]$ such that:

$$U(f|_{[a,c]}, \mathcal{P}_L) - L(f|_{[a,c]}, \mathcal{P}_L) < \frac{\epsilon}{2}, \qquad U(f|_{[c,b]}, \mathcal{P}_R) - L(f|_{[c,b]}, \mathcal{P}_R) < \frac{\epsilon}{2}$$

Now, we form a partition $\mathcal{P}$ of $[a,b]$ by concatenating $\mathcal{P}_L$ and $\mathcal{P}_R$. We again have that:

$$\begin{aligned}L(f, \mathcal{P}) &= L(f|_{[a,c]}, \mathcal{P}_L) + L(f|_{[c,b]}, \mathcal{P}_R), \\ U(f, \mathcal{P}) &= U(f|_{[a,c]}, \mathcal{P}_L) + U(f|_{[c,b]}, \mathcal{P}_R),\end{aligned}$$

so

$$\begin{aligned}U(f, \mathcal{P}) - L(f, \mathcal{P}) = &\Big[U(f|_{[a,c]}, \mathcal{P}_L) - L(f|_{[a,c]}, \mathcal{P}_L)\Big] \\ &+ \Big[U(f|_{[c,b]}, \mathcal{P}_R) - L(f|_{[c,b]}, \mathcal{P}_R)\Big] < \epsilon.\end{aligned}$$

Which implies f is integrable. We further deduce that:

$$\begin{aligned}\int_a^b f(x)dx &\leq U(f, \mathcal{P}) = U(f|_{[a,c]}, \mathcal{P}_L) + U(f|_{[c,b]}, \mathcal{P}_R) \\ &< \int_a^c f(x)dx + \int_c^b f(x)dx + \epsilon,\end{aligned}$$

Similarly, we have:

$$\begin{aligned}\int_a^b f(x)dx &\geq L(f, \mathcal{P}) = L(f|_{[a,c]}, \mathcal{P}_L) + L(f|_{[c,b]}, \mathcal{P}_R) \\ &> \int_a^c f(x)dx + \int_c^b f(x)dx - \epsilon.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we conclude that:

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

□

It is useful to extend slightly our definition of the integral to allow for the case where $a \geq b$. This is purely a convention, but we define:

$$\int_a^b f(x)dx := -\int_b^a f(x)dx \quad \text{for } a > b, \qquad \int_a^a f(x)dx := 0.$$

The next Lemma we shall require is slightly technical, but important.

Lemma 2.7. *If $A \subset \mathbb{R}^n$, and $f : A \to \mathbb{R}$, then:*

$$\sup_{x\in A} |f(x)| - \inf_{x\in A} |f(x)| \leq \sup_{x\in A} f(x) - \inf_{x\in A} f(x).$$

Proof. We consider three possible cases:

a) $f(x) \geq 0$ for all $x \in A$. In this case, $|f(x)| = f(x)$ and we have

$$\sup_{x\in A} |f(x)| - \inf_{x\in A} |f(x)| = \sup_{x\in A} f(x) - \inf_{x\in A} f(x).$$

b) $f(x) \leq 0$ for all $x \in A$. In this case, $|f(x)| = -f(x)$ and $\sup_{x\in A} |f(x)| = -\inf_{x\in A} f(x)$, $\inf_{x\in A} |f(x)| = -\sup_{x\in A} f(x)$ so again:

$$\sup_{x\in A} |f(x)| - \inf_{x\in A} |f(x)| = \sup_{x\in A} f(x) - \inf_{x\in A} f(x).$$

c) Finally, suppose $\inf_{x\in A} f(x) < 0 < \sup_{x\in I} f(x)$. Then

$$\sup_{x\in A} |f(x)| = \max\{\sup_{x\in A} f(x), -\inf_{x\in A} f(x)\}$$

where both terms are positive, and we have:

$$\sup_{x\in A} |f(x)| - \inf_{x\in A} |f(x)| \leq \sup_{x\in A} |f(x)| \leq \sup_{x\in A} f(x) - \inf_{x\in A} f(x)$$

□

Theorem 2.8. *Suppose $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ are both integrable. Then so is:*

i) $|f|$,

ii) f^2,

iii) fg.

Proof. i) Fix $\epsilon > 0$. Since f is integrable, there exists a partition $\mathcal{P} = (x_0, \ldots, x_k)$ such that:

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) = \sum_{j=0}^{k-1} \left[\sup_{x\in[x_j,x_{j+1}]} f(x) - \inf_{x\in[x_j,x_{j+1}]} f(x) \right] \Delta x_j < \epsilon$$

Now let us estimate:

$$\begin{aligned} U(|f|,\mathcal{P}) - L(|f|,\mathcal{P}) &= \sum_{j=0}^{k-1} \left[\sup_{x\in[x_j,x_{j+1}]} |f(x)| - \inf_{x\in[x_j,x_{j+1}]} |f(x)| \right] \Delta x_j \\ &\leq \sum_{j=0}^{k-1} \left[\sup_{x\in[x_j,x_{j+1}]} f(x) - \inf_{x\in[x_j,x_{j+1}]} f(x) \right] \Delta x_j \\ &< \epsilon, \end{aligned}$$

where we have used Lemma 2.7 to go from the first line to the second. We conclude that $|f|$ is integrable.

ii) Since f is integrable, it is bounded, say $|f(x)| \leq M$. Note that:

$$\left|f(x)^2 - f(y)^2\right| = |f(x) - f(y)|\,|f(x) + f(y)| \leq 2M\,|f(x) - f(y)|\,.$$

Since this holds for any x, y, in particular we conclude that for any interval $I \subset \mathbb{R}$:

$$\sup_{x\in I} f(x)^2 - \inf_{y\in I} f(y)^2 \leq 2M\left(\sup_{x\in I} f(x) - \inf_{y\in I} f(y)\right).$$

Now, fix $\epsilon > 0$. Since f is integrable, there exists a partition $\mathcal{P}$ such that:

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) = \sum_{j=0}^{k-1} \left[\sup_{x\in[x_j,x_{j+1}]} f(x) - \inf_{x\in[x_j,x_{j+1}]} f(x) \right] \Delta x_j < \frac{\epsilon}{1+2M}.$$

We can estimate:

$$\begin{aligned} U(f^2,\mathcal{P}) - L(f^2,\mathcal{P}) &= \sum_{j=0}^{k-1} \left[\sup_{x\in[x_j,x_{j+1}]} f(x)^2 - \inf_{x\in[x_j,x_{j+1}]} f(x)^2 \right] \Delta x_j \\ &\leq 2M \sum_{j=0}^{k-1} \left[\sup_{x\in[x_j,x_{j+1}]} f(x) - \inf_{x\in[x_j,x_{j+1}]} f(x) \right] \Delta x_j \\ &< \frac{2M}{1+2M}\epsilon < \epsilon. \end{aligned}$$

We conclude that f^2 is integrable.

iii) Note that:

$$fg = \frac{1}{4}\left[(f+g)^2 - (f-g)^2\right]$$

Applying previous results, we can see that the right-hand side is integrable, hence so is the left-hand side.

