Mathematics Learning Centre

The University of Sydney

Introduction to Integration
Part 2: The Definite Integral

Mary Barnes

Contents

1 Introduction

This unit deals with the *definite integral.* It explains how it is defined, how it is calculated and some of the ways in which it is used.

We shall assume that you are already familiar with the process of finding indefinite integrals or primitive functions (sometimes called anti-differentiation) and are able to 'anti-differentiate' a range of elementary functions. If you are not, you should work through **Introduction to Integration Part I: Anti-Differentiation**, and make sure you have mastered the ideas in it before you begin work on this unit.

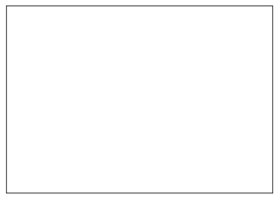
1.1 Objectives

By the time you have worked through this unit you should:

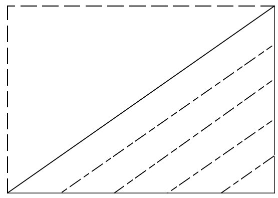
- Be familiar with the definition of the definite integral as the limit of a sum;
- Understand the rule for calculating definite integrals;
- Know the statement of the Fundamental Theorem of the Calculus and understand what it means;
- Be able to use definite integrals to find areas such as the area between a curve and the x-axis and the area between two curves;
- Understand that definite integrals can also be used in other situations where the quantity required can be expressed as the limit of a sum.

2 Finding Areas

Areas of plane (i.e. flat!) figures are fairly easy to calculate *if they are bounded by straight lines.*
The area of a rectangle is clearly the length times the breadth.
The area of a right-angled triangle can be seen to be half the area of a rectangle (see the diagram) and so is half the base times the height.

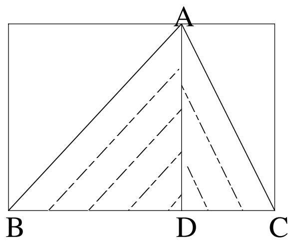


Area of rectangle
=length $\times$ breadth

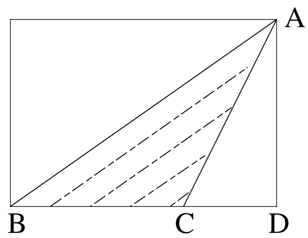


Area of triangle
$=\frac{1}{2}$ area of rectangle
$=\frac{1}{2}$ length $\times$ breadth

The areas of other triangles can be found by expressing them as the sum or the difference of the areas of right angled triangles, and from this it is clear that for any triangle this area is half the base times the height.

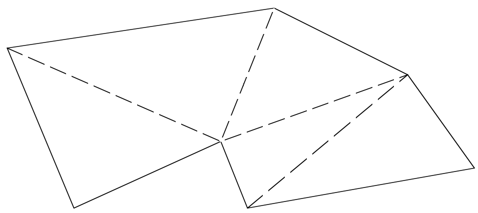


Area of $\triangle$ ABC
= area of $\triangle$ ABD
+ area of $\triangle$ ACD



Area of $\triangle$ ABC
= area of $\triangle$ ABD
− area of $\triangle$ ACD

Using this, we can find the area of any figure bounded by straight lines, by dividing it up into triangles (as shown).



Areas *bounded by curved lines* are a much more difficult problem, however. In fact, although we all feel we know intuitively what we mean by the area of a curvilinear figure, it is actually quite difficult to define precisely. The area of a figure is quantified by asking 'how many units of area would be needed to cover it?' We need to have some unit of area in mind (e.g. one square centimetre or one square millimetre) and imagine trying to cover the figure with little square tiles. We can also imagine cutting these tiles in halves, quarters etc. In this way a rectangle, and hence any figure bounded by straight lines, can be dealt with, but a curvilinear figure can never be covered exactly.

We are therefore forced to rely on the notion of limit in order to define areas of curvilinear figures.

To do this, we make some simple assumptions which most people will accept as intuitively obvious. These are:-

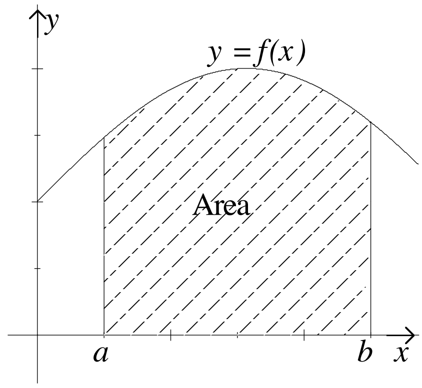
1. If one figure is a subset of a second figure, then the area of the first will be less than or equal to that of the second.

2. If a figure is divided up into non-overlapping pieces, the area of the whole will be the sum of the areas of the pieces.

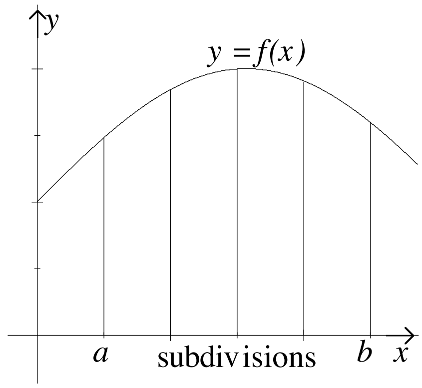
Using these assumptions, we can approximate to curved figures by means of polygons (figures with straight line boundaries), and hence *define* the area of the curved figure as the limit of the areas of the polygons as they 'approach' the curved figure (in some sense yet to be made precise).

3 Areas Under Curves

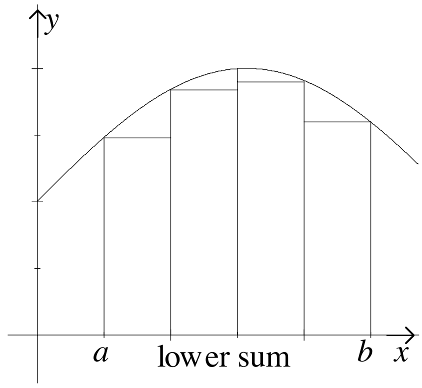
Let us suppose that we are given a positive function $f(x)$ and we want to find the area enclosed between the curve $y = f(x)$, the x-axis and the lines $x = a$ and $x = b$. (The shaded area in the diagram.)
If the graph of $y = f(x)$ is not a straight line we do not, at the moment, know how to calculate the area precisely.



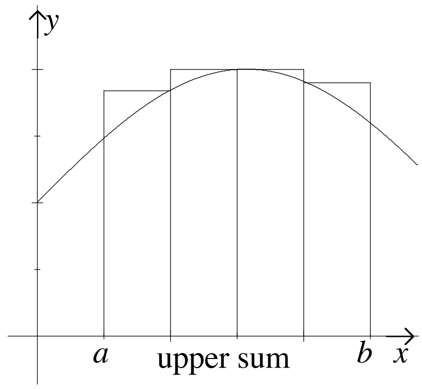
We can, however, approximate to the area as follows: **First** we divide the area up into strips as shown, by dividing the interval from a to b into equal subintervals, and drawing vertical lines at these points.



Next we choose the *least* value of $f(x)$ in each subinterval and construct a rectangle with that as its height (as in the diagram). The sum of the areas of these rectangles is clearly *less* than the area we are trying to find. This sum is called a *lower sum*.



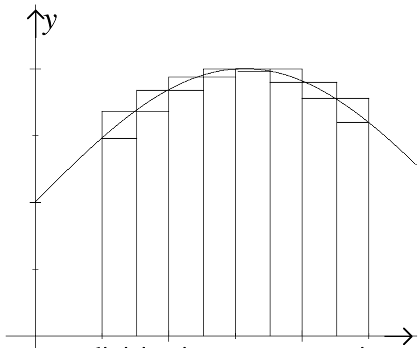
Then we choose the greatest value of $f(x)$ in each subinterval and construct a rectangle with *that* as its height (as in the diagram opposite). The sum of the areas of these rectangles is clearly *greater* than the area we are trying to find. This sum is called an *upper sum*.



