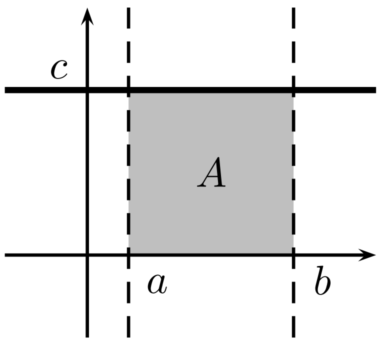
13 The definite integral

13.1 Areas under graphs of functions

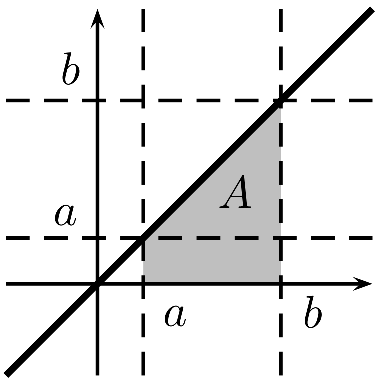
In this section we will compute **areas** under graphs of functions. More specifically, we will look at the area between the horizontal axis, the vertical lines through the points $(a, 0)$ and $(b, 0)$, and the graph of a function f between a and b where $f(x) \geq 0$ for all $x \in [a, b]$.

As a first example, take $f(x) = c$ for some constant $c \in \mathbb{R}$.



Here, the area A is simply $A = c \cdot (b - a)$.

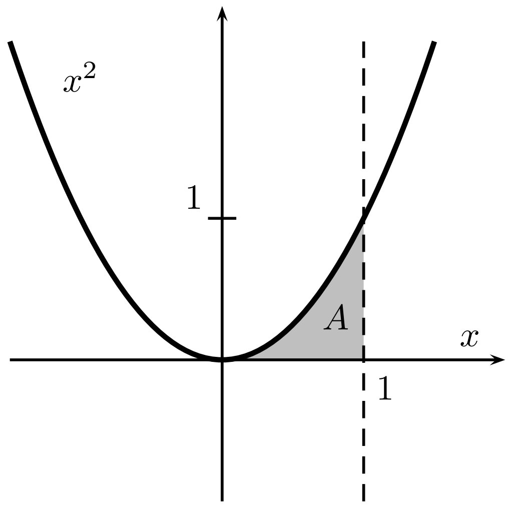
Things get slightly more exciting when we consider non-constant functions. For $f(x) = x$ we are faced with the following situation:



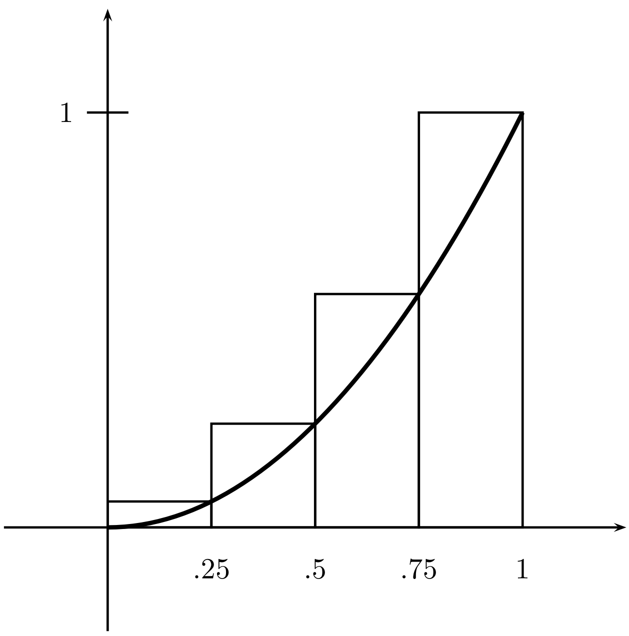
In this case, the total area is made up of a rectangle and a triangle. We find that

$$A = (b - a) \cdot a + \frac{1}{2}(b - a) \cdot (b - a) = \frac{1}{2}(b - a)(b + a) = \frac{1}{2}(b^2 - a^2).$$

The situation is considerably more challenging for functions that have 'curved' graphs. For $f(x) = x^2$, we do not have any obvious way to compute the area A.



However, we can find an approximation. To estimate the area under the curve of $f(x) = x^2$ between $x = 0$ and $x = 1$, we divide the interval $[0, 1]$ into four strips as shown in the following figure.



Each of the four strips has width $1/4$. The total area of the four strips is

$$\frac{1}{4}\left(\left(\frac{1}{4}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{3}{4}\right)^2 + 1^2\right) = \frac{15}{32} = 0.46875.$$

From the picture, it is clear that the area of the four strips is bigger than the actual area under the curve. We can get a better approximation by increasing the number of strips. For ten strips, we find

$$\frac{1}{10}\left(0.1^2 + 0.2^2 + 0.3^2 + 0.4^2 + 0.5^2 + 0.6^2 + 0.7^2 + 0.8^2 + 0.9^2 + 1^2\right) = \frac{77}{200} = 0.385.$$

While this is certainly still bigger than the actual area A, it is better than our first approximation. We now have $A < 0.385 < 0.46875$.

It is obvious how we can improve the estimate for A: we just use a larger number of strips. How many are enough? Computations get more and more cumbersome as the number of strips increases. With a little help from a computer we might get quite far, but what we have already learned about limits can get us even further.