□

Exercise 4.4. Let $A \subset [-1,1]$ be a *finite* set, and let $\chi_A : [-1,1] \to \{0,1\}$ be the *characteristic function* of A. That is:

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

Show that χ_A is integrable, and:

$$\int_{-1}^{1} \chi_A(x)dx = 0.$$

Exercise 4.5. Suppose that $a < c < b$ and suppose $f : [a,b] \to \mathbb{R}$ is a bounded function. Let $\mathcal{P}$ be a partition of $[a,b]$ of the form:

$$\mathcal{P} = (a, x_1, \ldots, x_{l-1}, x_l = c, x_{l+1}, \ldots, x_{k-1}, b).$$

and define $\mathcal{P}_L$ and $\mathcal{P}_R$ to be partitions of $[a,c]$ and $[c,b]$ respectively, given by:

$$\mathcal{P}_L = (a, x_1, \ldots, x_{l-1}, c), \qquad \mathcal{P}_R = (c, x_{l+1}, \ldots, x_{k-1}, b).$$

Show that:

$$L(f,\mathcal{P}) = L(f|_{[a,c]}, \mathcal{P}_L) + L(f|_{[c,b]}, \mathcal{P}_R),$$
$$U(f,\mathcal{P}) = U(f|_{[a,c]}, \mathcal{P}_L) + U(f|_{[c,b]}, \mathcal{P}_R).$$

Exercise 4.6. Suppose $f : [a,b] \to \mathbb{R}$, $g : [a,b] \to \mathbb{R}$ are integrable.

a) Show that:

$$(b-a) \inf_{x\in[a,b]} f(x) \leq \int_a^b f(x)dx \leq (b-a) \sup_{x\in[a,b]} f(x).$$

b) Establish the estimate:

$$\left| \int_a^b f(x)dx \right| \leq (b-a) \sup_{x\in[a,b]} |f(x)|$$

c) Show that if $0 \leq f(x)$ for all $x \in [a,b]$ then:

$$0 \leq \int_a^b f(x)dx$$

d) Show that if $f(x) \leq g(x)$ for all $x \in [a,b]$ then:

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx$$

e) Prove that:

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|\, dx$$

Exercise 4.7 (*). Consider Thomae's function $f : [0, 1] \to \mathbb{R}$:

$$f(x) = \begin{cases} 1 & x = 0, \\ \frac{1}{q} & x = \frac{p}{q} \in \mathbb{Q}, \text{ where } \mathrm{hcf}(p, q) = 1, \\ 0 & x \notin \mathbb{Q}. \end{cases}$$

Show that f is integrable.

2.1.4 Integrability of continuous functions

So far, we've only positively established that constant functions are integrable. Now we shall establish the important result that all continuous functions are integrable. This means that the set of all integrable functions is big enough to be interesting.

Theorem 2.9. *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$. Then f is integrable.*

Proof. Fix $\epsilon > 0$. Since f is continuous on $[a, b]$, it is uniformly continuous. Thus there exists $\delta > 0$ such that for all $x, y \in [a, b]$ with $|x - y| < \delta$ we have $|f(x) - f(y)| < \frac{1}{2(b-a)}\epsilon$. Pick a partition $\mathcal{P} = (x_0, \ldots, x_k)$ such that $\Delta x_j < \delta$ for all $j = 0, \ldots k - 1$. We can do this by (for example) setting:

$$x_j = a + \frac{b-a}{k} j, \qquad j = 0, \ldots, k,$$

where k is chosen sufficiently large that $(b - a) < k\delta$. Now, by construction, for $x, y \in [x_j, x_{j+1}]$ we have:

$$|f(x) - f(y)| < \frac{1}{2(b-a)}\epsilon,$$

which implies that:

$$\sup_{x \in [x_j, x_{j+1}]} f(x) - \inf_{x \in [x_j, x_{j+1}]} f(x) \le \frac{1}{2(b-a)}\epsilon < \frac{1}{b-a}\epsilon.$$

Now, we can estimate:

$$\begin{aligned} U(f, \mathcal{P}) - L(f, \mathcal{P}) &= \sum_{j=0}^{k-1} \left[\sup_{x \in [x_j, x_{j+1}]} f(x) - \inf_{x \in [x_j, x_{j+1}]} f(x) \right] \Delta x_j \\ &< \sum_{j=0}^{k-1} \frac{\epsilon}{b-a}(x_{j+1} - x_j) = \epsilon, \end{aligned}$$

as we have a telescoping sum. Thus for any $\epsilon > 0$, we have found a partition $\mathcal{P}$ such that:

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon,$$

which implies f is integrable. □

Remark. Note that for the proof to go through it is in fact sufficient for f to be uniformly continuous on (a, b). In view of Example 2.2 we can always modify f at the endpoints without affecting the integrability or the value of the integral.

Exercise 5.1. Suppose that $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, and satisfies:

$$\int_a^b f(t)dt = 0, \qquad 0 \le f(x), \text{ for all } x \in [a, b].$$

a) Show that $f(x) = 0$, for all $x \in [a, b]$.

b) Does the result hold if we do not assume f is continuous?

2.1.5 Riemann sums

While we've introduced a definition of what it means for a function to be integrable, and we've defined the integral, in practice it is useful to introduce a slightly different approach to find the integral, the Riemann sum. We will need to first refine slightly our criterion for integrability.

Lemma 2.10. *Suppose $f : [a, b] \to \mathbb{R}$ is integrable. Then given $\epsilon > 0$, there exists $\delta > 0$ such that for any partition $\mathcal{P}$ satisfying* $\operatorname{mesh}\mathcal{P} < \delta$ *we have:*

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon.$$

(*) Proof. Since f is integrable, we have $|f(x)| \le M$ for some M, and we know that there exists a partition $\mathcal{Q} = (y_0, \dots, y_m)$ such that:

$$U(f, \mathcal{Q}) - L(f, \mathcal{Q}) < \frac{\epsilon}{4}.$$

Suppose that $\mathcal{P}$ is any partition satisfying mesh $\mathcal{P} < \delta$ and consider $\mathcal{R}$ to be the common refinement of $\mathcal{P}, \mathcal{Q}$. Since a, b are common endpoints, there are at most $m - 1$ endpoints of $\mathcal{Q}$ that are distinct from the endpoints of $\mathcal{P}$. Therefore, at most $m - 1$ intervals in $\mathcal{P}$ contain endpoints in $\mathcal{Q}$. Those intervals of $\mathcal{P}$ which do not contain an endpoint in $\mathcal{Q}$ are left unchanged in $\mathcal{R}$. Let us consider:

$$L(f, \mathcal{R}) - L(f, \mathcal{P})$$

Those terms in the sum involving intervals which are common to $\mathcal{P}$, $\mathcal{R}$ will cancel. There will remain at most $3(m - 1)$ contributions to consider. Each interval has length at most δ, and we have $|f(x)| \le M$, so that the contribution to the sums from these intervals is at most $3(m - 1)\delta M$, so that we find:

$$L(f, \mathcal{R}) - L(f, \mathcal{P}) \le 3(m - 1)\delta M < \frac{\epsilon}{4},$$

where we take $\delta = \frac{\epsilon}{20m(M+1)}$. An identical argument leads to the estimate:

$$U(f, \mathcal{P}) - U(f, \mathcal{R}) < \frac{\epsilon}{4}.$$

Now, since $\mathcal{R}$ is a refinement of $\mathcal{Q}$, we have:

$$L(f, \mathcal{P}) > L(f, \mathcal{R}) - \frac{\epsilon}{4} > L(f, \mathcal{Q}) - \frac{\epsilon}{4} > U(f, \mathcal{Q}) - \frac{\epsilon}{2},$$

and similarly

$$U(f,\mathcal{P}) < U(f,\mathcal{R}) + \frac{\epsilon}{4} < U(f,\mathcal{Q}) + \frac{\epsilon}{4} < L(f,\mathcal{Q}) + \frac{\epsilon}{2},$$

so that we have:

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) - \epsilon < L(f,\mathcal{Q}) - U(f,\mathcal{Q}) \leq 0$$

which gives us that:

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < \epsilon.$$

for any $\mathcal{P}$ with mesh $\mathcal{P} < \delta$. □

Definition 2.2. A *tagged partition* $(\mathcal{P},\tau)$ is a partition $\mathcal{P} = (x_0 \dots, x_k)$, together with a choice of k points $\tau = (t_0, \dots, t_{k-1})$ such that $t_j \in [x_j, x_{j+1}]$ for $j = 0, \dots, k-1$. The *mesh* of the tagged partition is again the length of the longest subinterval, $\max_{j=0,\dots,k-1} \Delta x_j$.

If $f : [a,b] \to \mathbb{R}$ is an integrable function, we introduce the Riemann sum of f with respect to the tagged partition:

$$R(f,\mathcal{P},\tau) = \sum_{j=0}^{k-1} f(t_j)\Delta x_j$$

Clearly, since for $t \in [x_j, x_{j+1}]$ we have:

$$m_j \leq f(t) \leq M_j$$

we deduce:

$$L(f,\mathcal{P}) \leq R(f,\mathcal{P},\tau) \leq U(f,\mathcal{P}).$$

Theorem 2.11. *Let $f : [a,b] \to \mathbb{R}$ be an integrable function. Suppose $((\mathcal{P}_l,\tau_l))_{l=0}^{\infty}$ is a sequence of tagged partitions with mesh $\mathcal{P}_l \to 0$ as $l \to \infty$. Then we have:*

$$R(f,\mathcal{P}_l,\tau_l) \to \int_a^b f(x)dx.$$

Proof. Fix $\epsilon > 0$. By Lemma 2.10 there exists $\delta > 0$ such that for any partition $\mathcal{P}$ satisfying mesh $\mathcal{P} < \delta$ we have:

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < \frac{\epsilon}{2}.$$

Since mesh $\mathcal{P}_l \to 0$ as $l \to \infty$, there exists $N \in \mathbb{N}$ such that for all $l \geq N$ we have mesh $\mathcal{P}_l < \delta$, so we have:

$$\int_a^b f(x) - \frac{\epsilon}{2} < L(f,\mathcal{P}_l) \leq R(f,\mathcal{P}_l,\tau_l) \leq U(f,\mathcal{P}_l) < \int_a^b f(x) + \frac{\epsilon}{2},$$

which implies:

$$\left| R(f,\mathcal{P}_l,\tau_l) - \int_a^b f(x) \right| < \epsilon.$$ □

Example 2.4. Let us use the Riemann sum approach to calculate:

$$\int_0^1 x dx.$$

We know that $f : x \mapsto x$ is integrable, because it is continuous. We will take as our sequence of tagged partitions the uniform partitions with tags at the left endpoint:

$$\mathcal{P}_l = \left(0, \frac{1}{l}, \frac{2}{l}, \ldots, \frac{l-1}{l}, 1\right), \qquad \mathcal{T}_l = \left(0, \frac{1}{l}, \frac{2}{l}, \ldots, \frac{l-1}{l}\right)$$

The Riemann sum is:

$$R(f, \mathcal{P}_l, \mathcal{T}_l) = \sum_{j=0}^{l-1} \frac{j}{l} \times \frac{1}{l} = \frac{1}{l^2} \sum_{j=0}^{l-1} j = \frac{1}{l^2} \times \frac{l(l-1)}{2} = \frac{1}{2} - \frac{1}{2l^2}.$$

We have that:

$$R(f, \mathcal{P}_l, \mathcal{T}_l) \to \frac{1}{2} = \int_0^1 x dx,$$

as $l \to \infty$, which agrees with what we would expect from considering the area under the graph.

Exercise 5.2. By choosing a suitable sequence of partitions to approximate the integral by a Riemann sum, show that:

$$\int_0^1 x^2 dx = \frac{1}{3}.$$

Exercise 5.3. By rewriting the left-hand side as a Riemann sum, show that:

$$\lim_{k \to \infty} \sum_{j=0}^{k-1} \frac{k}{j^2 + k^2} = \int_0^1 \frac{1}{1+x^2} dx.$$

2.2 The fundamental theorem of calculus

Of course, when it comes to actually calculating integrals, you know that in practise the best approach is to find the anti-derivative of the integrand. The fact that this procedure works and gives the same answer as finding the area under the graph is non-trivial, and goes by the name of the *Fundamental Theorem of Calculus.*

Lemma 2.12. *Suppose that $f : [a, b] \to \mathbb{R}$ is integrable, and define the function $F : [a, b] \to \mathbb{R}$ by:*

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

Then F is uniformly continuous on $[a, b]$.

Proof. Since f is integrable, by Theorem 2.6 the function F is well defined. As f is bounded, there exists M such that $|f(x)| \leq M$. For $x, y \in [a,b]$ we can estimate by Exercise 4.6 *b*):

$$|F(x) - F(y)| = \left| \int_x^y f(t)dt \right| \leq |x-y|\, M$$

so if we take $|x-y| < \frac{\epsilon}{1+M}$, we have $|F(x) - F(y)| < \epsilon$, which is the statement that $F : [a,b] \to \mathbb{R}$ is uniformly continuous. □

Theorem 2.13 (Fundamental Theorem of Calculus)**.** *Suppose that* $f : [a,b] \to \mathbb{R}$ *is continuous, and define the function* $F : [a,b] \to \mathbb{R}$ *by:*

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

Then F *is differentiable on* (a,b)*, and:*

$$F'(x) = f(x).$$

Proof. First note the following two facts, which hold for $x, y \in [a,b]$ with $x \neq y$:

$$F(y) - F(x) = \int_x^y f(t)dt, \qquad \frac{1}{y-x}\int_x^y 1dt = 1.$$

These results follow from Theorem 2.6 and Example 2.1 respectively.