Thus we have 'sandwiched' the area we want to find in between an upper sum and a lower sum. Both the upper sum and the lower sum are easily calculated because they are sums of areas of rectangles.

Although we still can't say precisely what the area under the curve is, we know between what limits it lies.

If we now increase the number of strips the area is divided into, we will get new upper and lower sums, which will be closer to one another in size and so closer to the area which we are trying to find. In fact, the larger the number of strips we take, the smaller will be the difference between the upper and lower sums, and so the better approximation either sum will be to the area under the curve.



division into narrower strips

It can be shown that if $f(x)$ is a 'nice' function (for example, a continuous function) the difference between the upper and lower sums *approaches zero* as the number of strips the area is subdivided into approaches infinity.

We can thus define the area under the curve to be:

> the limit of either the upper sum or the lower sum, as the number of subdivisions tends to infinity (and the width of each subdivision tends to zero).

Thus finding the area under a curve boils down to finding the *limit of a sum*.

Now let us introduce some notation so that we can talk more precisely about these concepts.

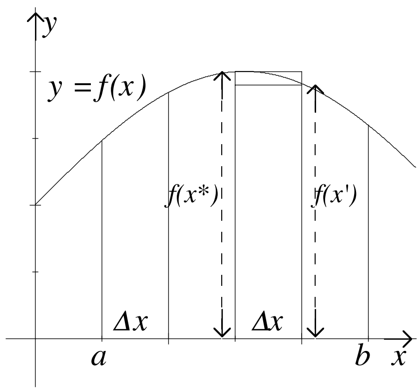
Let us suppose that the interval $[a, b]$ is divided into n equal subintervals each of width Δx. Suppose also that the greatest value of $f(x)$ in the ith subinterval is $f(x_i^*)$ and the least value is $f(x_i^{'})$.
Then the upper sum can be written as:

$$f(x_1^*)\Delta x + f(x_2^*)\Delta x + ... + f(x_n^*)\Delta x$$

or, using summation notation: $\sum_{i=1}^{n} f(x_i^*)\Delta x$.
Similarly, the lower sum can be written as:

$$f(x_1^{'})\Delta x + f(x_2^{'})\Delta x + ... + f(x_n^{'})\Delta x$$

or, using summation notation: $\sum_{i=1}^{n} f(x_i^{'})\Delta x$.



With this notation, and letting A stand for the area under the curve $y = f(x)$ from $x = a$ to $x = b$, we can express our earlier conclusions in symbolic form.

The area lies between the lower sum and the upper sum and can be written as follows:

$$\sum_{i=1}^{n} f(x_i^{'})\Delta x \leq A \leq \sum_{i=1}^{n} f(x_i^*)\Delta x.$$

The area is equal to the limit of the lower sum or the upper sum as the number of subdivisions tends to infinity and can be written as follows:

$$A = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^{'})\Delta x$$

or

$$A = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^{*})\Delta x.$$

3.1 What is the point of all this?

Well, **firstly** it enables us to define precisely what up till now has only been an imprecise intuitive concept, namely, the area of a region with curved lines forming part of its boundary.

Secondly it indicates how we may calculate approximations to such an area. By taking a fairly large value of n and finding upper or lower sums we get an approximate value for the area. The difference between the upper and lower sums tells us how accurate this approximation is. This, unfortunately, is not a very good or very practical way of approximating to the area under a curve. If you do a course in Numerical Methods you will learn much better ways, such as the Trapezoidal Rule and Simpson's Rule.

Thirdly it enables us to calculate areas *precisely*, provided we know how to find finite sums and evaluate limits. This however can be difficult and tedious, so we need to look for better ways of finding areas. This will be done in Section 5.

At this stage, many books ask students to do exercises calculating upper and lower sums and using these to estimate areas. Frequently students are also asked to find the limits of these sums as the number of subdivisions approaches infinity, and so find exact areas. We shall not ask you to do this, as it involves a great deal of computation.

3.2 Note about summation notation

The symbol $\sum$ (pronounced 'sigma') is the capital letter S in the Greek alphabet, and stands for 'sum'.

The expression $\sum_{i=1}^{4} f(i)$ is read 'the sum of $f(i)$ from $i = 1$ to $i = 4$', or 'sigma from $i = 1$ to 4 of $f(i)$'.

In other words, we substitute 1, 2, 3 and 4 in turn for i and add the resulting expressions.

Thus,	$\sum_{i=1}^{4} x_i$	stands for	$x_1 + x_2 + x_3 + x_4$,
	$\sum_{i=1}^{5} i^2$	stands for	$1^2 + 2^2 + 3^2 + 4^2 + 5^2$,
and	$\sum_{i=1}^{2} f(x_1)\Delta x$	stands for	$f(x_1)\Delta x + f(x_2)\Delta x$.

4 The Definition of the Definite Integral

The discussion in the previous section led to an expression of the form

$$A = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x \qquad (1)$$

where the interval $[a, b]$ has been divided up into n equal subintervals each of width Δx and where x_i is a point in the ith subinterval. This is a very clumsy expression, and mathematicians have developed a simpler notation for such expressions. We denote them by

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

which is read as 'the integral from a to b of $f(x)dx$'.

The $\int$ sign is an elongated 's' and stands for 'sum', just as the $\sum$ did previously. The difference is that in this case it means 'the limit of a sum' rather than a finite sum. The dx comes from the Δx as we pass to the limit, just as happened in the definition of $\frac{dy}{dx}$.

Thus the definite integral is defined as *the limit of a particular type of sum* i.e. sums like that given in (1) above, as the width of each subinterval approaches zero and the number of subintervals approaches infinity.

4.1 Notes

1. Although we used the area under a curve as the motivation for making this definition, the definite integral is **not** defined to be the area under a curve but simply the limit of the sum (1).

2. Initially, when discussing areas under curves, we introduced the restriction that $f(x)$ had to be a positive function. This restriction is not necessary for the definition of a definite integral.

3. The definition can be made more general, by removing the requirement that all the subintervals have to be of equal widths, but we shall not bother with such generalisations here.

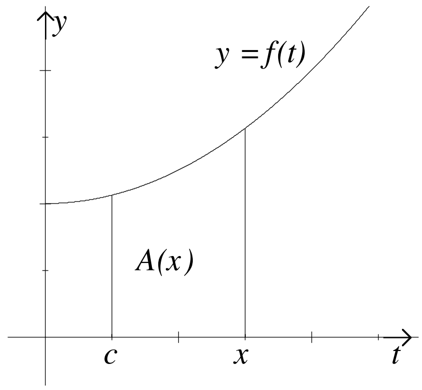
4. Sums such as (1) are called *Riemann sums* after the mathematician Georg Riemann who first gave a rigorous definition of the definite integral.

5. The definition of a definite integral requires that $f(x)$ should be *defined* everywhere in the interval $[a, b]$ and that the limit of the Riemann sums should exist. This will always be the case if f is a continuous function.

5 The Fundamental Theorem of the Calculus

So far, we have defined definite integrals but have not given any practical way of calculating them. Nor have we shown any connection between definite integrals and differentiation.

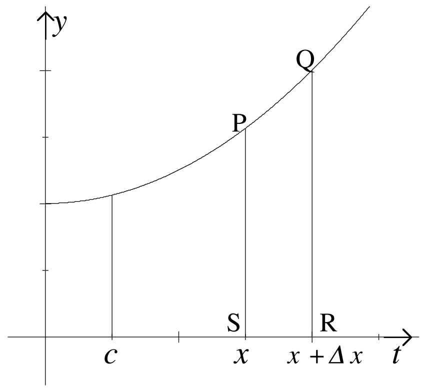
Let us consider the special case where $f(t)$ is a continuous *positive* function, and let us consider the area under the curve $y = f(t)$ from some fixed point $t = c$ up to the variable point $t = x$. For different values of x we will get different areas. This means that the area is a function of x. Let us denote the area by $A(x)$. Clearly, $A(x)$ increases as x increases. Let us try to find the *rate* at which it increases, that is, the derivative of $A(x)$ with respect to x.