We first derive a formula for the area of n strips and then compute the limit as n tends to ∞. If the interval $[0,1]$ is divided into n equal strips, each strip has width $1/n$. The height of the first strip is $(1/n)^2$, the height of the second strip is $(2/n)^2$, and so on. The total area covered by the strips is

$$\frac{1}{n}\left(\frac{1^2}{n^2}+\frac{2^2}{n^2}+\frac{3^2}{n^2}+\cdots+\frac{n^2}{n^2}\right)=\frac{1}{n^3}\sum_{k=1}^{n}k^2.$$

As we have established previously, $\displaystyle\sum_{k=1}^{n}k^2=\frac{n(n+1)(2n+1)}{6}$, and so the approximate area is

$$\frac{(n+1)(2n+1)}{6n^2}.$$

In the limit as $n\to\infty$ we get

$$\begin{aligned} A &= \lim_{n\to\infty}\frac{(n+1)(2n+1)}{6n^2}=\lim_{n\to\infty}\left(\frac{1}{3}\cdot\frac{n+1}{n}\cdot\frac{2n+1}{2n}\right)\\ &= \frac{1}{3}\lim_{n\to\infty}\left(1+\frac{1}{n}\right)\lim_{n\to\infty}\left(1+\frac{1}{2n}\right)=\frac{1}{3}. \end{aligned} \tag{13.1}$$

This is the process that we will now generalise.

13.2 Lower and upper sums: the integral

We will first define a way of describing how we break an interval up into ‘strips’. We will think about how to estimate the value of the area. In fact we will estimate it in two ways, one of which always gives an underestimate (or the exact value) while the other always gives an overestimate (or the exact value). The exact value must then be ‘sandwiched’ between these estimates.

Definition 13.1 *A **partition** of the interval $[a,b]$ is a finite collection of points $x_i\in[a,b]$ such that $a=x_0<x_1<x_2<\cdots<x_{n-1}<x_n=b$.*

Example 13.1 Which of the following sets are partitions of $[0,1]$?

$$\{0,1\},\quad \{0,\tfrac{1}{2},1\},\quad \{0,\tfrac{1}{3},1\},\quad \{0,\tfrac{1}{2},\tfrac{1}{3},1\},\quad \{0,\tfrac{1}{4},\tfrac{1}{2},\tfrac{3}{4}\},\quad \{0,\tfrac{1}{3},\tfrac{1}{3},\tfrac{2}{3},1\}.$$

The sets

$$\{0,1\}, \quad \{0,\frac{1}{2},1\}, \quad \{0,\frac{1}{3},1\}, \quad \{0,\frac{1}{2},\frac{1}{3},1\},$$

are partitions of $[0,1]$. Note that sets are not ordered, though the elements of each set can be ordered.

However, the set $\{0,\frac{1}{4},\frac{1}{2},\frac{3}{4}\}$ is not a partition of $[0,1]$ because it does not include the end point 1, and the set $\{0,\frac{1}{3},\frac{1}{3},\frac{2}{3},1\}$ is not a partition, because one of the points is repeated.

A partition $P = \{x_0, x_1, \cdots, x_{n-1}, x_n\}$ of $[a,b]$ breaks up the interval $[a,b]$ into n closed subintervals

$$[x_0,x_1],\ [x_1,x_2],\ \cdots, [x_{n-1},x_n]$$

of lengths $\Delta x_1,\ \Delta x_2,\ \ldots, \Delta x_n$ where $\Delta x_i = x_i - x_{i-1}$.

A function f that is continuous on $[a,b]$ is also continuous on each subinterval $[x_{i-1},x_i]$. It then has a minimum value m_i and a maximum value M_i on the interval $[x_{i-1},x_i]$.[4]

Definition 13.2 *Consider a function f that is continuous on $[a,b]$ and a partition $P = \{x_0, x_1, \cdots, x_{n-1}, x_n\}$ of $[a,b]$.*

The number

$$L(f,P) = \sum_{i=1}^{n} m_i \Delta x_i$$

is called the ***lower sum*** *of f for P.*

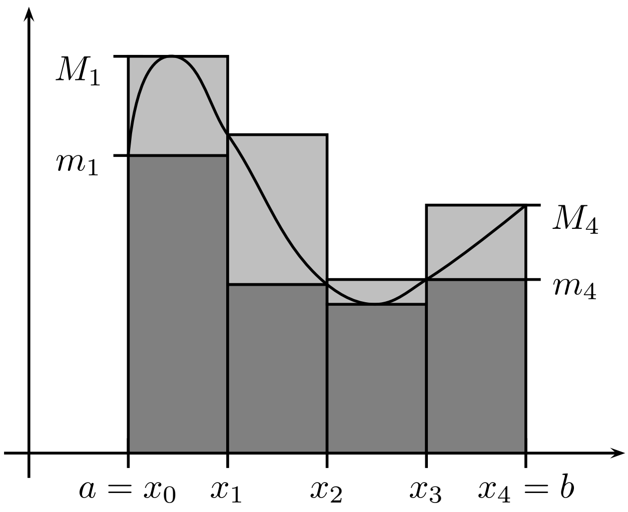
The number

$$U(f,P) = \sum_{i=1}^{n} M_i \Delta x_i$$

is called the ***upper sum*** *of f for P.*

Here, $\Delta x_i = x_i - x_{i-1}$.

[4]Although this seems intuitively obvious, it is not easy to prove rigorously. It corresponds to what's sometimes called the Extreme Value Theorem; it appears as Theorem 3 in Chapter 7 of SPIVAK's book, and is proved in Chapter 8.



