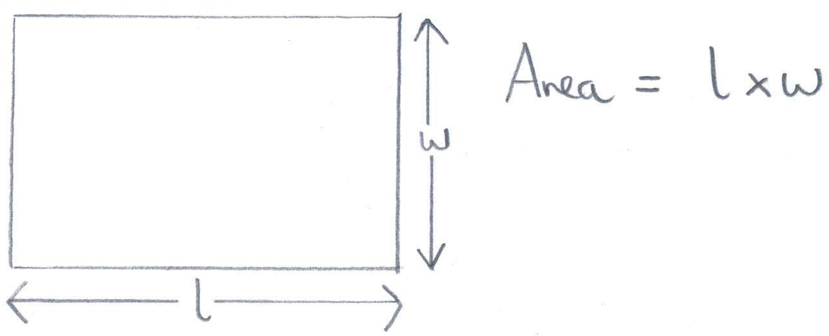
0.2 Motivation: The Problem of Measure

> *Mathematics is facts; just as houses are made of stones, so is mathematics made of facts; but a pile of stones is not a house and a collection of facts is not mathematics.*
>
> Henri Poincaré

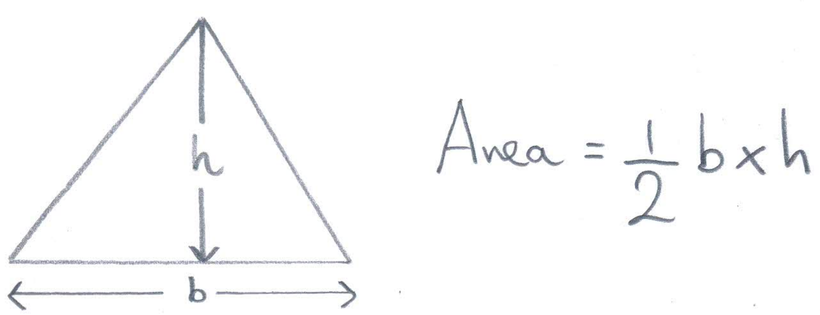
The theme of this module is arguably how to assign a size to certain sets — usually shapes and solids (you will probably disagree with this in time!). In everyday life this is usually pretty straightforward; we

- count: $\{a, b, c, \ldots, x, y, z\}$ has 26 letters.
- take measurements: length (in one dimension), area (in two dimensions), volume (in three measurements) or time;
- calculate: rates of radioactive decay.

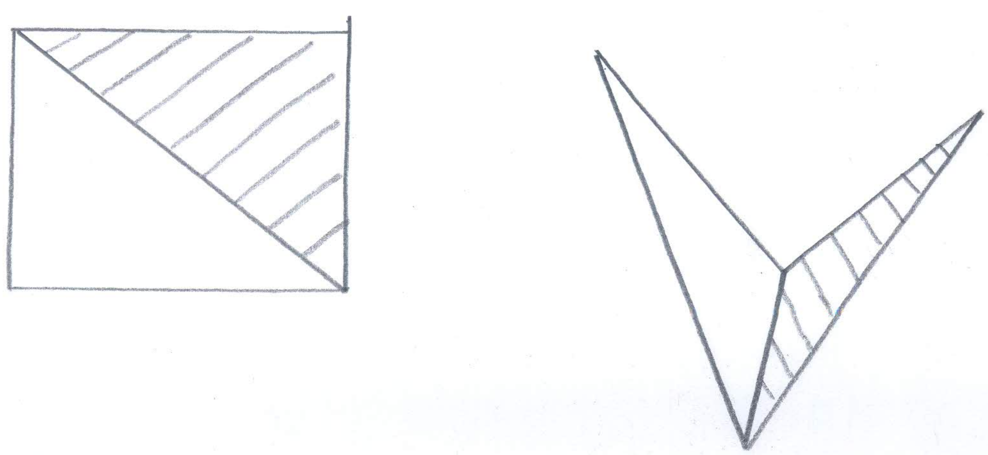
In each case we compare (and express the outcome) with respect to some base unit; most of the measurements just mentioned are supposed to be intuitively clear. Nevertheless, let's have a closer look at areas:



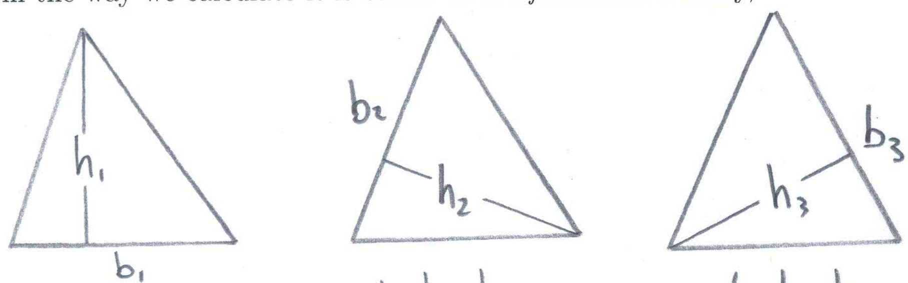
An even more flexible shape than the rectangle is the triangle:



Triangles are actually more basic than rectangles since we can represent every rectangle, and actually and odd-shaped quadrangle, as the 'sum' of two non-overlapping triangles:



In doing so we have *tacitly* assumed a few things. For the triangles we have chosen a *particular* base line and the corresponding height arbitrarily. But the concept of *area* should not depend on such a choice and the calculation this choice entails. Independence of the area from the way we calculate it is called *well-definedness*. Plainly,



$$\frac{1}{2} h_1 b_1 = \frac{1}{2} h_2 b_2 = \frac{1}{2} h_3 b_3$$

Notice that this allows us the most convenient method to work out in the area. In calculating the area of a quadrangle we actually used two assumptions:

- the area of non-overlapping (disjoint[1]) sets can be added, i.e.

$$A(X \cup Y) = A(X) + A(Y) \quad \text{if } X \cap Y = \emptyset$$

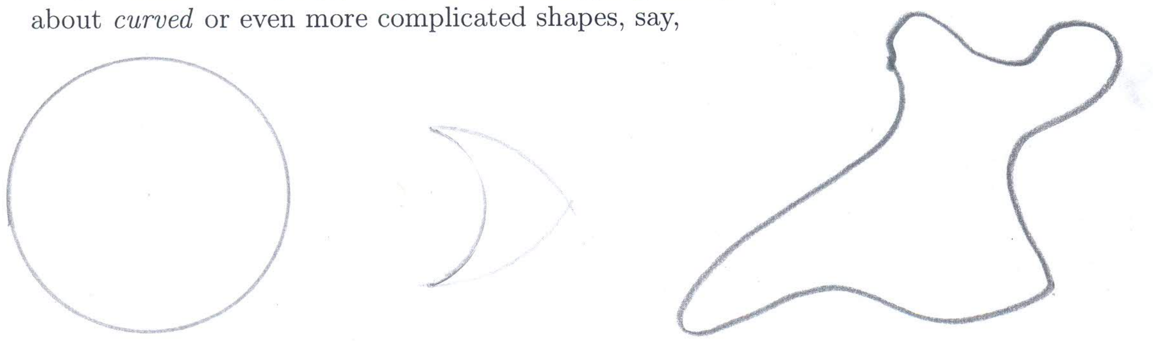
- congruent triangles have the same area[2].

This shows that the least we should expect from a reasonable area measure A is that it is

well-define, take values in $[0, \infty]$, and $A(\emptyset) = 0$;
additive, i.e. $A(X \cup Y) = A(X) + A(Y)$ whenever $X \cap Y = \emptyset$.

An additional property is that area is *invariant* under congruences.

The above rules allow us to measure arbitrarily odd-looking *polygons*[3] using the following recipe: dissect the polygon into non-overlapping triangles and add their areas. But what about *curved* or even more complicated shapes, say,



[1] empty intersection

[2] [Ex:] argue using the idea of congruent triangles why the area *should* be half the base times the perpendicular height — this argument here takes the area of a triangle as fundamental

[3] a figure formed by three or more points in the plane joined by line segments

Here is *one* possibility for the circle: inscribe a regular j-sided polygonal[4] into the circle, subdivide it into congruent triangles, find the area of each of these slices and then add all j pieces:

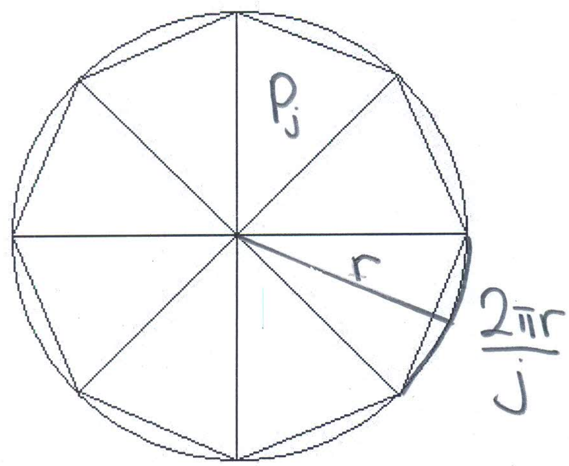


Figure 1: The area of the circle is approximated by triangles. Note the similarity with differentiation — we approximate the slope of the tangent and then take a limit. Here we repeat the trick.