At this point, recall how we find derivatives from first principles:
Given a function $f(x)$, we let x change by an amount Δx, so that $f(x)$ changes to $f(x+\Delta x)$. The derivative of $f(x)$ is the limit of

$$\frac{f(x+\Delta x) - f(x)}{\Delta x} \text{ as } \Delta x \to 0.$$

We shall go through this process with $A(x)$ in place of $f(x)$.



When we increase x by Δx, $A(x)$ increases by the area of the figure PQRS. That is, (see the diagram)

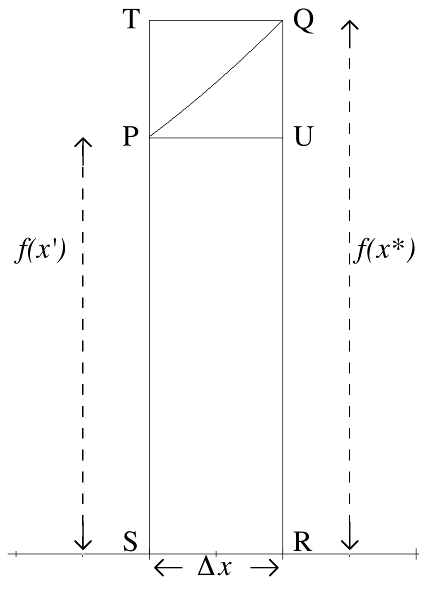
$$A(x+\Delta x) - A(x) = \text{area PQRS}.$$

Now that the area PQRS is bounded by a curved line at the top, but it can be seen to lie in between the areas of two rectangles:

$$\text{area PURS} < \text{area PQRS} < \text{area TQRS}.$$

Both of these rectangles have width Δx. Let the height of the larger rectangle be $f(x^*)$ and the height of the smaller rectangle $f(x')$. (In other words, x^* and x' are the values of x at which $f(x)$ attains its maximum and minimum values in the interval from x to $x + \Delta x$.)

Thus area PURS $= f(x')\Delta x$ and area TQRS $= f(x^*)\Delta x$.
So, $f(x')\Delta x \le A(x+\Delta x) - A(x) \le f(x^*)\Delta x$.



Now if we divide these inequalities all through by Δx, we obtain

$$f(x') \leq \frac{A(x+\Delta x) - A(x)}{\Delta x} \leq f(x^*).$$

Finally, if we let $\Delta x \to 0$, both $f(x')$ and $f(x^*)$ approach $f(x)$, and so the expression in the middle must also approach $f(x)$, that is, the derivative of $A(x)$, $\frac{dA}{dx} = f(x)$.

This result provides the link we need between differentiation and the definite integral.

If we recall that the area under the curve $y = f(t)$ from $t = a$ to $t = x$ is equal to $\int_a^x f(t)dt$, the result we have just proved can be stated as follows:

$$\frac{d}{dx}\int_a^x f(t)dt = f(x). \qquad (2)$$

This is the **Fundamental Theorem of the Calculus**.

In words

> If we differentiate a definite integral with respect to the upper limit of integration, the result is the function we started with.

You may not actually use this result very often, but it is important because we can derive from it the rule for calculating definite integrals:

Let us suppose that $F(x)$ is an anti-derivative of $f(x)$. That is, it is a function whose derivative is $f(x)$. If we anti-differentiate both sides of the equation (2) we obtain

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

Now we can find the value of c by substituting $x = a$ in this expression.

Since $\displaystyle\int_a^a f(t)dt$ is clearly equal to zero, we obtain

$$0 = F(a) + c, \qquad \text{and so} \qquad c = -F(a).$$

Thus $\displaystyle\int_a^x f(t)dt = F(x) - F(a)$, or, letting $x = b$,

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

This tells us how to evaluate a definite integral

- first, find an anti-derivative of the function
- then, substitute the upper and lower limits of integration into the result and subtract.

Note A convenient short-hand notation for $F(b) - F(a)$ is $[F(x)]_a^b$.

To see how this works in practice, let us look at a few examples:

i Find $\int_0^1 x^2 dx$.

An anti-derivative of x^2 is $\frac{1}{3}x^3$, so we write

$$\int_0^1 x^2 dx = \left[\frac{1}{3}x^3\right]_0^1 = \frac{1}{3}(1)^3 - \frac{1}{3}(0)^3 = \frac{1}{3}.$$

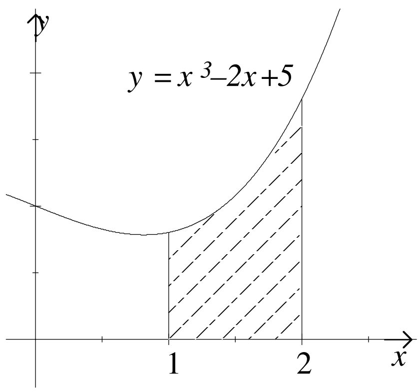
ii Find $\int_0^\pi \sin t dt$.

$$\begin{aligned}\int_0^\pi \sin t dt &= [-\cos t]_0^\pi \\ &= -\cos(\pi) + \cos 0 \\ &= -(-1) + 1 \\ &= 2.\end{aligned}$$

iii Find the area enclosed between the x-axis, the curve $y = x^3 - 2x + 5$ and the ordinates $x = 1$ and $x = 2$.

In a question like this it is always a good idea to draw a rough sketch of the graph of the function and the area you are asked to find. (See below)

If the required area is A square units, then



$$\begin{aligned}A &= \int_1^2 \left(x^3 - 2x + 5\right) dx \\ &= \left[\frac{x^4}{4} - x^2 + 5x\right]_1^2 \\ &= (4 - 4 + 10) - \left(\frac{1}{4} - 1 + 5\right) \\ &= 5\frac{3}{4}.\end{aligned}$$

Exercises 5

1. **a.** $\left[2x^3\right]_2^4$

b. $\left[\frac{1}{x^2}\right]_1^3$

c. $\left[\sqrt{x}\right]_9^{16}$

d. $[\ln x]_2^4$

2. **a.** $\int_4^9 \frac{1}{\sqrt{x}} dx$

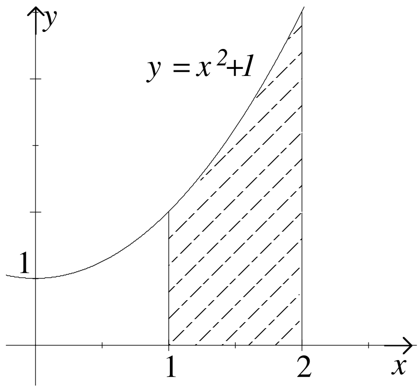
b. $\int_{\frac{-\pi}{2}}^{\frac{\pi}{2}} \cos t dt$

c. $\displaystyle\int_1^2 \frac{1}{y^2}dy$

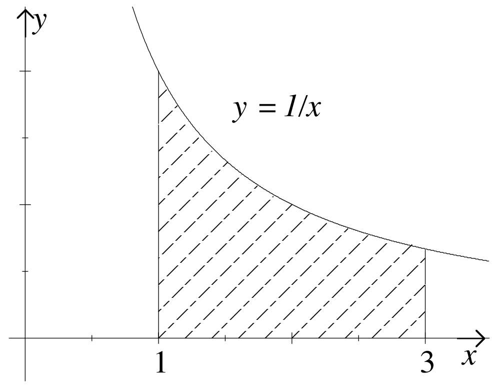
d. $\displaystyle\int_{-2}^{-1}(s^2+2s+2)ds$

3. Find the area of the shaded region in each of the diagrams below:

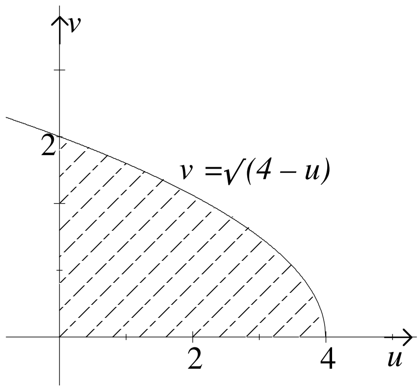
a.