Definition 13.3 ***The Definite Integral.*** *A function f is* ***integrable*** *on an interval $[a, b]$ if there is exactly one number I such that*

$$L(f, P) \leq I \leq U(f, P)$$

for all partitions P of $[a, b]$. This number I is then called the ***definite integral*** *of f on $[a, b]$ and is denoted by*

$$I = \int_a^b f.$$

By its definition, the integral I is a number that depends on the function f and on the numbers a and b. The numbers a and b are called the **lower and upper limit of integration**, respectively. The function f is called the **integrand**. The symbol $\displaystyle\int$, the **integral sign**, was introduced by LEIBNIZ. It is an elongated S, reminiscent of $\sum$, the Greek S in Definition 13.2.

An alternative notation for the integral is in wide use. One often writes

$$\int_a^b f = \int_a^b f(x)\,\mathrm{d}x.$$

The expressions on the left- and right-hand sides mean exactly the same. This alternative notation is particularly useful when we want to write integrals of simple functions. For example, we can write

$$\int_a^b x^2\,\mathrm{d}x$$

without the need to introduce a name for the function $x \mapsto x^2$ first. However, it should be noted that in the expression

$$I = \int_a^b f(x)\,\mathrm{d}x$$

the variable x is merely a *dummy*. The value I of the integral does *not* depend on x. We could have used any letter (other than a, b, and f, which are already in use!):

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b f(x)\,\mathrm{d}x = \int_a^b f(u)\,\mathrm{d}u = \int_a^b f(\xi)\,\mathrm{d}\xi = \cdots$$

In this example, t, x, u, and ξ are all dummy variables. This is similar to $\sum_{r=1}^{5} \frac{r}{r^2+1} = \frac{1}{2} + \frac{2}{5} + \frac{3}{10} + \frac{4}{17} + \frac{5}{26}$ in which r is a dummy, or $\lim\limits_{x \to 0} \frac{\sin x}{x} = 1$ where x could be replaced by any other letter.

Lemma 13.1 *Given a function f and an interval $[a,b]$, suppose that we can find a real number I such that $L(f,P) \le I \le U(f,P)$ for all partitions P of $[a,b]$, as in Definition 13.3. If, given any $\epsilon > 0$, we can find a partition P_ϵ such that $|U(f,P_\epsilon) - L(f,P_\epsilon)| < \epsilon$, then I is unique; thus the conditions of Definition 13.3 are satisfied and $\displaystyle\int_a^b f = I$.*

Proof: We proceed by contradiction. Suppose that I is not unique, so there exist real numbers I_1 and I_2 such that $I_1 < I_2$ and that $L(f,P) \le I_1 \le U(f,P)$ and $L(f,P) \le I_2 \le U(f,P)$ for all partitions P of $[a,b]$. We thus have

$$L(f,P) \le I_1 < I_2 \le U(f,P) \implies U(f,P) - L(f,P) \ge I_2 - L(f,P) \ge I_2 - I_1.$$

Thus, if we take $\epsilon < I_2 - I_1$ then there is no partition P such that $U(f,P) - L(f,P)| < \epsilon$. The proof of the Lemma follows by contradiction. □

Intuitively, this means that if we can show that by a suitable choice of partitions we can make the upper and lower sums U and L arbitrarily close, then there is a unique value "sandwiched" between them. This will be useful below, because it provides a convenient way of proving the uniqueness of a proposed value of the definite integral.

Example 13.2 If f is a constant function, $f(x) = c$ for all $x \in [a,b]$, then

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

To see this, take an arbitrary partition $P = \{x_0, x_1, \cdots, x_{n-1}, x_n\}$ of $[a,b]$. Since f is constant on $[a,b]$, it is also constant on each subinterval $[x_{i-1}, x_i]$. Thus all m_i and all M_i are equal to c. It follows that

$$\begin{aligned} L(f,P) &= \sum_{i=1}^{n} m_i \Delta x_i = \sum_{i=1}^{n} c \Delta x_i = c \sum_{i=1}^{n} \Delta x_i \\ &= c\left((x_1 - x_0) + (x_2 - x_1) + \cdots + (x_{n-1} - x_{n-2}) + (x_n - x_{n-1})\right) \\ &= c(x_n - x_0) = c(b-a) \end{aligned}$$

because the only terms in the sum that do not cancel are $x_n = b$ and $-x_0 = -a$. Similarly,

$$U(f,P) = \sum_{i=1}^{n} M_i \Delta x_i = c(b-a).$$

Clearly, the only candidate for the integral is $c(b-a)$ as

$$c(b-a) = L(f,P) \le c(b-a) \le U(f,P) = c(b-a).$$

Because this holds for all partitions P, we have shown that

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

Example 13.3 By a slightly more complicated argument we can determine the value of $\displaystyle\int_a^b f$ where $f(x) = cx$ for all $x \in [a,b]$ and where c is a positive constant.