Now fix $\epsilon > 0$ and $x \in (a,b)$. Since f is continuous at x, there exists $\delta > 0$ such that if $|y-x| < \delta$ we have $|f(x) - f(y)| < \epsilon$. Suppose $y \neq x$ satisfies this condition. Let us estimate:

$$\begin{aligned}\left| \frac{F(y)-F(x)}{y-x} - f(x) \right| &= \left| \frac{1}{y-x}\int_x^y f(t)dt - f(x) \right| \\ &= \left| \frac{1}{y-x}\int_x^y \left(f(t) - f(x)\right)dt \right|\end{aligned}$$

Now, for t between x and y, we know that $|x-t| < \delta$, so $|f(t) - f(x)| < \epsilon$, and we can further estimate:

$$\begin{aligned}\left| \frac{1}{y-x}\int_x^y \left(f(t) - f(x)\right)dt \right| &= \frac{1}{|y-x|}\left| \int_x^y f(t) - f(x)dt \right| \\ &\leq \frac{1}{|y-x|}\,|y-x| \sup_{t\in[x,y]} |f(t) - f(x)| < \epsilon.\end{aligned}$$

This is precisely the condition that:

$$\lim_{y\to x} \frac{F(y)-F(x)}{y-x} = f(x).$$

□

Remark. Again, it is actually sufficient for $f : (a,b) \to \mathbb{R}$ to be uniformly continuous, rather than requiring continuity on $[a,b]$.

An immediate corollary of this result is the following:

Corollary 2.14. *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous and that $F : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, differentiable on (a, b) and satisfies $F'(x) = f(x)$ for all $x \in (a, b)$. Then:*

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

Proof. Define the function $G : [a, b] \to \mathbb{R}$ by:

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

By the previous Theorem we have that G is continuous on $[a, b]$, differentiable on (a, b) and $G'(x) = F'(x)$ for $x \in (a, b)$. By the Mean Value Theorem we have that:

$$F(x) = G(x) + C,$$

for some constant $C \in \mathbb{R}$ (See Exercise A.6). We calculate:

$$\begin{aligned} F(b) - F(a) &= G(b) - G(a) \\ &= \int_a^b f(t)dt - \int_a^a f(t)dt \\ &= \int_a^b f(t)dt. \end{aligned}$$

□

This justifies the usual calculations that one does when computing integrals: it is enough to find an anti-derivative and then evaluate it at the endpoints of the integral.

Exercise 5.4. Define $f : [-1, 1] \to \mathbb{R}$ by:

$$f(x) := \begin{cases} 0 & x \leq 0, \\ 1 & x > 0. \end{cases}$$

a) Show that f is integrable, and compute:

$$F(x) = \int_{-1}^x f(t)dt$$

b) Show that $F : [-1, 1] \to \mathbb{R}$ is *not* differentiable at $x = 0$.

Exercise 5.5 (*)**.** Suppose $f : [a, b] \to \mathbb{R}$ is integrable and that $F : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$, differentiable on (a, b) and satisfies $F'(x) = f(x)$ for all $x \in (a, b)$.

a) Let $\mathcal{P} = (x_0, \ldots, x_k)$ be a partition. Show that for each subinterval $[x_j, x_{j+1}]$ with $j = 0, \ldots, k - 1$ there exists $t_j \in [x_j, x_{j+1}]$ such that:

$$F(x_{j+1}) - F(x_j) = f(t_j)\Delta x_j.$$

b) Show that for the tagged partition $(\mathcal{P}, \tau)$ with $\tau = (t_0, \ldots, t_{k-1})$ we have:

$$R(f, \mathcal{P}, \tau) = F(b) - F(a).$$

c) Deduce that the assumption in Corollary 2.14 that f is continuous may be weakened to integrability.

2.2.1 Change of variables formula

We can use the Fundamental Theorem of Calculus to prove a useful result, sometimes called the change of variables formula or substitution formula.

Theorem 2.15. *Let $I_1 := [\alpha, \beta], I_2 := [\gamma, \delta]$ be finite intervals. Suppose that $\phi : I_1 \to (\gamma, \delta)$ is a differentiable function with ϕ' continuous on (α, β). Suppose that $f : I_2 \to \mathbb{R}$ is continuous. Then if $[a, b] \subset (\alpha, \beta)$:*

$$\int_{\phi(a)}^{\phi(b)} f(t)dt = \int_a^b f(\phi(t))\phi'(t)dt.$$

Proof. Since $f(t)$ is continuous, by Theorem 2.13 there exists $F : [\gamma, \delta] \to \mathbb{R}$ with $F'(x) = f(x)$ for all $x \in (\gamma, \delta)$. By the chain rule, we have that the function $F \circ \phi : [\alpha, \beta] \to \mathbb{R}$ is differentiable on (α, β), with:

$$(F \circ \phi)'(x) = F'(\phi(x))\phi'(x), \qquad \text{for all } x \in (\alpha, \beta).$$

We now apply Corollary 2.14 to find:

$$\begin{aligned}\int_a^b f(\phi(t))\phi'(t)dt &= \int_a^b (F \circ \phi)'(t)dt \\ &= (F \circ \phi)(b) - (F \circ \phi)(a) = F(\phi(b)) - F(\phi(a)) \\ &= \int_{\phi(a)}^{\phi(b)} f(t)dt.\end{aligned}$$ □

The slight subtlety regarding the domains of definition in the statement of the theorem above is due to the fact that we need to ensure that ϕ' is continuous on the closed interval $[a, b]$ in order to conclude that the product $f(\phi(t))\phi'(t)$ is integrable on this interval, and to allow us to invoke the Fundamental Theorem of Calculus in the form that we have proved it.

Exercise 5.6. a) By making the change of variables given by $\phi(t) = \frac{1}{t}$, show that:

$$\int_a^b \frac{1}{1+t^2}dt = \int_{1/b}^{1/a} \frac{1}{1+t^2}dt,$$

for $0 < a < b$.

b) Setting $a = -1, b = 1$ in the above formula, we obtain:

$$\int_{-1}^{1} \frac{1}{1+x^2} dx = -\int_{-1}^{1} \frac{1}{1+x^2} dx.$$

What has gone wrong?

Exercise 5.7 (*)**.** Let $I_1 := [\alpha, \beta], I_2 := [\gamma, \delta]$ be finite intervals. Suppose that $\phi : I_1 \to I_2$ is continuous on I_1, and differentiable on (α, β) with ϕ' integrable on I_1. Suppose that $f : I_2 \to \mathbb{R}$ is continuous. Show that:

$$\int_{\phi(\alpha)}^{\phi(\beta)} f(t)dt = \int_{\alpha}^{\beta} f(\phi(t))\phi'(t)dt.$$

2.3 Interchanging limits

It is often the case that we wish to interchange various limits with integrals. For example, suppose that we have a sequence of integrable functions $f_j : [a, b] \to \mathbb{R}$ such that $f_j \to f$ in some appropriate sense. When can we say that:

$$\lim_{j\to\infty} \int_a^b f_j(x)dx = \int_a^b f(x)dx?$$

We have to be cautious in situations like this. Integration is a limiting operation, and experience teaches us that interchanging different limiting operations can lead to difficulties.

2.3.1 Pointwise convergence does not commute with integration

For $j \in \mathbb{N}$, consider the function $f_j : [0, 1] \to \mathbb{R}$ given by:

$$f_j(x) = \begin{cases} 2^{2j}x & 0 \leq x < 2^{-j}, \\ 2^j\left(2 - 2^j x\right) & 2^{-j} \leq x < 2^{-(j-1)} \\ 0 & x \geq 2^{-(j-1)}, \end{cases}$$

For each j, the graph of this function is an isosceles triangle, with base of length $2^{-(j-1)}$ and height 2^j. Therefore, we have:

$$\int_0^1 f_j(x)dx = 1.$$

However, we have that for all $x \in [0, 1]$:

$$\lim_{j\to\infty} f_j(x) = 0.$$

This is an example of a sequence of integrable functions for which $f_j(x) \to f(x)$ for all $x \in [a, b]$, but:

$$\lim_{j\to\infty} \int_a^b f_j(x)dx \neq \int_a^b f(x)dx.$$

This tells us that in general, pointwise convergence does *not* commute with integration.

Exercise 5.8. By modifying the example of Section 2.3.1, show that there exists a family $(f_j)_{j=0}^{\infty}$ of integrable functions $f_j : [0,1] \to \mathbb{R}$ such that:

$$f_j(x) \to 0 \qquad \text{for all } x \in [0,1],$$

while:

$$\int_0^1 f_j(x)dx \to \infty.$$

2.3.2 Uniform convergence does commute with integration

It may or may not surprise you to discover that the situation is much better if instead of pointwise convergence we instead consider *uniform* convergence. From our previous work we know that this would already be enough to rule out the bad behaviour inherent in our previous example.

Theorem 2.16. *Suppose that $(f_i)_{i=0}^{\infty}$ is a sequence of integrable functions $f : [a,b] \to \mathbb{R}$ such that $f_i \to f$ uniformly on $[a,b]$. Then f is integrable, and*

$$\lim_{i\to\infty} \int_a^b f_i(x)dx = \int_a^b f(x)dx.$$

(*) Proof. 1. Fix $\epsilon > 0$. By the uniform convergence of $(f_i)_{i=0}^{\infty}$ to f, there exists $N \in \mathbb{N}$ such that for any $i \geq N$ we have:

$$|f_i(x) - f(x)| < \frac{\epsilon}{2(b-a)},$$

for all $x \in [a,b]$. Henceforth assume $i \geq N$. Now, if $I \subset [a,b]$ is any subinterval, we have that:

$$\begin{aligned}
\sup_{x\in I} f(x) &= \sup_{x\in I} \left(f(x) - f_i(x) + f_i(x)\right) \\
&\leq \sup_{x\in I} \left(f(x) - f_i(x)\right) + \sup_{x\in I} f_i(x) \\
&\leq \sup_{x\in I} f_i(x) + \frac{\epsilon}{2(b-a)}.
\end{aligned}$$

From here, we conclude that for any partition $\mathcal{P}$:

$$\begin{aligned}
U(f,\mathcal{P}) &= \sum_{j=0}^{k-1} \sup_{x\in[x_j,x_{j+1}]} f(x)\Delta x_j \\
&\leq \sum_{j=0}^{k-1} \left[\sup_{x\in[x_j,x_{j+1}]} f_i(x) + \frac{\epsilon}{2(b-a)} \right] \Delta x_j \\
&= U(f_i,\mathcal{P}) + \frac{\epsilon}{2(b-a)} \sum_{j=0}^{k-1} (x_{j+1} - x_j) \\
&= U(f_i,\mathcal{P}) + \frac{\epsilon}{2}
\end{aligned}$$

Taking the infimum over $\mathcal{P}$ and using that f_i is integrable, we conclude:

$$\overline{\int_a^b} f(x)dx \leq \int_a^b f_i(x)dx + \frac{\epsilon}{2} < \int_a^b f_i(x)dx + \epsilon.$$

2. Similarly, if $I \subset [a,b]$ is any interval, we have that:

$$\begin{aligned}\inf_{x\in I} f(x) &= \inf_{x\in I} (f(x) - f_i(x) + f_i(x)) \\ &\geq \inf_{x\in I} (f(x) - f_i(x)) + \inf_{x\in I} f_i(x) \\ &\geq \inf_{x\in I} f_i(x) - \frac{\epsilon}{2(b-a)}.\end{aligned}$$

so that:

$$L(f,\mathcal{P}) \geq L(f_i,\mathcal{P}) - \frac{\epsilon}{2}.$$

and taking the supremum, we have:

$$\underline{\int_a^b} f(x)dx \geq \int_a^b f_i(x)dx - \frac{\epsilon}{2} > \int_a^b f_i(x)dx - \epsilon.$$

3. Putting this together, we have:

$$\int_a^b f_i(x)dx - \epsilon < \underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx < \int_a^b f_i(x)dx + \epsilon$$

from which we first conclude:

$$0 \leq \overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx < 2\epsilon.$$

Since ϵ is arbitrary, we must have:

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx = \int_a^b f(x)dx$$

and f is integrable. Moreover, we have:

$$\left|\int_a^b f_i(x)dx - \int_a^b f(x)dx\right| < \epsilon,$$

so indeed:

$$\lim_{i\to\infty} \int_a^b f_i(x)dx = \int_a^b f(x)dx. \qquad \square$$

In Theorem 1.3, we established that if $(f_i)_{i=0}^{\infty}$ is a sequence of uniformly continuous maps $f_i : (a,b) \to \mathbb{R}$ such that $f_i \to f$ uniformly, then $f : (a,b) \to \mathbb{R}$ is uniformly continuous. We are now in a position to extend this to the situation where the f_i's are *differentiable* functions, provided the derivatives also converge uniformly.

Theorem 2.17. *Suppose $(f_i)_{i=0}^{\infty}$ is a sequence of uniformly continuous, differentiable maps $f_i : (a,b) \to \mathbb{R}$ such that each $f_i' : (a,b) \to \mathbb{R}$ is uniformly continuous. Suppose further that:*

$$f_i \to f, \qquad f_i' \to g$$

uniformly as $i \to \infty$. Then f is differentiable on (a,b) and $f' = g$.

Proof. Let us pick $c \in (a,b)$. By Corollary 2.14, we may write:

$$f_i(x) - f_i(c) = \int_c^x f_i'(t)dt.$$

Now, since f_i converges uniformly (hence pointwise), as $i \to \infty$ we have:

$$f_i(x) - f_i(c) \to f(x) - f(c).$$

Further, since the convergence of the f_i' is uniform we know that g is continuous, and by Theorem 2.16 we have:

$$\lim_{i\to\infty} \int_c^x f_i'(x)dx = \int_c^x g(t)dt,$$

so that:

$$f(x) = f(c) + \int_c^x g(t)dt.$$

Applying Theorem 2.13, we have that f is differentiable on (a,b) with $f' = g$. □

It is clear that we can iterate this result to show that if $f_i^{(m)} \to g_m$ for $m = 0, \ldots, k$, then $f := g_0$ will be $k-$times differentiable and $f^{(m)} = g_m$. Combining these results with the Weierstrass $M-$test we have:

Corollary 2.18. *Suppose that $(a_i)_{i=0}^{\infty}$ is a sequence of real numbers $a_i \in \mathbb{R}$, and suppose there exists $r > 0$ such that the sum*

$$\sum_{i=0}^{\infty} |a_i|\, r^i$$

converges. Then the power series:

$$f(x) = \sum_{i=0}^{\infty} a_i x^i$$

converges uniformly to a smooth function on the interval $(-r,r)$, and we can differentiate under the sum.

Proof. We already know from Corollary 1.8 that the sum converges uniformly. We need to show that $f(x)$ is smooth. Fix $k \in \mathbb{N}$ and consider the partial sums

$$S_j(x) = \sum_{i=0}^{j} a_i x^i.$$

for $j \geq k$. These are smooth on $(-r, r)$ with:

$$S_j^{(m)}(x) = \sum_{i=k}^{j} \frac{i!}{(i-m)!} a_i x^{i-m},$$

for $m \leq k$. Let us pick ρ with $0 < \rho < r$. If $x \in [-\rho, \rho]$, then we have:

$$\left| \frac{i!}{(i-m)!} a_i x^{i-m} \right| \leq \frac{i!}{(i-m)!} |a_i| \rho^{i-m} \leq \rho^{-m} \left[\left(\frac{\rho}{r}\right)^i \frac{i!}{(i-m)!} \right] |a_i| r^i = \lambda_i |a_i| r^i,$$

where

$$\lambda_i = \rho^{-m} \left(\frac{\rho}{r}\right)^i \frac{i!}{(i-m)!} \to 0$$

as $i \to \infty$, since $\rho/r < 1$, and $i!/(i-m)!$ is a polynomial in i, so eventually is beaten by the exponential that it's multiplied by. By the comparison test, we have that

$$\sum_{i=0}^{\infty} \lambda_i |a_i| r^i$$

converges, so by the Weierstrass $M-$test, we deduce that $S_j^{(m)}$ converges uniformly on $(-\rho, \rho)$, thus by the previous result, we have that $f(x)$ is $k-$times differentiable on $(-\rho, \rho)$, and moreover:

$$f^{(m)}(x) = \sum_{i=0}^{\infty} \frac{i!}{(i-m)!} a_i x^{i-m}$$

for $m \leq k$ and $x \in [-\rho, \rho]$. Since ρ<r and k were arbitrary, the result follows. □

2.3.3 Improper integrals

It is often the case that the integrals one wishes to consider do not fall into the class allowed by our definition. Either the integrand, or the range of integration may be unbounded. Typically the best approach to handle these problems is to consider limits of problems that we do understand. We will give here a few examples of the situations that can occur, and urge caution in working with them.

Example 2.5. Consider the integral:

$$\int_0^1 \frac{1}{\sqrt{x}} dx.$$

Since $x \mapsto x^{-\frac{1}{2}}$ is unbounded on the interval $[0, 1]$, it is not integrable on this interval. However, for any $\epsilon > 0$, it is integrable on $[\epsilon, 1]$. We are perfectly justified in calculating:

$$\int_\epsilon^1 \frac{1}{\sqrt{x}} dx = 2\left(1 - \sqrt{\epsilon}\right).$$

We therefore will define the *improper integral*:

$$\int_0^1 \frac{1}{\sqrt{x}}dx := \lim_{\epsilon\to 0}\int_\epsilon^1 \frac{1}{\sqrt{x}}dx = 2.$$

Of course, this sort of limit may or may not exist in general. For example:

$$\int_\epsilon^1 \frac{1}{x}dx = -\log\epsilon \to \infty \qquad \text{as } \epsilon \to 0.$$

In this case, we would say that the integral

$$\int_0^1 \frac{1}{x}dx$$

does not exist, or fails to converge.

Example 2.6. We can also consider integrating a function which is continuous and diverges at a single point. For example:

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

In this case we split up the integral into two regions as:

$$\int_{-1}^1 \frac{1}{\sqrt{|x|}}dx = \int_{-1}^0 \frac{1}{\sqrt{|x|}}dx + \int_0^1 \frac{1}{\sqrt{|x|}}dx,$$

and we demand that *both* the improper integrals on the right hand side exist in the sense of Example 2.5. In this case, that is so by the previous part and we have:

$$\int_{-1}^1 \frac{1}{\sqrt{|x|}}dx = 4.$$

One must be careful here. This is a stronger condition than:

$$\lim_{\epsilon\to 0}\left(\int_{-1}^{-\epsilon} \frac{1}{\sqrt{|x|}}dx + \int_\epsilon^1 \frac{1}{\sqrt{|x|}}dx\right)$$

existing. For example, it is true that:

$$\lim_{\epsilon\to 0}\left(\int_{-1}^{-\epsilon} \frac{1}{x}dx + \int_\epsilon^1 \frac{1}{x}dx\right) = \lim_{\epsilon\to 0} 0 = 0, \tag{2.2}$$

however since the limits:

$$\lim_{\epsilon\to 0}\left(\int_{-1}^{-\epsilon} \frac{1}{x}dx\right), \qquad \lim_{\epsilon\to 0}\left(\int_\epsilon^1 \frac{1}{x}dx\right),$$

do not exist, we say that the integral:

$$\int_{-1}^1 \frac{1}{x}dx \tag{2.3}$$

does not converge. The expression (2.2) is sometimes called the Cauchy Principal Value of the integral (2.3). Such quantities often arise naturally in complex analysis.

Example 2.7. One also wishes to consider integrating over an infinite range, that is one wishes to give meaning to expressions such as:

$$\int_0^\infty f(x)dx.$$

This again is not covered by our definition of integration. We define the improper integral to be:

$$\int_a^\infty f(x)dx := \lim_{R\to\infty} \int_a^R f(x)dx,$$

if the limit exists. For example:

$$\int_1^\infty \frac{1}{x^2}dx = \lim_{R\to\infty} \int_a^R \frac{1}{x^2}dx = \lim_{R\to\infty} \left[1 - \frac{1}{R}\right] = 1.$$

For a doubly infinite range, we define:

$$\int_{-\infty}^\infty f(x)dx = \int_{-\infty}^0 f(x)dx + \int_0^\infty f(x)dx,$$

provided both improper integrals on the right hand side exist. Again, note that in general the existence of the limit:

$$\lim_{r\to\infty} \int_{-R}^R f(x)dx$$

is a weaker condition than the condition that both $\int_{-\infty}^0 f(x)dx, \int_0^\infty f(x)dx$ exist.