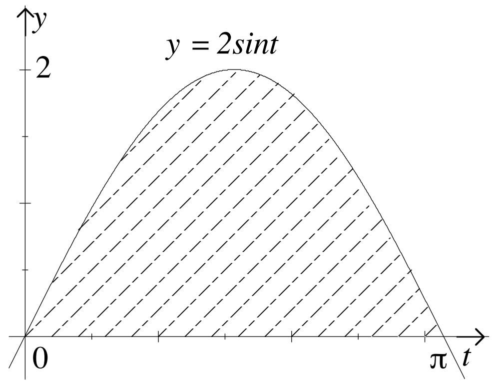
b.



c.



d.



4. Evaluate

a. $\displaystyle\int_0^1 xe^{x^2}dx$

b. $\displaystyle\int_{-2}^{-1}\frac{1}{3-x}dx$

c. $\displaystyle\int_0^{\frac{\pi}{2}}\sin 2y dy$

d. $\displaystyle\int_1^5 \frac{t}{4+t^2}dt$

6 Properties of the Definite Integral

Some simple properties of definite integrals can be derived from the basic definition, or from the Fundamental Theorem of the Calculus. We shall not give formal proofs of these here but you might like to think about them, and try to explain, to yourself or someone else, why they are true.

a.

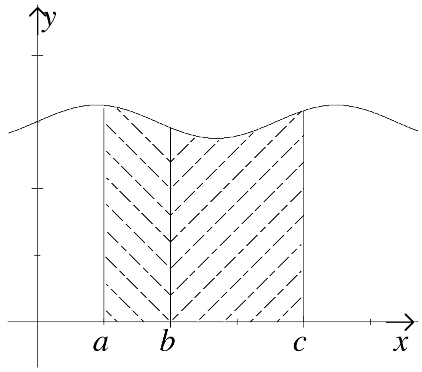
$$\int_a^a f(x)dx = 0.$$

If the upper and lower limits of the integral are the same, the integral is zero. This becomes obvious if we have a positive function and can interpret the integral in terms of 'the area under a curve'.

b. If $a \le b \le c$,

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

This says that the integral of a function over the union of two intervals is equal to the sum of the integrals over each of the intervals. The diagram opposite helps to make this clear if $f(x)$ is a positive function.



c.

$$\int_a^b cf(x)dx = c\int_a^b f(x)dx \quad \text{for any constant } c.$$

This tells us that we can move a constant past the integral sign, but *beware*: we can only do this with constants, never with variables!

d.

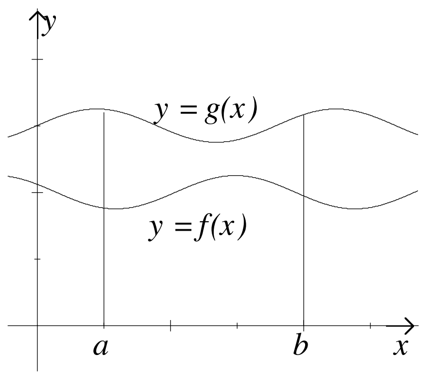
$$\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

That is, the integral of a sum is equal to the sum of the integrals.

e. If $f(x) \le g(x)$ in $[a, b]$ then

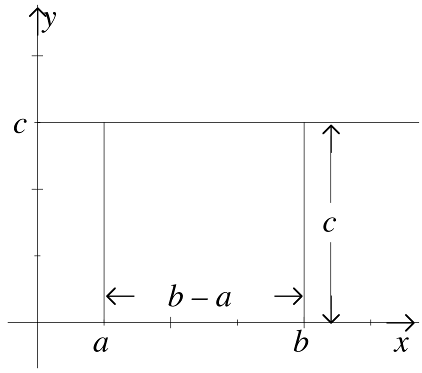
$$\int_a^b f(x)dx \le \int_a^b g(x)dx.$$

That is, integration preserves inequalities between functions. The diagram opposite explains this result if $f(x)$ and $g(x)$ are positive functions.



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

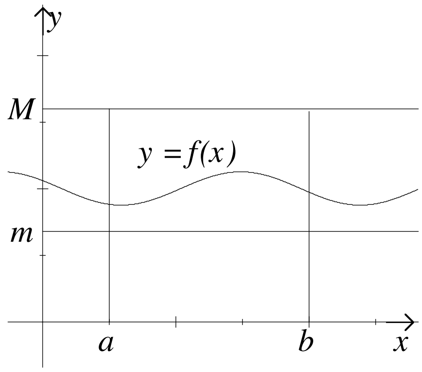
This tells us that the integral of a constant is equal to the product of the constant and the range of integration. It becomes obvious when we look at the diagram with $c > 0$, since the area represented by the integral is just a rectangle of height c and width $b - a$.



g. We can combine (e) and (f) to give the result that, if M is any upper bound and m any lower bound for $f(x)$ in the interval $[a, b]$, so that $m \le f(x) \le M$, then

$$m(b-a) \le \int_a^b f(x)dx \le M(b-a).$$

This, too, becomes clear when $f(x)$ is a positive function and we can interpret the integral as the area under the curve.



h. Finally we extend the definition of the definite integral slightly, to remove the restriction that the lower limit of the integral must be a smaller number than the upper limit. We do this by specifying that

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

For example,

$$\int_2^1 f(x)dx = -\int_1^2 f(x)dx.$$

7 Some Common Misunderstandings

7.1 Arbitrary constants

When you first learned how to find indefinite integrals (anti-derivatives), you probably also learned that it was important to remember always to add an arbitrary constant to the answer.

There is no arbitrary constant in a definite integral.

If we interpret a definite integral as an area, it is clear that its value is a fixed number (the number of units of area in the region). There is no ambiguity, and so no need to add an arbitrary constant - in fact, it is wrong to do so.

When we apply the Fundamental Theorem of the Calculus to finding a definite integral, however, the possibility of an arbitrary constant appears to arise.

For example, in calculating $\int_1^2 x^2 dx$, we have to find an anti-derivative for x^2. The most natural choice would be $\frac{1}{3}x^3$, but instead of that we could choose $\frac{1}{3}x^3 + c$, where c is any constant.

Then,

$$\int_0^1 x^2 dx = \left[\frac{1}{3}x^3 + c\right]_0^1 = \left(\frac{1}{3}(1)^3 + c\right) - \left(\frac{1}{3}(0)^3 + c\right) = \frac{1}{3}.$$

Note that the constants cancel one another out, and we get the same answer as we did before. Thus we might as well take the simplest course, and forget about arbitrary constants when we are calculating definite integrals.

7.2 Dummy variables

What is the difference between $\int_a^b f(x)dx$ and $\int_a^b f(t)dt$?

Let's work them both out in a special case.

$$\int_2^4 \frac{1}{x} dx = [\ln x]_2^4 = \ln 4 - \ln 2.$$

$$\int_2^4 \frac{1}{t} dt = [\ln t]_2^4 = \ln 4 - \ln 2.$$

So **both** integrals give the same answer.

It is clear that the value of a definite integral depends on the *function* and the *limits of integration* but not on the actual variable used. In the process of evaluating the integral, we substitute the upper and lower limits for the variable and so the variable doesn't appear in the answer. For this reason we call the variable in a definite integral a **dummy variable** - we can replace it with any other variable without changing a thing.