In the next step increase $j \to j + 1$ by increasing by one the number of points on the circumference and repeat the above procedure. Eventually[5],

$$\text{Area} = \lim_{j \to \infty} j \times \text{Area}(P_j) = \pi r^2$$

Again, there are a few problems: does the limit exist? Is it submissible to subdivide a set into arbitrarily many subsets — each of vanishingly small area? Is the procedure independent of the particular subdivision? In fact, nothing should have stopped us from paving the circle with ever smaller squares! For a reasonable notion of area measure the answer to these questions must be assumed to be *yes*. However, finite additivity is not enough for this and we have to use instead:

$$\text{Area}\left(\dot{\bigcup_{j=1}^{\infty}} P_j\right) = \sum_{j=1}^{\infty} \text{Area}(P_j)$$

It can be show that an area measure satisfying all these conditions is powerful enough to cater for all our everyday needs and for much more. We will also that this good notion of area measure allows us to introduce integrals, basically starting with the naïve (but valid) idea that the integral of a positive function should be the same as the area of the set between the graph of the function and the x-axis.

[4]made from points spaced at an equal distance around the circle

[5][Ex]: find this limit by approximating the bases by the arc-length between each point — and the perpendicular height by the radius, r.

Chapter 1

Integration

Although this may seem a paradox, all exact science is dominated by the idea of approximation.

Bertrand Russell

In this chapter we rigourously define the integral of *continuous functions defined on a closed interval* and explore some its properties.

1.1 The Definite Integral: Riemann sums

Suppose that $f : [a, b] \to \mathbb{R}$ is a continuous function. If $f \geq 0$, then one can approximate the area under $y = f(x)$ on $[a, b]$ by drawing rectangles:

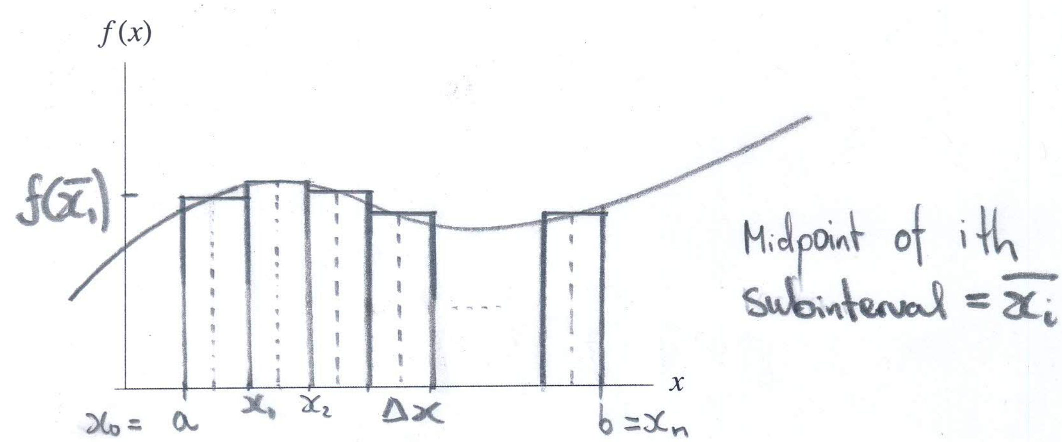


Figure 1.1: Follow the process below to approximate the area under the curve using rectangles.

(i) Divide the interval $[a, b]$ into $n \geq 2$ equal pieces.

(ii) Draw a rectangle on each subinterval with height equal to the value of $f(x)$ at the midpoint of each interval.