In all cases of improper integrals one has to be very cautious in applying the results that we have established for the standard integral, as these may or may not be true depending on the circumstances. The best approach is to explicitly reduce to the standard integral, and then attempt to take a limit.

Exercise 6.1. Suppose that $f : [0,\infty) \to [0,\infty)$ is a non-negative, continuous, monotone decreasing function, and let $N, M \in \mathbb{Z}$.

a) Show that:

$$\sum_{n=N}^{M} f(n) \geq \int_N^{M+1} f(x)dx.$$

b) Show that:

$$\sum_{n=N+1}^{M} f(n) \leq \int_N^{M} f(x)dx.$$

c) Establish the inequality:

$$\int_N^{M+1} f(x)dx \leq \sum_{n=N}^{M} f(n) \leq f(N) + \int_N^{M} f(x)dx$$

d) Show that
$$\sum_{n=N}^{\infty} f(n)$$
converges if and only if the improper integral:
$$\int_N^{\infty} f(x)dx$$
exists.

e) Show that the harmonic series:
$$\sum_{n=1}^{\infty} \frac{1}{n}$$
diverges.

Exercise 6.2. Let $n \in \mathbb{N}$ with $n \geq 2$. By choosing a suitable partition to approximate the integral:
$$\int_1^n \log x \, dx$$
by upper and lower sums, show that:
$$\sum_{k=1}^{n-1} \log k \leq n \log n - n + 1 \leq \sum_{k=2}^{n} \log k.$$
Deduce that:
$$n^n e^{-n+1} \leq n! \leq (n+1)^{n+1} e^{-n}.$$

2.4 Examples and applications

2.4.1 Complex and vector valued integrals

In some circumstances we wish to integrate a function which maps some interval $[a, b]$ to $\mathbb{C}$ or alternatively to $\mathbb{R}^m$. We can incorporate this into our existing set-up quite simply. Suppose $f : [a, b] \to \mathbb{C}$. Then we can write:
$$f(t) = u(x) + iv(x)$$
We say that f is integrable if both of the functions $u, v : [a, b] \to \mathbb{R}$ are integrable, and we write:
$$\int_a^b f(t)dt = \int_a^b u(t)dt + i \int_a^b v(t)dt.$$

In a similar way, if $f : [a, b] \to \mathbb{R}^m$, then we can write:
$$f(t) = e_1 f^1(t) + e_2 f^2(t) + \ldots + e_m f^m(t)$$
where $\{e_i\}_{i=1}^m$ is the standard basis for $\mathbb{R}^m$ and $f_i : [a, b] \to \mathbb{R}$ are real valued functions for $i = 1, \ldots, m$. We say that f is integrable if each f_i is integrable, and we define:
$$\int_a^b f(t)dt = e_1 \int_a^b f^1(t)dt + e_2 \int_a^b f^2(t)dt + \ldots + e_m \int_a^b f^m(t)dt$$

Exercise 6.3. Suppose $f : [a,b] \to \mathbb{R}^m$ is integrable.

a) Show that if $x, y \in [a,b]$:

$$\big|\, ||f(x)|| - ||f(y)||\, \big| \leq \sum_{l=1}^{m} \left| f^l(x) - f^l(y) \right|,$$

where $f(t) = e_1 f^1(t) + e_2 f^2(t) + \cdots + e_m f^m(t)$.

b) Deduce that if I is any subinterval of $[a,b]$:

$$\sup_{x\in I} ||f(x)|| - \inf_{y\in I} ||f(y)|| \leq \sum_{l=1}^{m} \left(\sup_{x\in I} f^l(x) - \inf_{y\in I} f^l(y) \right).$$

c) Prove that $||f|| : [a,b] \to \mathbb{R}$ is integrable.

Exercise 6.4. Let $f : [a,b] \to \mathbb{R}^m$ be an integrable function, and write

$$f(t) = e_1 f^1(t) + e_2 f^2(t) + \cdots + e_m f^m(t)$$

for $f^i : [a,b] \to \mathbb{R}$.

a) For a tagged partition $(\mathcal{P}, \tau)$, show that:

$$R(f, \mathcal{P}, \tau) = e_1 R(f^1, \mathcal{P}, \tau) + e_2 R(f^2, \mathcal{P}, \tau) + \ldots + e_m R(f^m, \mathcal{P}, \tau)$$

b) Show that if $((\mathcal{P}_l, \tau_l))_{l=0}^{\infty}$ is a sequence of tagged partitions of $[a,b]$ with mesh $\mathcal{P}_l \to 0$, then

$$R(f, \mathcal{P}_l, \tau_l) \to \int_a^b f(x)dx,$$

as $l \to \infty$.

c) Show that for a tagged partition $(\mathcal{P}, \tau)$:

$$||R(f, \mathcal{P}, \tau)|| \leq R(||f||, \mathcal{P}, \tau).$$

d) Deduce that:

$$\left|\left| \int_a^b f(t)dt \right|\right| \leq \int_a^b ||f(t)||\, dt.$$

Note that an identical argument shows that if $f : [a,b] \to \mathbb{C}$ then:

$$\left| \int_a^b f(t)dt \right| \leq \int_a^b |f(t)|\, dt.$$

Exercise 6.5 (*).

Suppose $f, g, h : [a,b] \to \mathbb{R}^m$ are integrable. We define:

$$(f,g) := \int_a^b \langle f(t), g(t) \rangle \, dt, \qquad ||f||_{L^2} = \sqrt{(f,f)}.$$

a) Show that:

$$(f,g) = (g,f), \qquad (f+h,g) = (f,g) + (h,g), \qquad (\lambda f, g) = \lambda (f,g),$$

where $\lambda \in \mathbb{R}$.

b) For $t \in \mathbb{R}$, show that:

$$||f + tg||_{L^2}^2 = ||f||_{L^2}^2 + 2t(f,g) + t^2 \, ||g||_{L^2}^2 \geq 0. \tag{2.4}$$

c) By thinking of (2.4) as a quadratic in t and considering its possible roots, deduce the *Cauchy-Schwartz* inequality:

$$|(f,g)| \leq ||f||_{L^2} \, ||g||_{L^2} \, .$$

When does equality hold?

d) Deduce that:

$$||f+g||_{L^2} \leq ||f||_{L^2} + ||g||_{L^2} \, .$$

2.4.2 (*)The spaces $C^k([a,b])$

Recall that in Section 1.2.3, we defined the space $C^0([a,b])$ to consist of all continuous functions $f : [a,b] \to \mathbb{R}$, equipped with the norm function:

$$||f||_{C^0} = \max_{x \in [a,b]} |f(x)| \, .$$

We showed that this satisfied certain conditions analogous to those of the norm in Euclidean space. We also defined Cauchy sequences with respect to this norm and showed that they converge uniformly to a continuous function. We're now in a position to upgrade this definition to consider functions which have a certain amount of differentiability.