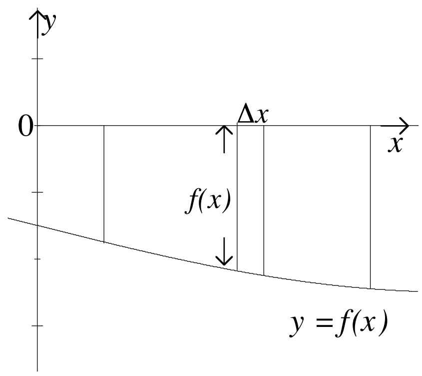
Thus,

$$\int_a^b f(x)dx = \int_a^b f(y)dy = \int_a^b f(t)dt = \int_a^b f(\theta)d\theta.$$

8 Another Look at Areas

We have defined the definite integral $\int_a^b f(x)dx$ as the limit of a particular type of sum, without placing any restrictions on whether the function $f(x)$ is positive or negative.

We know that, if $f(x)$ is positive, $\int_a^b f(x)dx$ is equal to the area between the curve $y = f(x)$, the x-axis and the ordinates $x = a$ and $x = b$, (which we refer to as 'the area under the curve'). The natural question to ask now is: what does $\int_a^b f(x)dx$ equal if $f(x)$ is negative? Can we represent it as an area in this case too; perhaps 'the area above the curve'?
If we go back to the definition of $\int_a^b f(x)dx$ as the limit of a sum, we can see clearly that if $f(x)$ is always negative then each of the terms $f(x_i)\Delta x$ will also be negative (since Δx is positive).



So the sum $\sum_{i=1}^{n} f(x_i)\Delta x$ will be a sum of negative terms and so will be negative too. And when we let n approach infinity and pass to the limit, that will be negative also.

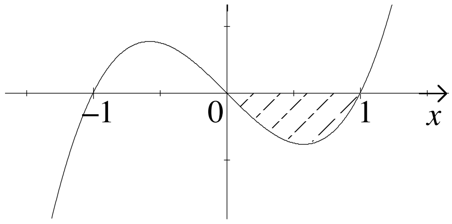
Thus, if $f(x)$ is negative for x between a and b, $\displaystyle\int_a^b f(x)dx$ will also be negative.

Now areas are, by definition, positive. Remember that, in section 1, we explained that we can measure the area of a region by counting the number of little square tiles (each of unit area) needed to cover it. Since we can't cover a region with a negative number of tiles (it doesn't make sense to talk of it) we can't have a negative area. On the other hand, if we ignore the fact that each of the terms $f(x)\Delta x$ is negative, and consider its **numerical value** only, we can see that it is numerically equal to the area of the rectangle shown. And, if we go through the usual process, adding up the areas of all the little rectangles and taking the limit, we find that $\displaystyle\int_a^b f(x)dx$ is **numerically equal** to the area between the curve and the x-axis.

So to find the area, we calculate $\displaystyle\int_a^b f(x)dx$, which will turn out to be negative, and then take its numerical (i.e. absolute) value.

To see this more clearly, let's look at an example. Consider the curve, $y = x(x^2 - 1)$. This is a cubic curve, and cuts the x-axis at -1, 0 and 1. A sketch of the curve is shown below. Let us find the shaded area. First we calculate the definite integral $\int_0^1 x(x^2 - 1)dx$.

$$\begin{aligned}
\int_0^1 x(x^2-1)dx &= \int_0^1 (x^3 - x)dx \\
&= \left[\frac{1}{4}x^4 - \frac{1}{2}x^2\right]_0^1 \\
&= \left(\frac{1}{4} - \frac{1}{2}\right) - (0 - 0) \\
&= -\frac{1}{4}.
\end{aligned}$$



Since $x(x^2 - 1)$ is negative when x lies between 0 and 1, the definite integral is also negative, as expected. We can conclude that the area required is $\frac{1}{4}$ square units.

As a check, let us find the area of the other 'loop' of the curve, i.e. the area between the curve and the x-axis from -1 to 0. Since $x(x^2 - 1)$ is positive for this range of values of x, the area will be given by

$$\begin{aligned}
\int_{-1}^0 x(x^2-1)dx &= \left[\frac{1}{4}x^4 - \frac{1}{2}x^2\right]_{-1}^0 \\
&= (0 - 0) - \left(\frac{1}{4} - \frac{1}{2}\right) \\
&= \frac{1}{4}.
\end{aligned}$$

This is the answer we would expect, since a glance at the diagram shows that the curve has 'point symmetry' about the origin. If we were to rotate the whole graph through 180°, the part of the curve to the left of the origin would fit exactly on top of the part to the right of the origin, and the unshaded loop would fit on top of the shaded loop. So the areas of the two loops are the same.

Now let us calculate $\int_{-1}^1 x(x^2 - 1)dx$.

$$\begin{aligned}
\int_{-1}^1 x(x^2-1)dx &= \left[\frac{1}{4}x^4 - \frac{1}{2}x^2\right]_{-1}^1 \\
&= \left(\frac{1}{4} - \frac{1}{2}\right) - \left(\frac{1}{4} - \frac{1}{2}\right) \\
&= 0.
\end{aligned}$$

This makes it very clear that

a definite integral does not always represent the area under a curve.

We have found that

1. If $f(x)$ is **positive** between a and b, then $\int_a^b f(x)dx$ does represent the area under the curve.

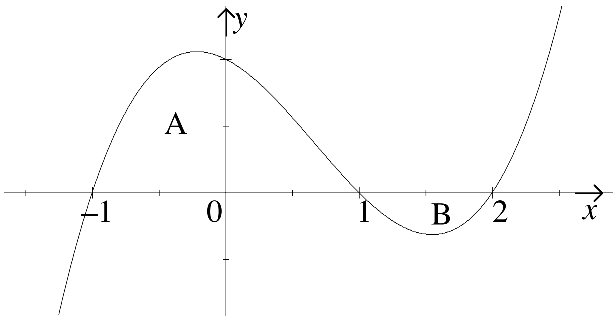
2. If $f(x)$ is **negative** between a and b, then $\left|\int_a^b f(x)dx\right|$ represents the area **above** the curve, since the value of $\int_a^b f(x)dx$ is negative.

3. If $f(x)$ is **sometimes positive and sometimes negative** between a and b, then $\int_a^b f(x)dx$ measures the **difference in area** between the part above the x-axis and the part below the x-axis. (In the example above, the two areas were equal, and so the difference came out to be zero.)

Let's look at another example.

Consider the function $y = (x+1)(x-1)(x-2) = x^3 - 2x^2 - x + 2$.

This is a cubic function, and the graph crosses the x-axis at -1, 1 and 2. A sketch of the graph is shown.



The area marked A is given by

$$
\begin{aligned}
\int_{-1}^{1}(x^3 - 2x^2 - x + 2)dx &= \left[\frac{1}{4}x^4 - \frac{2}{3}x^3 - \frac{1}{2}x^2 + 2x\right]_{-1}^{1} \\
&= \left(\frac{1}{4} - \frac{2}{3} - \frac{1}{2} + 2\right) - \left(\frac{1}{4} + \frac{2}{3} - \frac{1}{2} - 2\right) \\
&= -\frac{4}{3} + 4 = 2\frac{2}{3}.
\end{aligned}
$$

So the area of A is $2\frac{2}{3}$ square units.

The area marked B can be found by evaluating

$$\int_1^2 \left(x^3 - 2x^2 - x + 2\right) dx.$$

This works out as $-\frac{5}{12}$. (The details of the calculation are left to you.)

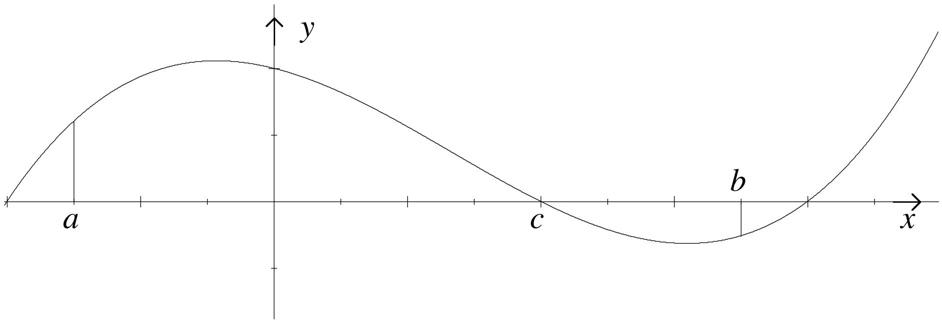
So the area of B is $\frac{5}{12}$ square units.