Suppose that the length of each subinterval is Δx. Then we have

$$n \times \Delta x = b - a$$

$$\Rightarrow \Delta x = \frac{b-a}{n}$$

In particular we have $x_1 = a + \Delta x$, $x_2 = a + 2\Delta x, \ldots$, $x_i = a + i\Delta x$. Let $\overline{x_i}$ be the midpoint of the ith subinterval. In this notation we have that the area under the curve is approximated by

$$A \approx f(\bar{x}_1)\Delta x + f(\bar{x}_2)\Delta x + \cdots + f(\bar{x}_n)\Delta x$$

$$= \sum_{i=1}^{n} f(\bar{x}_i)\Delta x.$$

Intuitively, one expects that if we choose a larger n (i.e., more subintervals, and consequently narrower rectangles) then the total area of the rectangles is a better approximation of the area under $y = f(x)$. We take the limit as $n \to \infty$ to therefore define this area:

1.1.1 Definition

Let $f : [a, b] \to \mathbb{R}$ be continuous. Then *the integral of f on* $[a, b]$, in the notation above, is given by:

$$\int_a^b f(x), \, dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(\overline{x_i})\Delta x.$$

Remarks

1. The sums on the right-hand side here is known as a Riemann sum. That f is continuous is a sufficient condition for the convergence of such a sum.

2. What if $f(x) \not\geq 0$??

3. Here the function $f(x)$ is the *integrand*, the a is the *lower limit of integration* and b is the *upper limit of integration.* When a and b are constants, then the definite integral is a number and does not depend on x; in fact

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

In situations like this where a variable such as x, t, s appears but does not affect the value of the expression, the variables x, t, s are called *dummy variables.*

Further Remarks

This notion of integration has been known since the middle of the 19th century. The theory has since undergone several revolutions particularly the introduction of the Darboux Integral and the even more powerful Lebesgue Integral which extend this definition in rigour and, in the case of the Lebesgue Integral, to a much broader class of functions than just the continuous functions. The Darboux Integral appeals to arbitrary partitions of $[a, b]$ and instead of looking at the midpoints of the subintervals it instead focusses on maxima and minima of the function on these subintervals and constructs upper and lower bounds for the integral/area. Then the limit is taken over all partitions of $[a, b]$ — if the upper and lower bounds agree then this is defined as the integral. One consequence of this is that for functions that are Riemann integrable we don't need to look at $\bar{x}_i$ but any point in $[x_{i-1}, x_i]$.

1.1.2 Proposition

Suppose that $f : [a, b] \to \mathbb{R}$, $g : [a, b] \to \mathbb{R}$ are continuous and $k \in \mathbb{R}$, with $a < b$. Then we have the following:

1. $$\int_a^b kf(x)\,dx = k\int_a^b f(x)\,dx.$$

2. $$\int_a^b [f(x) \pm g(x)]\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx.$$

3. **If $f(x) \geq 0$ for all $x \in [a, b]$** *then*

$$\int_a^b f(x)\,dx \geq 0.$$

4. **If $f(x) \geq g(x)$ for all $x \in [a, b]$**, *then*

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

5. $$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

6. **Suppose that f, g are continuous on a closed interval containing a, b and $c \in \mathbb{R}$:**

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

7. **Where $m \in \mathbb{R}$ and $M \in \mathbb{R}$ are the minimum and maximum of f on $[a, b]$:**

$$m(b-a) \leq \int_a^b f(x)\,dx \leq M(b-a). \tag{1.1}$$

8.

$$\int_a^b f(x)\,dx = (b-a)f(c) \quad \text{for some } c \in [a,b].$$

The Mean Value Theorem for Integrals.

Remark

Here we use properties of infinite limits that we didn't prove (but sometimes used) in MS2001. It would be a good exercise to recast Proposition 2.1.4 (Calculus of Limits) it terms of the limit as $x \to \infty$... these facts remain true in this case.

Proof. 1. In the appropriate notation:

$$\int_a^b kf(x)dx = \lim_{n\to\infty} \sum_{i=1}^n kf(\bar{x}_i)\Delta x$$

$$= \lim_{n\to\infty} k \sum_{i=1}^n f(\bar{x}_i)\Delta x$$

$$= k \times \lim_{n\to\infty} \sum_{i=1}^n f(\bar{x}_i)\Delta x = k\int_a^b f(x)\,dx.$$

2. Here we prove for the '+ case ('−'?). Again;

$$\int_a^b (f(x)+g(x))dx = \lim_{n\to\infty} \sum_{i=1}^n (f(\bar{x}_i)+g(\bar{x}_i))\Delta x$$

$$= \lim_{n\to\infty}\left(\sum_{i=1}^n f(\bar{x}_i)\Delta x + \sum_{i=1}^n g(\bar{x}_i)\Delta x\right)$$

Now the limit of a sum is the sum of the limits:

$$= \lim_{n\to\infty}\sum_{i=1}^{n} f(\bar{x}_i)\Delta x + \lim_{n\to\infty}\sum_{i=1}^{n} g(\bar{x}_i)\Delta x$$