We consider an arbitrary partition $P = \{x_0, x_1, \cdots, x_{n-1}, x_n\}$ of $[a,b]$. Since the function f is increasing on each subinterval $[x_{i-1}, x_i]$, we have $m_i = cx_{i-1}$ and $M_i = cx_i$ for each i. Thus,

$$\begin{aligned} L(f,P) = \sum_{i=1}^{n} m_i \Delta x_i &= \sum_{i=1}^{n} cx_{i-1}(x_i - x_{i-1}) \\ &= c\left[x_0x_1 + x_1x_2 + \cdots + x_{n-1}x_n - \left(x_0^2 + x_1^2 + \cdots + x_{n-1}^2\right)\right] \end{aligned}$$

and

$$\begin{aligned} U(f,P) = \sum_{i=1}^{n} M_i \Delta x_i &= \sum_{i=1}^{n} cx_i(x_i - x_{i-1}) \\ &= c\left[x_1^2 + x_1^2 + \cdots + x_n^2 - (x_0x_1 + x_1x_2 + \cdots + x_{n-1}x_n)\right]. \end{aligned}$$

To estimate the integral, we will define I to lie between the upper and lower sums, and show that this satisfies Definition 13.3. It is convenient to choose

$$\begin{aligned} I = \frac{1}{2}(L+U) &= \frac{1}{2}c\Big[(x_0x_1 - x_0^2) + (x_1x_2 - x_1^2) + \cdots + (x_{n-1}x_n - x_{n-1}^2) \\ &\qquad + (x_1^2 - x_1x_0) + (x_2^2 - x_2x_1) + \cdots + (x_n^2 - x_nx_{n-1})\Big] \\ &= \frac{1}{2}c(x_n^2 - x_0^2) = \frac{1}{2}c(b^2 - a^2). \end{aligned}$$

Since $L \le I \le U$, part of Definition 13.3 is satisfied, and it remains to show that I is unique. To do so, we use Lemma 13.1 and show that we can find a partition P such that we can make $|U - L|$ as small as we like. We let P_N be the partition with N equal sub-intervals, so that $\Delta x_i = \Delta x = (b-a)/N$ for all i. Then we have

$$L(f, P_N) = c\Delta x \sum_{i=1}^{N} x_{i-1} \qquad \text{and} \qquad U(f, P_N) = c\Delta x \sum_{i=1}^{N} x_i$$

and so

$$U(f, P_N) - L(f, P_N) = c\Delta x(x_n - x_0) = c(b-a)\frac{(b-a)}{N} = \frac{c(b-a)^2}{N}.$$

Thus, given any $\epsilon > 0$, if we take some N such that $N > c(b-a)^2/\epsilon$ then we must have $|U(f, P_N) - L(f, P_N)| < \epsilon$, i.e. by choosing N large enough we can make $|U - L|$ as small as we like. It follows that the definite integral is defined and equal to I.

The examples we have seen so far may suggest that every function is integrable. However, there are exceptions!

Example 13.4 The function

$$f(x) = \begin{cases} 0 & \text{if} \quad x \text{ is irrational} \\ 1 & \text{if} \quad x \text{ is rational} \end{cases}$$

is *not* integrable.

Every interval $[x_{i-1}, x_i]$ contains both irrational and rational numbers (you can think about how to prove this!), and so $m_i = 0$ and $M_i = 1$. For any partition P of an interval $[a, b]$, the lower sum $L(f, P) = 0$ and the upper sum $U(f, P) = b - a$, so any number $N \in [0, b-a]$ satisfies

$$L(f, P) \leq N \leq U(f, P).$$

In particular, this means that there is no *unique* number I with this property: the integral does not exist.

13.3 Some useful theorems on the definite integral

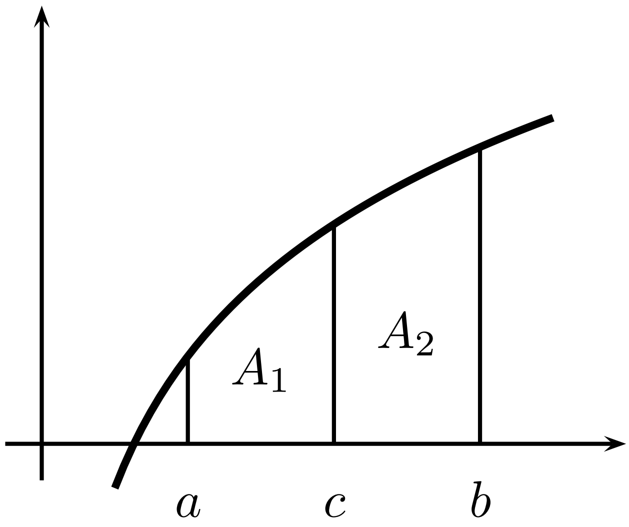
If $f(x) \geq 0$ for all $x \in [a, b]$, then by construction of the integral, the area A under the curve of f between a and b is

$$A = \int_a^b f.$$

This suggests the following theorem.

Theorem 13.1 *If f is integrable on $[a,b]$, then for any $c \in (a,b)$ (that is, $a < c < b$) f is integrable on $[a,c]$ and $[c,b]$. Conversely, if f is integrable on $[a,c]$ and $[c,b]$, then it is also integrable on $[a,b]$. We then have*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$



Proof: For nonnegative functions f this is clear from the interpretation of the integral as an area. A proper proof for a general f can be given by resorting to appropriate partitions and upper and lower sums. This is omitted here; see the textbooks for details. $\square$

So far we have always assumed that $a < b$. To be able to compute definite integrals with arbitrary lower and upper limits of integration, we make the following definitions.

Definition 13.4

$$\int_a^b f = -\int_b^a f \quad \text{if } a > b \qquad \text{and} \qquad \int_a^a f = 0.$$

With these definitions the identity $\int_a^b f = \int_a^c f + \int_c^b f$ of Theorem 13.1 holds for all numbers a, b, and c even if $a < c < b$ is not true. This can be proved by checking individually all the different possible orderings of a, b, and c. (Try $a = b$.)

Theorem 13.2 *If f is integrable on $[a,b]$, then cf is integrable on $[a,b]$ for any constant c and*

$$\int_a^b cf = c \cdot \int_a^b f.$$

Proof: If $c = 0$ the result is obvious: both sides of the equation are 0. If $c > 0$, we have for any partition P that

$$L(cf, P) = cL(f, P) \qquad \text{and} \qquad U(cf, P) = cU(f, P)$$

because if m_i is the minimum of f on $[x_{i-1}, x_i]$, then cm_i is the minimum of cf on $[x_{i-1}, x_i]$, and if M_i is the maximum of f on $[x_{i-1}, x_i]$, then cM_i is the maximum of cf on $[x_{i-1}, x_i]$. Since f is integrable, $\int_a^b f = I$ is the unique number such that for all partitions P

$$\begin{array}{rrcl} & L(f,P) & \leq \quad I \quad \leq & U(f,P) \\ \iff & cL(f,P) & \leq \quad cI \quad \leq & cU(f,P) \\ \iff & L(cf,P) & \leq \quad cI \quad \leq & U(cf,P). \end{array}$$

The last inequality means that the (unique) number cI is the integral $\int_a^b cf$, which proves the theorem for $c \geq 0$.

For $c < 0$, the minimum of cf on $[x_{i-1}, x_i]$ is cM_i, and its maximum is cm_i (the greatest value of f, M_i, once multiplied by a negative c, becomes the least value of cf). This implies that

$$L(cf, P) = cU(f, P) \qquad \text{and} \qquad U(cf, P) = cL(f, P).$$

Then

$$\begin{array}{rrcl} & L(f,P) & \leq \quad I \quad \leq & U(f,P) \\ \iff & cL(f,P) & \geq \quad cI \quad \geq & cU(f,P) \\ \iff & U(cf,P) & \geq \quad cI \quad \geq & L(cf,P) \\ \iff & L(cf,P) & \leq \quad cI \quad \leq & U(cf,P), \end{array}$$

which proves the theorem for $c < 0$. □

Theorem 13.3 *If f and g are integrable on $[a, b]$, then $f + g$ is integrable on $[a, b]$ and*

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g.$$

Proof: This proof is a little bit trickier than the one for Theorem 13.2. It relies on the fact that on any of the subintervals of a partition

$$m_i^{f+g} \geq m_i^f + m_i^g \qquad \text{and} \qquad M_i^{f+g} \leq M_i^f + M_i^g.$$

This implies

$$L(f, P) + L(g, P) \leq L(f + g, P) \qquad \text{and} \qquad U(f + g, P) \leq U(f, P) + U(g, P),$$

and so

$$L(f,P)+L(g,P)\le L(f+g,P)\le \int_a^b (f+g)\le U(f+g,P)\le U(f,P)+U(g,P).$$

Because f and g are integrable, for any $\epsilon>0$ there exist partitions P^f and P^g such that $U(f,P^f)-L(f,P^f)<\epsilon/2$ and $U(g,P^g)-L(g,P^g)<\epsilon/2$. Thus if we choose a partition P that contains both P^f and P^g, then $U(f,P)+U(g,P)-(L(f,P)+L(g,P))<\epsilon$. This shows that the integral exists. But because f and g are integrable, we also have

$$L(f,P)+L(g,P)\le \int_a^b f+\int_a^b g\le U(f,P)+U(g,P).$$

Comparing the last two displayed equations, and using Lemma 13.1 to establish uniqueness, shows that

$$\int_a^b (f+g)=\int_a^b f+\int_a^b g$$

as required. □

14 The Fundamental Theorem of Calculus

We are now ready to make the connection between integration and differentiation, and to put it to work.

14.1 The two Fundamental Theorems

We start by looking at the function F defined by $F(x) = \int_a^x f$.

Theorem 14.1 ***The First Fundamental Theorem of Calculus.***
If f is continuous on $[a, b]$, then F defined on $[a, b]$ as

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

is continuous and differentiable at all $c \in (a, b)$ and

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

Proof: We need to look at

$$\lim_{h \to 0} \frac{F(c+h) - F(c)}{h}.$$

From Theorem 13.1,

$$\int_a^c f + \int_c^{c+h} f = \int_a^{c+h} f,$$

and so we find that

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

Denote by m_h and M_h, respectively, the minimum and maximum value of f on $[c, c+h]$. (For simplicity, we have here assumed that $h > 0$. For $h < 0$, we would have to look at the interval $[c+h, c]$.) Denote by P_h the trivial partition of $[c, c+h]$, $P_h = \{c, c+h\}$. We then have

$$L(f, P_h) = m_h \cdot h \leq \int_c^{c+h} f \leq M_h \cdot h = U(f, P_h).$$

Because $h \neq 0$, we can divide by it and obtain

$$m_h \leq \frac{F(c+h) - F(c)}{h} \leq M_h.$$

Now, since f is continuous on $[c, c+h]$, we have

$$\lim_{h \to 0} m_h = f(c) = \lim_{h \to 0} M_h$$

(you can prove this yourself starting from the definition of continuity) and hence

$$F'(c) = \lim_{h \to 0} \frac{F(c+h) - F(c)}{h} = f(c).$$

This shows that F is differentiable on $[a, b]$, and it then follows by Theorem 11.2 that it is also continuous on $[a, b]$. $\square$

Remark: We have proved the Fundamental Theorem for continuous functions f. The theorem is actually more general: it holds for all *integrable* functions f, and $F'(c) = f(c)$ at all points c where f is continuous. The proof of the more general form of the theorem requires techniques that we have not yet developed.