If we calculate $\int_{-1}^{2}(x^3 - 2x^2 - x + 2)dx$ the answer will be the difference between the area of A and the area of B, that is, $2\frac{1}{4}$ square units. (Check it out for yourself.)

If we want the total area enclosed between the curve and the x-axis we must **add** the area of A and the area of B.

i.e. $2\frac{2}{3} + \frac{5}{12} = 3\frac{1}{12}$ square units.

WARNING In working out area problems you should *always* sketch the curve first. If the function is sometimes positive and sometimes negative in the range you are interested in, it may be necessary to divide the area into two or more parts, as shown below.



The area between the curve and the x-axis from a to b is NOT equal to $\int_a^b f(x)dx$.

Instead, it is $\int_a^c f(x)dx + |\int_c^b f(x)dx|$.

Before you can calculate this, you must find the value of c, i.e. find the point where the curve $y = f(x)$ crosses the x-axis.

Exercises 8

1. Find the area enclosed by the graph of $y = 3x^2(x - 4)$ and the x-axis.

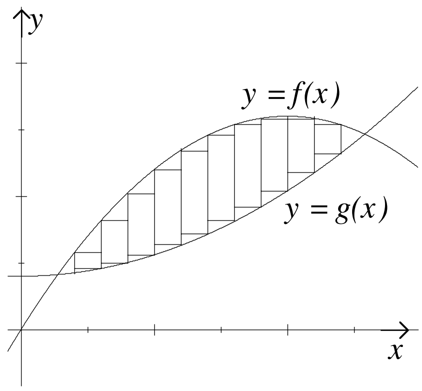
2. i Find the value of $\int_0^{2\pi} \sin x dx$.

ii Find the area enclosed between the graph of $y = \sin x$ and the x-axis from $x = 0$ to $x = 2\pi$.

3. Find the total area enclosed between the graph of $y = 12x(x + 1)(2 - x)$ and the x-axis.

9 The Area Between Two Curves

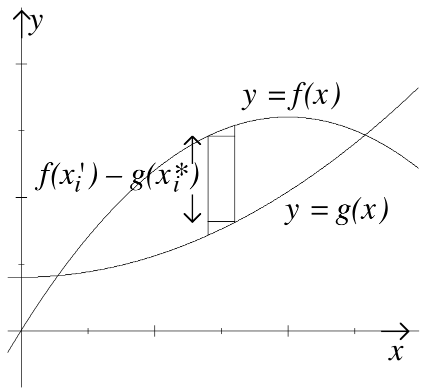
Sometimes we want to find, not the area between a curve and the x-axis, but the area enclosed between two curves, say between $y = f(x)$ and $y = g(x)$.
We can approach this problem in the same way as before by dividing the area up into strips and approximating the area of each strip by a rectangle. The lower sum is found by calculating the area of the interior rectangles as shown in the diagram.



The height of each interior rectangle is equal to the difference between the least value of $f(x)$, $f(x')$, and the greatest value of $g(x)$, $g(x^*)$, in the rectangle. The area of the ith rectangle is $(f(x_i') - g(x_i^*))\Delta x$.

$$\text{The lower sum} = \sum_{i=1}^{n}(f(x_i') - g(x_i^*))\Delta x.$$

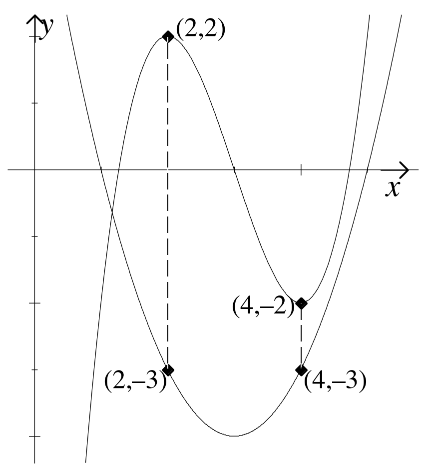
The upper sum can be found in the same way. The area enclosed between the curves is sandwiched between the lower sum and the upper sum.



When we pass to the limit as $\Delta x \to 0$, we get

$$\text{Area enclosed between the curves} = \int_a^b (f(x) - g(x))dx.$$

Note that the height is *always* $f(x) - g(x)$, even when one or both of the curves lie below the x-axis.
For example, if for some value of x, $f(x) = 2$ and $g(x) = -3$, the distance between the curves is $f(x) - g(x) = 2 - (-3) = 5$, or, if $f(x) = -2$ and $g(x) = -3$, the distance between the curves is $(-2) - (-3) = 1$ (see the diagram).



So, to find the area enclosed between two curves, we must:

1. Find where the curves intersect.

2. Find which is the upper curve in the region we are interested in.

3. Integrate the function (upper curve − lower curve) between the appropriate limits.

In other words, if two curves $f(x)$ and $g(x)$ intersect at $x = a$ and $x = b$, and $f(x) \geq g(x)$ for $a \leq x \leq b$, then

$$\text{Area enclosed between the curves} = \int_a^b (f(x) - g(x))dx.$$

Exercises 9

(Remember to draw a diagram first, before beginning any problem.)

1. Find the area enclosed between the parabola $y = x(x - 2)$ and the line $y = -x + 2$.

2. Find the area enclosed between the two parabolas $y = x^2 - 4x + 2$ and $y = 2 - x^2$.

3. Check that the curves $y = \sin x$ and $y = \cos x$ intersect at $\frac{\pi}{4}$ and $\frac{5\pi}{4}$, and find the area enclosed by the curves between these two point.

4. i Sketch the graphs of the function $y = 6 - x - x^2$ and $y = x^3 - 7x + 6$.

ii Find the points of intersection of the curves.

iii Find the total area enclosed between them.

10 Other Applications of the Definite Integral

The problem with which we introduced the idea of the definite integral was that of finding the area under a curve. As a result, most people tend to think of definite integrals always in terms of area. But it is important to remember that the definite integral is actually defined as the limit of a sum:

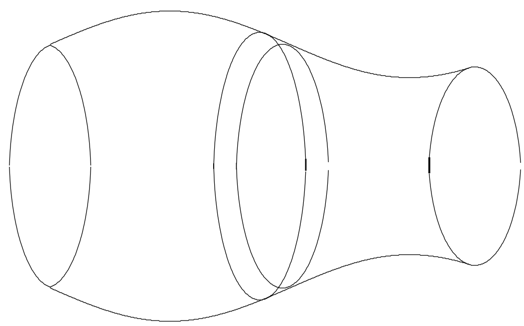
$$\lim_{n\to\infty}\sum_{i=1}^{n} f(x_i)\Delta x$$

and that any other problem which can be approximated by a similar sum will give rise to a definite intregral when we take the limit.

Examples

1. *Volume of a solid*

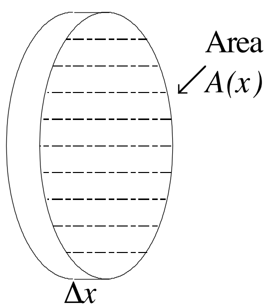
If we want to find the volume of a solid, we can imagine it being put through a bread slicer, and cut into slices of thickness Δx. If $A(x)$ is the cross sectional area at distance x along the x-axis, the volume of the slice will be approximately $A(x)\Delta x$, and the total volume of the solid will be approximately



$$\sum_{i=1}^{n} A(x_i)\Delta x.$$

When we pass to the limit as $\Delta x \to 0$ and $n \to \infty$, this becomes the definite integral

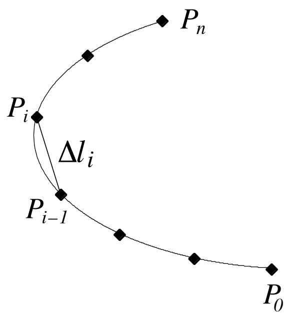
$$\int_a^b A(x)dx.$$



2. *Length of a curve*

We can approximate to the length of a curve by dividing it up into segments, as shown, and approximating the length of each segment by replacing the curved line with a straight line joining the end points. If the length of the ith straight line segment is Δl_i, the total length of the curve will be approximately

$$\sum_{i=1}^{n} \Delta l_i.$$



If we take the limit of this sum as the length of each segment approaches zero and the number of segments approaches infinity, we again get a definite integral. The details are rather complicated and are not given here.