Definition 2.3. $C^k([a,b])$ is the set of all $k-$times differentiable functions $f : [a,b] \to \mathbb{R}$ such that f is continuous on $[a,b]$ and $f^{(m)} : (a,b) \to \mathbb{R}$ is uniformly continuous for $m = 1, \ldots, k$. We define a function $||\cdot||_{C^k}$ by:

$$||f||_{C^k} = \sum_{m=0}^{k} \sup_{x \in (a,b)} \left| f^{(m)}(x) \right|,$$

for all $f \in C^k([a,b])$. We call $||\cdot||_{C^k}$ the *norm* on $C^k([a,b])$.

It is not immediately obvious, but for any $f \in C^k([a,b])$, we have that $||f||_{C^k} < \infty$. This follows from the following Lemma:

Lemma 2.19. *Suppose $f : (a, b) \to \mathbb{R}$ is uniformly continuous. Then f is bounded.*

Proof. Since f is uniformly continuous, there exists $\delta > 0$ such that if $x, y \in (a, b)$ with $|x - y| < \delta$, then $|f(x) - f(y)| < 1$. Let

$$y_j = a + \frac{j}{N}(b - a), \qquad j = 1, 2, \ldots, N$$

where N is sufficiently large that $\frac{(b-a)}{N} < \delta$. Set $M = \max\{f(y_1), f(y_2), \ldots, f(y_N)\}$. For any $x \in (a, b)$ there exists some j such that $|x - y_j| < \delta$, so we conclude:

$$|f(x)| = |f(x) - f(y_j) + f(y_j)| \le |f(x) - f(y_j)| + |f(y_j)| < 1 + M,$$

so f is bounded on (a, b). □

Lemma 2.20. *The the norm on $C^k([a, b])$ has the following properties:*

i) For all $f \in C^k([a, b])$, we have $||f||_{C^k} \ge 0$, with $||f||_{C^0} = 0$ if and only if $f \equiv 0$.

ii) For all $f \in C^k([a, b])$, if $a \in \mathbb{R}$, then $||af||_{C^0} = |a|\, ||f||_{C^k}$

iii) The norm satisfies the triangle inequality:

$$||f + g||_{C^k} \le ||f||_{C^0} + ||g||_{C^k}\,, \qquad \forall\ f, g, \in C^k. \tag{2.5}$$

Proof. i) It is clear that since each term in the sum defining $||f||_{C^k}$ is the supremum of a set of non-negative numbers, it must be non-negative, thus $||f||_{C^k}$ is non-negative. If $||f||_{C^k} = 0$, then in particular:

$$\sup_{x \in [a,b]} |f(x)| = 0,$$

so that $0 \le |f(x)| \le 0$ and $f(x) = 0$ for all $x \in [a, b]$.

ii) If $a \in \mathbb{R}$, then:

$$||af||_{C^k} = \sum_{m=0}^{k} \sup_{x \in (a,b)} \left| af^{(m)}(x) \right| = \sum_{m=0}^{k} \sup_{x \in (a,b)} |a| \left| f^{(m)}(x) \right| = |a|\, ||f||_{C^k}\,.$$

iii) We estimate:

$$\begin{aligned} ||f + g||_{C^k} &= \sum_{m=0}^{k} \sup_{x \in (a,b)} \left| f^{(m)}(x) + g^{(m)}(x) \right| \le \sum_{m=0}^{k} \sup_{x \in (a,b)} \left(\left| f^{(m)}(x) \right| + \left| g^{(m)}(x) \right| \right) \\ &\le \sum_{m=0}^{k} \left(\sup_{x \in (a,b)} \left| f^{(m)}(x) \right| + \sup_{x \in (a,b)} \left| g^{(m)}(x) \right| \right) \\ &= \sum_{m=0}^{k} \sup_{x \in (a,b)} \left| f^{(m)}(x) \right| + \sum_{m=0}^{k} \sup_{x \in (a,b)} \left| g^{(m)}(x) \right| = ||f||_{C^k} + ||g||_{C^k}\,. \end{aligned}$$

Here we have used the triangle inequality for $|\cdot|$, as well as the fact that $\sup(a + b) \le \sup a + \sup b$. □

The observant reader will note that the proof here is almost identical to that of the properties of the norm on $C^0([a,b])$.

Now, we can give a definition of convergence in $C^k([a,b])$ as follows.

Definition 2.4. A sequence $(f_j)_{j=0}^{\infty}$ of functions $f_j \in C^k([a,b])$ converges to $f \in C^k([a,b])$ if the following holds: for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $j \geq N$, then

$$||f_j - f||_{C^k} < \epsilon.$$

A sequence $(f_j)_{j=0}^{\infty}$ of functions $f_j \in C^k([a,b])$ is *Cauchy* if the following holds: for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $i, j \geq N$, then

$$||f_i - f_j||_{C^k} < \epsilon.$$

Notice that since $||\cdot||_{C^k}$ is a sum of positive terms, we have in particular that:

$$\left|\left|f^{(m)}\right|\right|_{C^0} < ||f||_{C^k}$$

for all $m \leq k$, so that if $f_j \to f$ in C^k, then in particular $f_j^{(m)} \to f^{(m)}$ uniformly for each m. Similarly, if $(f_j)_{j=0}^{\infty}$ is a Cauchy sequence in $C^k([a,b])$, then $\left(f_j^{(m)}\right)_{j=0}^{\infty}$ is a Cauchy sequence[3] in $C^0([a,b])$ for each $m \leq k$.

Theorem 2.21. *A sequence $(f_j)_{j=0}^{\infty}$ of functions $f_j \in C^k([a,b])$ converges in $C^k([a,b])$ if and only if it is a Cauchy sequence.*

Proof. From the observations above, we see that $(f_j)_{j=0}^{\infty}$ converges in $C^k([a,b])$ if and only if $\left(f_j^{(m)}\right)_{j=0}^{\infty}$ converges in $C^0([a,b])$ for all $m \leq k$. By Theorem 1.6 we know that $\left(f_j^{(m)}\right)_{j=0}^{\infty}$ converges in $C^0([a,b])$ if and only if it is a Cauchy sequence in $C^0([a,b])$. Finally, we know that $\left(f_j^{(m)}\right)_{j=0}^{\infty}$ is a Cauchy sequence. $C^0([a,b])$ for all $m \leq k$ if and only if $(f_j)_{j=0}^{\infty}$ is a Cauchy sequence in $C^k([a,b])$. □

Thus the space $C^k([a,b])$ has the property of *completeness* that we introduced in Section 1.2.3 and so is another example of a Banach space.

[3] Strictly speaking, we should use the result of Exercise 3.6 to argue that a uniformly continuous function on (a,b) extends to a continuous function on $[a,b]$.