Example 14.1 Find $\dfrac{\mathrm{d}}{\mathrm{d}x} \displaystyle\int_a^x \cos^{17}(\sqrt{t})\,\mathrm{d}t$.

$$\frac{\mathrm{d}}{\mathrm{d}x} \int_a^x \cos^{17}(\sqrt{t})\,\mathrm{d}t = \cos^{17}(\sqrt{x}).$$

Example 14.2 Find $\dfrac{\mathrm{d}}{\mathrm{d}z} \displaystyle\int_z^7 3^{t^2}\,\mathrm{d}t$.

$$\frac{\mathrm{d}}{\mathrm{d}z} \int_z^7 3^{t^2}\,\mathrm{d}t = \frac{\mathrm{d}}{\mathrm{d}z}\left(-\int_7^z 3^{t^2}\,\mathrm{d}t\right) = -3^{z^2}.$$

Theorem 14.2 ***The Second Fundamental Theorem of Calculus.***
If f is continuous on $[a, b]$ and $f = g'$ for some function g, then

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

Proof: Let

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

then $F' = f = g'$ on $[a, b]$. But if F and g have the same derivative (the same slope!) everywhere, they can only differ by a constant C so that[5]

$$F = g + C.$$

[5]Two comments. (i) We haven't proved that if $F' = g'$ then $F - g$ must be a constant. The proof of this result lies beyond the syllabus, and involves the Mean Value Theorem. See, e.g., section 3.2 of STEWART, section 2.6 of ADAMS, or chapter 11 of SPIVAK. (ii) We are a bit sloppy about notation here: we use the symbol C to denote both the constant $C \in \mathbb{R}$ and the constant function $C(x) = C$.

We can find C by observing that $0 = F(a) = g(a) + C$, which yields $C = -g(a)$. Thus

$$F(x) = g(x) + C = g(x) - g(a).$$

This must also be true for $x = b$, which gives

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

$\square$

Comment. We will often find it useful to adopt the notation[6]

$$[g(x)]_a^b = g(b) - g(a).$$

These theorems enable us to compute many integrals without the need to compute any limits of sums. Whenever we know a function g whose derivative g' equals the integrand f, we can evaluate the integral straight away.

Example 14.3 Find $\displaystyle\int_0^1 f$ for $f(x) = x^2$.

We know that

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^3}{3} = x^2.$$

That is, with $g(x) = \dfrac{x^3}{3}$ we have $g'(x) = f(x) = x^2$. Hence

$$\int_0^1 f = \int_0^1 x^2 \mathrm{d}x = g(1) - g(0) = \frac{1^3}{3} - 0 = \frac{1}{3}.$$

This is of course the same result as (13.1), but obtained with considerably less effort.

Example 14.4 Find the area under the function $f(x) = \sin x$ between $x = 0$ and $x = \pi$.

---

[6] Both Adams and Spivak use a slightly different notation that means exactly the same:

$$g(x)\Big|_a^b = g(b) - g(a).$$

Our notation might use more ink, but it has the advantage of built-in brackets, so there can never be any doubt about what is intended. For example, in the Adams and Spivak notation,

$$1 + x\Big|_a^b \qquad \text{and} \qquad (1 + x)\Big|_a^b$$

mean $1 + b - a$ and $(1 + b) - (1 + a)$, respectively.

Because $\sin(x) \geq 0$ for $x \in [0, \pi]$, the area is simply the integral

$$\int_0^\pi \sin x \,\mathrm{d}x = (-\cos \pi) - (-\cos 0) = 1 - (-1) = 2$$

because $\dfrac{\mathrm{d}}{\mathrm{d}x}(-\cos x) = \sin x$.

14.2 Areas between curves

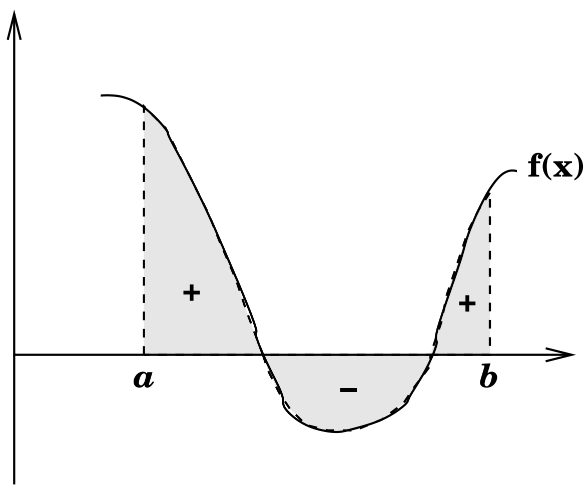
Consider a curve defined by a function f on an interval $[a, b]$ with $f(x) \geq 0$ for $x \in [a, b]$. We have already seen that the area between this positive curve and the x-axis is given by

$$A = \int_a^b f.$$

What happens if f changes sign in $[a, b]$? In general, if we calculate the integral

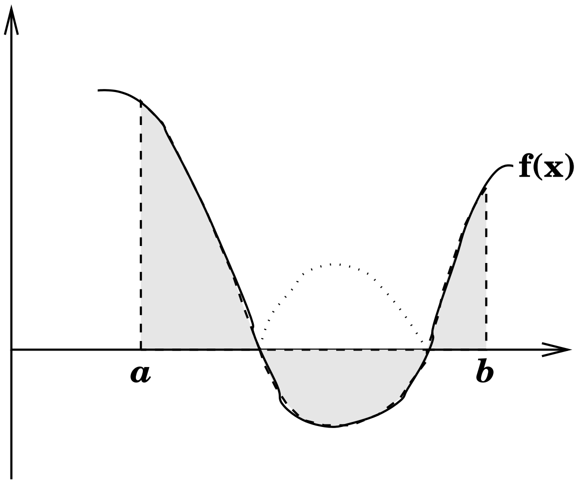
$$\int_a^b f$$

we may get a *negative* answer, because the integral represents the area above the x-axis minus the area below the x-axis.



If we want to calculate the total area between this curve and the x-axis, we take the modulus of the integrand (which is always non-negative), that is, we evaluate

$$A = \int_a^b |f|.$$



Example 14.5 Find the area between the graph of $f(x) = 2x^2 - 1$ and the x-axis between $x = 0$ and $x = 1$.

First, observe that $f = g'$ for $g(x) = \frac{2}{3}x^3 - x$. We need to compute

$$A = \int_0^1 |f|,$$

so we have to find $|f|$ first. Recall that

$$|x| = \begin{cases} x & \text{if} \quad x \geq 0 \\ -x & \text{if} \quad x < 0. \end{cases}$$

This means that to compute $|f|$ we have to find out where f is positive and where it is negative. Since $f(0) = -1$ and $f(1) = 1$, the function f changes sign on $[0, 1]$. To find out where, we consider

$$2x^2 - 1 = 0 \iff x^2 = \frac{1}{2} \iff x = \pm\frac{1}{\sqrt{2}}.$$

We conclude that $f(x) < 0$ for $x \in [0, \frac{1}{\sqrt{2}})$ and $f(x) \geq 0$ for $x \in [\frac{1}{\sqrt{2}}, 1]$. Hence

$$|f(x)| = \begin{cases} -f(x) & \text{if} \quad x \in [0, \frac{1}{\sqrt{2}}) \\ f(x) & \text{if} \quad x \in [\frac{1}{\sqrt{2}}, 1]. \end{cases}$$

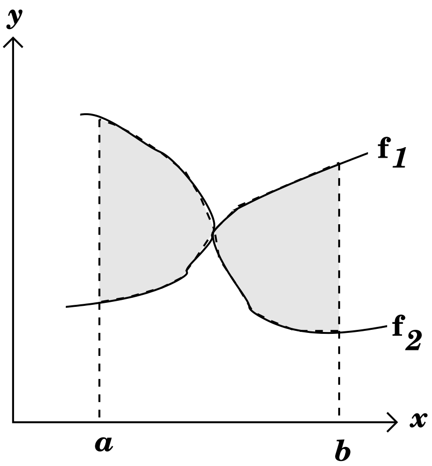
So

$$\begin{aligned} A &= \int_0^1 |f| = \int_0^{1/\sqrt{2}} (-f) + \int_{1/\sqrt{2}}^1 f = -\int_0^{1/\sqrt{2}} f + \int_{1/\sqrt{2}}^1 f \\ &= -\left\{g(1/\sqrt{2}) - g(0)\right\} + g(1) - g(1/\sqrt{2}) = g(0) + g(1) - 2g(1/\sqrt{2}) \\ &= 0 - \frac{1}{3} - 2\left(\frac{1}{3}\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\right) = -\frac{1}{3} - \sqrt{2}\left(\frac{1}{3} - 1\right) = \frac{2\sqrt{2} - 1}{3}. \end{aligned}$$

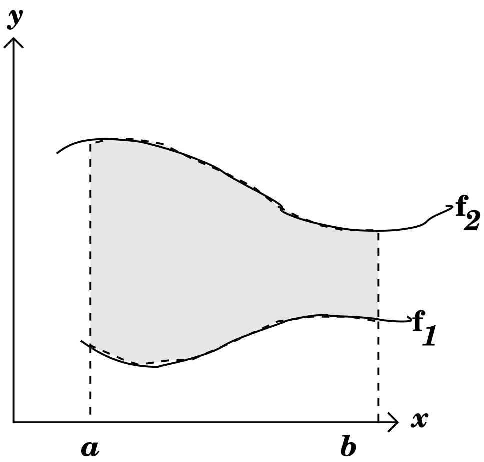
Note that, as it should be, $A > 0$. At the same time, $\int_0^1 f = g(1) - g(0) = -\frac{1}{3}$.

We have just computed the area between two curves: the graph of f and the x-axis. More generally, the area between two curves, given by the graphs of two functions f_1 and f_2, is given by

$$A = \int_a^b |f_2 - f_1|.$$



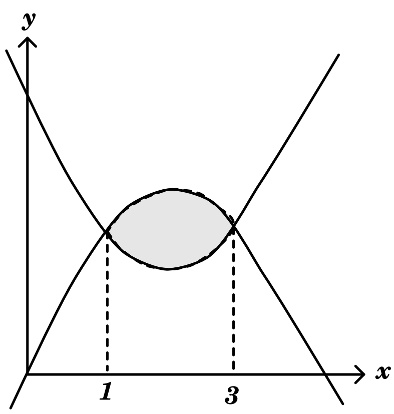
or



Example 14.6 Find the area A enclosed by the graphs of $f_1(x) = x^2 - 4x + 6$ and $f_2(x) = 4x - x^2$ between the two points where f_1 and f_2 intersect.

The curves intersect where

$$\begin{aligned} x^2 - 4x + 6 &= 4x - x^2 \\ \iff 2(x^2 - 4x + 3) &= 0 \\ \iff 2(x-3)(x-1) &= 0 \\ \iff x = 1 &\text{ or } x = 3. \end{aligned}$$



Hence
$$A = \int_1^3 [(4x - x^2) - (x^2 - 4x + 6)]\mathrm{d}x = \int_1^3 (-2x^2 + 8x - 6)\mathrm{d}x = \frac{8}{3}.$$

14.3 Improper integrals

So far, we have always assumed that the interval of integration $[a, b]$ is finite and that the function f is bounded on $[a, b]$. However, this is not always the case. Integrals over infinite intervals, or where the integrand is unbounded, are called **improper integrals**, and they need to be defined as extensions of the definite integrals we've already seen.

14.3.1 Integrals over infinite regions

Definition 14.1 *Consider the definite integral*
$$\int_a^b f(x)\,\mathrm{d}x$$
for any number $b > a$. If, as $b \to \infty$, this integral tends to a finite limit L, then we define the improper integral
$$\int_a^\infty f(x)\,\mathrm{d}x := \lim_{b\to\infty} \int_a^b f(x)\,\mathrm{d}x = L$$
*and say that it **converges** to L. If no limit exists, we say that the improper integral **diverges**.*

Similarly, we define
$$\int_{-\infty}^b f(x)\,\mathrm{d}x := \lim_{a\to-\infty} \int_a^b f(x)\,\mathrm{d}x$$
where this limit exists.

(Compare the terminology for infinite series in Chapter 7.)

Example 14.7 Evaluate $\displaystyle\int_1^\infty \frac{1}{x^2}\,\mathrm{d}x$.

$$\int_1^\infty \frac{1}{x^2}\,\mathrm{d}x = \lim_{b\to\infty} \int_1^b \frac{1}{x^2}\,\mathrm{d}x = \lim_{b\to\infty} \left[-\frac{1}{x}\right]_1^b = \lim_{b\to\infty} \left(1 - \frac{1}{b}\right) = 1.$$

Example 14.8 Evaluate $\displaystyle\int_{-\infty}^{1} \cos(x)\,\mathrm{d}x$.

$$\int_{-\infty}^{1} \cos x\,\mathrm{d}x = \lim_{a\to-\infty}\int_{a}^{1} \cos x\,\mathrm{d}x = \lim_{a\to-\infty}[\sin x]_a^1 = \lim_{a\to-\infty}(-\sin a)\,,$$

which does not exist (because $\sin a$ oscillates between 1 and -1 as $a \to -\infty$), and so the integral diverges.

14.3.2 Integrals of unbounded functions

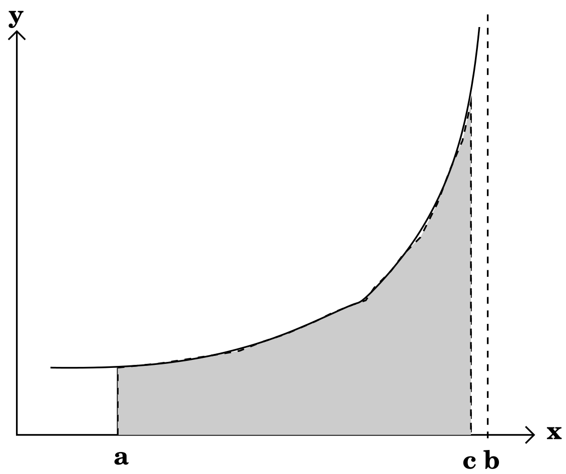
Definition 14.2 *For a function f that is continuous on $[a, b)$ but unbounded at b, we consider the definite integral*

$$\int_a^c f(x)\,\mathrm{d}x$$

for any number $a < c < b$. If, as $c \to b^-$, this integral tends to a finite limit L, then we say that the improper integral converges to L and we write

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{c\to b^-}\int_a^c f(x)\,\mathrm{d}x = L.$$

Otherwise, we say that the improper integral diverges.



Similarly, for an integral

$$\int_a^b f(x)\,\mathrm{d}x$$

where f is continuous on $(a, b]$ and unbounded at a, we consider the limit as $c \to a^+$,

$$\lim_{c\to a^+}\int_c^b f(x)\,\mathrm{d}x.$$

If this limit exists and equals L, then the improper integral converges to L, otherwise the improper integral diverges.

Example 14.9 Evaluate $\displaystyle\int_0^1 x^{-2/3}\,\mathrm{d}x$.

Note that the integrand is undefined at $x = 0$. We use the fact that $\dfrac{\mathrm{d}}{\mathrm{d}x}3x^{1/3} = x^{-2/3}$; then, by the FTC,

$$\begin{aligned}\int_0^1 x^{-2/3}\,\mathrm{d}x &= \lim_{c\to 0^+}\int_c^1 x^{-2/3}\,\mathrm{d}x \\ &= \lim_{c\to 0^+}\left[3x^{1/3}\right]_c^1 = \lim_{c\to 0^+}\left(3 - 3c^{1/3}\right) = 3.\end{aligned}$$

Example 14.10 Evaluate $\displaystyle\int_{-1}^0 \frac{1}{x^2}\,\mathrm{d}x$.

The integrand is undefined at $x = 0$.

$$\int_{-1}^0 \frac{1}{x^2}\,\mathrm{d}x = \lim_{c\to 0^-}\int_{-1}^0 \frac{1}{x^2}\,\mathrm{d}x = \lim_{c\to 0^-}\left[-\frac{1}{x}\right]_0^c = \lim_{c\to 0^-}\left(-\frac{1}{c} - 1\right),$$

which does not exist, so the integral diverges.