3. *Mass of a body of varying density*

Suppose we have a bar, rope or chain whose linear density (mass per unit length) varies. Let the density at distance x along the x-axis be $d(x)$. If we subdivide the object into small sections of length Δx, the total mass can be approximated by the sum

$$\sum_{i=1}^{n} d(x_i)\Delta x.$$

When we take the limit as $n \to \infty$, we obtain the definite integral

$$\int_a^b d(x)dx.$$

4. *Work done by a variable force*

In mechanics, the work done by a constant force is defined to be the product of the magnitude of the force and the distance moved in the direction of the force. If the force $F(x)$ is varying, we can approximate to the work by dividing up the distance into small subintervals. If these are small enough, we can regard the force as effectively constant throughout each interval and so the work done in moving through distance Δx is approximately $F(x)\Delta x$.

The total work is thus approximately $\sum_{i=1}^{n} F(x_i)\Delta x$ and when we take the limit as $n \to \infty$, we find that the work done in moving the force from $x = a$ to $x = b$ is

$$\int_a^b F(x)dx.$$

Many other examples could be given, but these four should be sufficient to illustrate the wide variety of applications of the definite integral.

11 Solutions to Exercises

Exercises 5

1. **a.** $2(4^3) - 2(2^3) = 112$

b. $\frac{1}{9} - \frac{1}{1} = -\frac{8}{9}$

c. $\sqrt{16} - \sqrt{9} = 1$

d. $\ln 4 - \ln 2 = \ln \frac{4}{2} = \ln 2$

2. **a.** $\displaystyle\int_4^9 x^{-\frac{1}{2}} dx = \left[2x^{\frac{1}{2}}\right]_4^9 = 2\sqrt{9} - 2\sqrt{4} = 2$

b. $\displaystyle\int_{\frac{-\pi}{2}}^{\frac{\pi}{2}} \cos t dt = [\sin t]_{\frac{-\pi}{2}}^{\frac{\pi}{2}} = \sin\frac{\pi}{2} - \sin\left(-\frac{\pi}{2}\right) = 1 - (-1) = 2$

c. $\displaystyle\int_1^2 y^{-2} dy = \left[-y^{-1}\right]_1^2 = -\frac{1}{2} - \left(-\frac{1}{1}\right) = \frac{1}{2}$

d. $\displaystyle\int_{-2}^{-1} (s^2 + 2s + 2) ds = \left[\frac{1}{3}s^3 + s^2 + 2s\right]_{-2}^{-1} = \left(-\frac{1}{3} + 1 - 2\right) - \left(-\frac{8}{3} + 4 - 4\right) = 1\frac{1}{3}$

3. **a.** Area $\displaystyle = \int_1^2 (x^2 + 1) dx = \left[\frac{1}{3}x^3 + x\right]_1^2 = \left(\frac{8}{3} + 2\right) - \left(\frac{1}{3} + 1\right) = 3\frac{1}{3}$

b. Area $\displaystyle = \int_1^3 \frac{1}{x} dx = [\ln x]_1^3 = \ln 3 - \ln 1 = \ln 3$

c. Area $\displaystyle = \int_0^4 \sqrt{(4-u)} du = -\int_0^4 (4-u)^{\frac{1}{2}}(-1) du = -\left[\frac{2}{3}(4-u)^{\frac{3}{2}}\right]_0^4 = 5\frac{1}{3}$

d. Area $\displaystyle = \int_0^\pi 2\sin t dt = [-2\cos t]_0^\pi = -2\cos\pi + 2\cos 0 = 4$

4. **a.** $\displaystyle\int_0^1 xe^{x^2} dx = \frac{1}{2}\int_0^1 e^{x^2} \cdot 2x dx = \frac{1}{2}\left[e^{x^2}\right]_0^1 = \frac{1}{2}(e-1)$

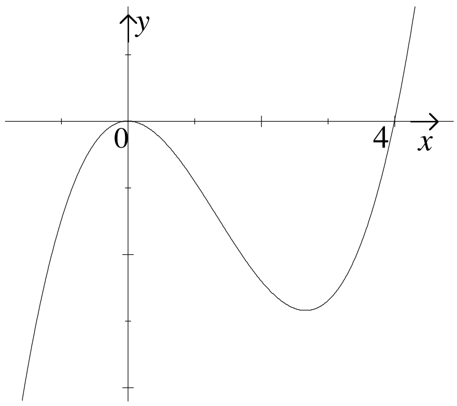
b. $\displaystyle\int_{-2}^{-1} \frac{1}{3-x} dx = -\int_{-2}^{-1} \frac{1}{3-x} \cdot (-1) dx = -[\ln(3-x)]_{-2}^{-1} = -(\ln 4 - \ln 5) = \ln 5 - \ln 4 = \ln\frac{5}{4}$

c. $\displaystyle\int_0^{\frac{\pi}{2}} \sin 2y dy = \frac{1}{2}\int_0^{\frac{\pi}{2}} \sin 2y \cdot 2dy = \frac{1}{2}[-\cos 2y]_0^{\frac{\pi}{2}} = \frac{1}{2}(-\cos\pi + \cos 0) = 1$

d. $\displaystyle\int_1^5 \frac{t}{4+t^2} dt = \frac{1}{2}\int_1^5 \frac{2t}{4+t^2} dt = \frac{1}{2}\left[\ln(4+t^2)\right]_1^5 = \frac{1}{2}\ln\frac{29}{5}$

Exercises 8

1. First, draw a graph.
The area is below the x-axis, so we first calculate $\int_0^4 3x^2(x-4)dx$.

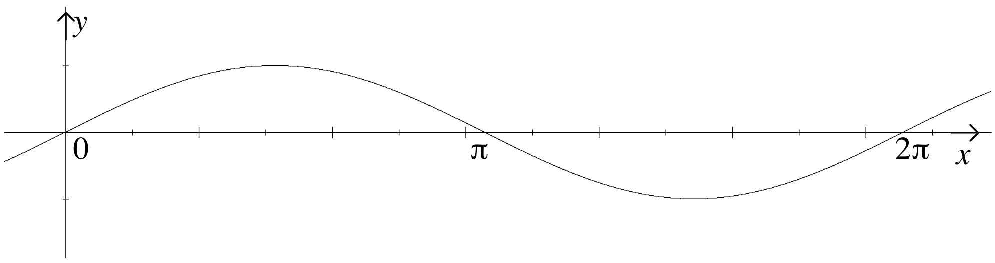


$$\begin{aligned}\int_0^4 3x^2(x-4)dx &= \int_0^4(3x^3-12x^2)dx\\ &= \left[\frac{3}{4}x^4-4x^3\right]_0^4\\ &= -64.\end{aligned}$$

The required area is therefore 64 units.

2. i $\displaystyle\int_0^{2\pi}\sin x dx = [-\cos x]_0^{2\pi} = -\cos 2\pi + \cos 0 = -1+1 = 0$

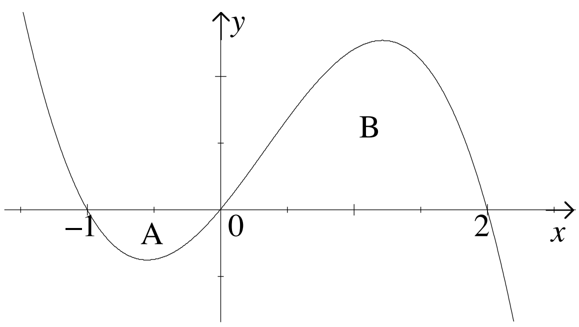
ii



From the graph we see that the area

$$\begin{aligned}\text{Area} &= \int_0^{\pi}\sin x dx + |\int_{\pi}^{2\pi}\sin x dx|\\ &= [-\cos x]_0^{\pi} + |[-\cos x]_{\pi}^{2\pi}|\\ &= (-\cos\pi+\cos 0) + |-\cos 2\pi + \cos\pi|\\ &= (-(-1)+1) + |-1+(-1)|\\ &= 4.\end{aligned}$$

3. The graph of the curve cuts the x-axis at -1, 0 and 2.
The total area = area A + area B.



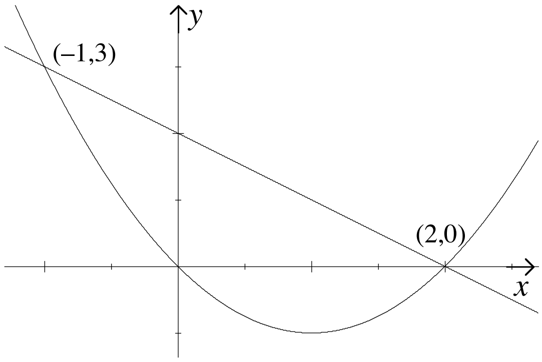
$$\begin{aligned}\text{Area A} &= |\int_{-1}^{0} 12x(x+1)(2-x)dx|\\ &= |\int_{-1}^{0}(-12x^3+12x^2+24x)dx|\\ &= |\left[-3^4+4x^3+12x\right]_{-1}^{0}|\\ &= |0-(-3-4+12)|\\ &= |-5| = 5.\end{aligned}$$

$$\begin{aligned}\text{Area B} &= \int_0^2 12x(x+1)(2-x)dx\\ &= \left[-3x^4+4x^3+12x^2\right]_0^2 = (-48+32+48) = 32.\end{aligned}$$

Therefore the total area is $5 + 32 = 37$ square units.

Exercises 9

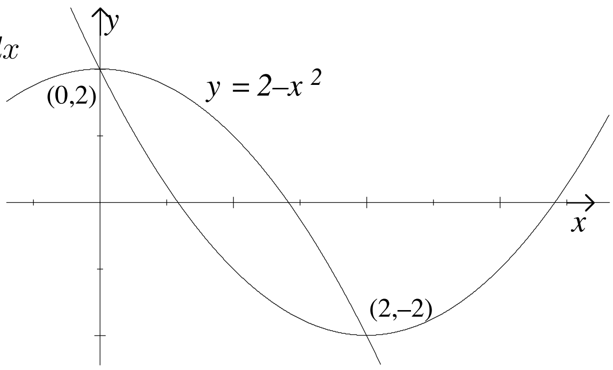
1. The curves $y = x^2 - 2x$ and $y = -x + 2$ intersect where $x^2 - 2x = -x + 2$. i.e. at $x = -1$ or $x = 2$.
The upper curve is $y = -x + 2$.



$$\begin{aligned}\text{Area} &= \int_{-1}^{2} ((-x+2) - (x^2 - 2x))dx \\ &= \int_{-1}^{2} (2 + x - x^2)dx \\ &= \left[2x + \frac{1}{2}x^2 - \frac{1}{3}x^3\right]_{-1}^{2} \\ &= (4 + 2 - \frac{8}{3}) - (-2 + \frac{1}{2} + \frac{1}{3}) \\ &= 4\frac{1}{2}.\end{aligned}$$

2. The curves intersect where $x^2 - 4x + 2 = 2 - x^2$ i.e. $2x^2 - 4x = 0$ i.e. $x = 0$ or $x = 2$.
The upper curve is $y = 2 - x^2$ (see sketch).

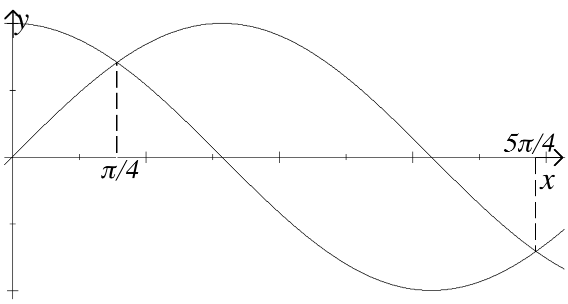
$$\begin{aligned}\text{Area} &= \int_{0}^{2} ((2 - x^2) - (x^2 - 4x + 2))dx \\ &= \int_{0}^{2} (4x - 2x^2)dx \\ &= \left[2x^2 + \frac{2}{3}x^3\right]_{0}^{2} \\ &= (8 - \frac{2}{3} \cdot 8) - 0 \\ &= 2\frac{2}{3}.\end{aligned}$$



3. When $x = \frac{\pi}{4}$, $\sin x = \frac{1}{\sqrt{2}}$ and $\cos x = \frac{1}{\sqrt{2}}$.

When $x = \frac{5\pi}{4}$, $\sin x = -\frac{1}{\sqrt{2}}$ and $\cos x = -\frac{1}{\sqrt{2}}$.

So the curves $y = \sin x$ and $y = \cos x$ intersect at $\frac{\pi}{4}$ and $\frac{5\pi}{4}$.



$$\begin{aligned}\text{Area} &= \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} (\sin x - \cos x)dx \\ &= [-\cos x - \sin x]_{\frac{\pi}{4}}^{\frac{5\pi}{4}}\end{aligned}$$

$$\begin{aligned}
&= (-\cos\frac{5\pi}{4} - \sin\frac{5\pi}{4}) + (\cos\frac{\pi}{4} + \sin\frac{\pi}{4}) \\
&= \frac{4}{\sqrt{2}} \\
&= 2\sqrt{2}.
\end{aligned}$$

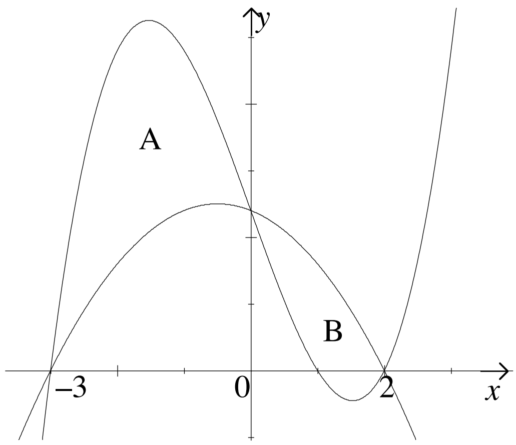
4. (i) and (ii) The curves are easier to sketch if we first find the points of intersection: they meet where $x^3 - 7x + 6 = 6 - x - x^2$.
That is,

$$x^3 + x^2 - 6x = 0$$

or

$$x(x-2)(x+3) = 0.$$

So the points of intersection are $(0, 6)$; $(2, 0)$; and $(-3, 0)$.
The first curve is an 'upside-down' parabola, and the second a cubic.
Total area = area A + area B.



$$\begin{aligned}
\text{Area A} &= \int_{-3}^{0} ((x^3 - 7x + 6) - (6 - x - x^2))dx \\
&= \int_{-3}^{0} (x^3 + x^2 - 6x)dx \\
&= \left[\frac{1}{4}x^4 + \frac{1}{3}x^3 - 3x^2\right]_{-3}^{0} \\
&= 15\frac{3}{4}.
\end{aligned}$$

$$\begin{aligned}
\text{Area B} &= \int_{0}^{2} ((6 - x - x^2) - (x^3 - 7x + 6))dx \\
&= \int_{0}^{2} (6x - x^2 - x^3)dx \\
&= 5\frac{1}{3}.
\end{aligned}$$

$\therefore$ the total area $= 15\frac{3}{4} + 5\frac{1}{3} = 21\frac{1}{12}$ square units.