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

3. Suppose $f(x) \geq 0$ so that

$$\int_a^b f(x)\,dx = \lim_{n\to\infty}\sum_{i=1}^{n} \underbrace{\underbrace{f(\bar{x}_i)}_{\geq 0}\underbrace{\Delta x}_{\geq 0}}_{\geq 0} \geq 0.$$

4. This is a corollary of the last proposition as $f(x) \geq g(x)$ is equivalent to $f(x) - g(x) \geq 0$:

$$\Rightarrow \int_a^b (f(x) - g(x))\,dx \geq 0$$

$$\Rightarrow \int_a^b f(x)\,dx - \int_a^b g(x)\,dx \geq 0$$

$$\Rightarrow \int_a^b f(x)\,dx \geq \int_a^b g(x)\,dx.$$

5. Note that $\Delta x = (b-a)/n$. If instead we integrate $b \to a$ we have $(\Delta x)' = (a-b)/n = -(b-a)/n = -\Delta x$. This is how we get a minus •

6. This is only a sketch of a proof using areas. Once we fix $[a, b]$, assume c differs from a and b, there are only three possibilities: $c < a < b$, $a < c < b$ and $a < b < c$:

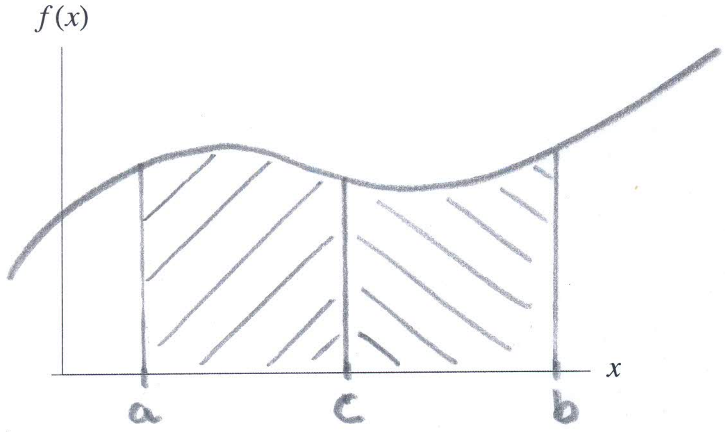


Figure 1.2: If $a < b < c$ it is clear that the integral (area) from $a \to b$ is comprised as that from $a \to c$ plus $c \to b$.

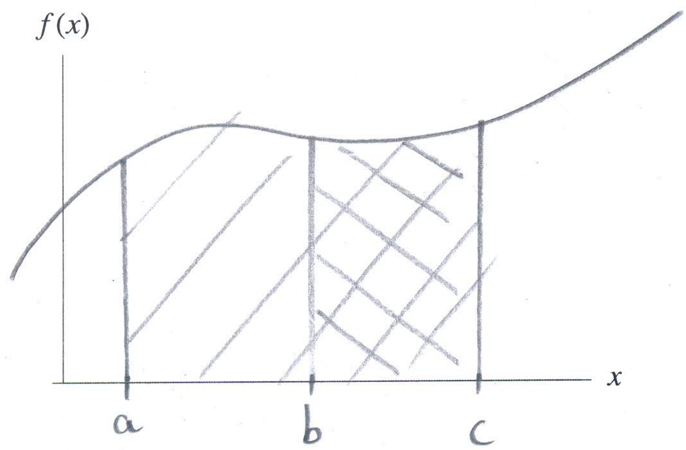


Figure 1.3: If $a < b < c$ then the extra area picked up integrating from $a \to c$ is negated (by the previous part) by the area from $c \to b$.

7. Because f is continuous on the *closed* interval $[a,b]$, it attains a minimum and maximum there, i.e., there are numbers m and M such that $m \leq f(x) \leq M$ for $a \leq x \leq b$ *and* there are points x_m and x_M in $[a,b]$ such that $f(x_m) = m,\ f(x_M) = M$:

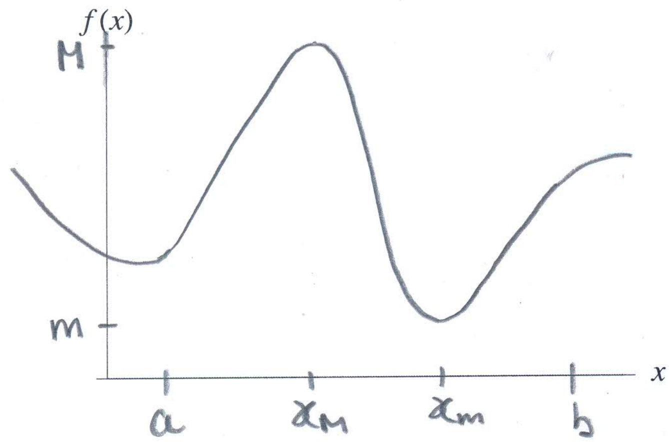


Figure 1.4: A continuous function on a closed interval attains it's absolute max and min on the interval.

Thus

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(\bar{x}_i)\Delta x \leq \lim_{n\to\infty} \sum_{i=1}^{n} M\,\Delta x$$
$$= M(b-a)$$

A similar proof holds to show the lower bound $m(b-a)$. Both are neatly exhibited in a picture:

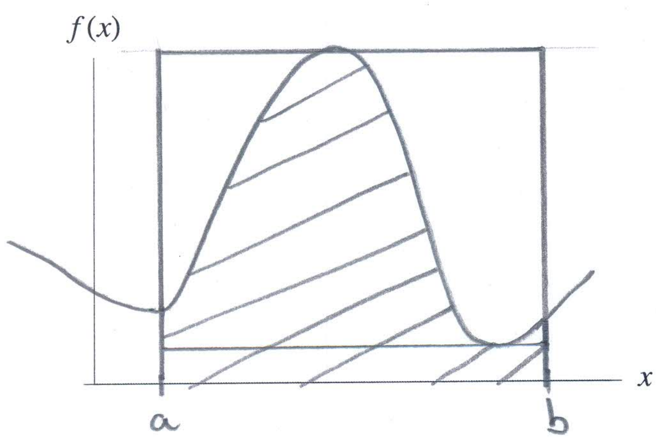


Figure 1.5: Clearly the area is bounded above and below by these rectangles.

8. Rewrite (1.1) as

$$m \leq \frac{1}{b-a}\int_a^b f(x)\,dx \leq M.$$

That is,

$$f(x_m) \leq \frac{1}{b-a}\int_a^b f(x)\,dx \leq f(x_M),$$

where $x_m, x_M \in [a,b]$. This says that the integral (a number) lies between the values that f takes at the points x_m and x_M. The *Intermediate Value Theorem* for continuous functions now implies that there exists a point c lying between x_m and x_M — therefore $c \in [a,b]$ — such that

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

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

Examples

1. Given that $\int_4^9 f(x)\,dx = 38$, can you deduce the value of $\int_9^4 f(t)\,dt$? Justify your answer.

Solution:

$$\int_9^4 f(t)\,dt = \int_9^4 f(x)\,dx = -\int_4^9 f(x)\,dx = -38$$

2. Use (1.1) to find largest and smallest possible values of

$$\int_1^5 (x-2)^2\,dx.$$

Solution: Using the Closed Interval Method, the maxima and minima of $f(x) = (x-2)^2$ on $[1,5]$ are found at endpoints, points where $f' = 0$ and points where f is not differentiable:

$$f'(x) = 2(x-2) = 0 \text{ at } 2.$$

$$f(1) = (1-2)^2 = +1 \qquad f(5) = (5-2)^2 = +9 = M$$

$$f(2) = (2-2)^2 = 0 = m$$

In the next two sections we shall examine how one computes the value of a definite integral — hopefully not from first principles!

Exercises

1. Find lower and upper bounds for:

(i) $\displaystyle\int_2^5 (3x+1)\,dx$ Ans: 21 and 48 (ii) $\displaystyle\int_{-1}^2 \frac{x}{x+2}\,dx$ Ans: -3 and $3/2$.

1.2 The First Fundamental Theorem of Calculus

This theorem makes precise the idea if we integrate a function, then differentiate the result, we get back the original function. Consider definite integrals where one of the limits of integration is a variable, not a constant. Then the value of the definite integral of the (continuous) integrand is also variable, i.e., it defines a new function — that is also an anti-derivative (see later) of the integrand. The *First Fundamental Theorem of calculus* gives a simple formula for differentiating such functions.

1.2.1 First Fundamental Theorem of Calculus

Let $f : [a,b] \to \mathbb{R}$ be continuous. Define $g : [a,b] \to \mathbb{R}$ by

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

Then g differentiable and hence continuous on (a,b) with derivative $g'(x) = f(x)$.

Proof. Fix $x \in (a,b)$. We calculate the derivative of g from first principles: