

Integration as summation

The second major component of the Calculus is called *integration*. This may be introduced as a means of finding areas using summation and limits. We shall adopt this approach in the present Unit. In later units, we shall also see how integration may be related to differentiation.

In order to master the techniques explained here it is vital that you undertake plenty of practice exercises so that they become second nature.

After reading this text, and/or viewing the video tutorial on this topic, you should be able to:

- find, by using rectangles above and below the curve, the area between the graph of a quadratic function, the x -axis, and two vertical lines;
- use the integral notation for certain types of area.

Contents

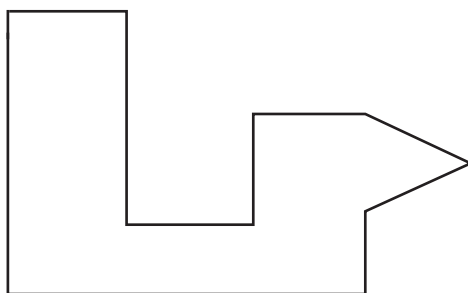
1. Introduction	2
2. Sub-dividing shapes	2
3. The area under the graph of a quadratic function	3
4. The area under a general graph	7

1. Introduction

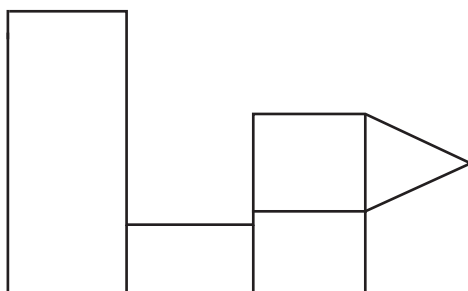
In this unit we are going to look at how we can find the area underneath the graph of a function. The technique we are going to use is called *integration*. The idea behind it is that we can find the area of a shape by dividing it into small shapes whose areas are easier to calculate. We then add up the areas of the small shapes. In most cases, we cannot make the division exactly, and so we estimate by using a slightly larger or slightly smaller area. As the small shapes get smaller and smaller, our estimates get better and better.

2. Sub-dividing shapes

Suppose we wish to find the area of a shape for which there is no exact formula. We can tackle such a problem by dividing up the shape into simpler pieces, where we do know a formula for the area of each piece. For example, we could find the area of this shape,



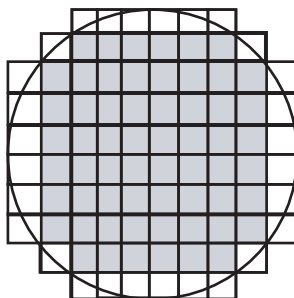
by dividing it up as shown like this:



The shape is now composed of rectangles and a triangle. There is a formula for the area of a rectangle, and another formula for the area of a triangle. So we can use these formulae to work out the individual areas of the small shapes, and hence the area of the whole shape.

Similarly, for a circle, there is a formula for the area, $A = \pi r^2$. But suppose we didn't know the formula for the area of a circle. What could we do?

One way might be to place a grid of squares over the circle. Then we could work out the area of each square, and add all these areas together. However, some of the squares will not coincide completely with the circle. Do we count such squares, or do we leave them out?

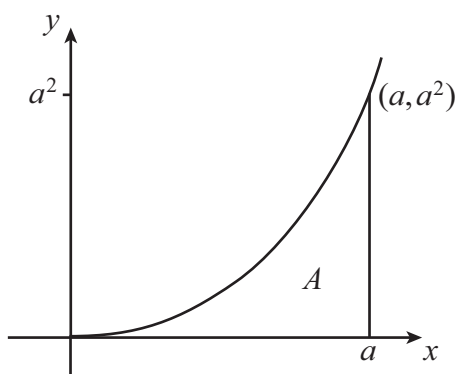


If we add together all the squares that contain any part of the circle, we will arrive at an over-estimate of the area of the circle. If we ignore all those squares that do not completely coincide with the circle, we will obtain an under-estimate of the area of the circle.

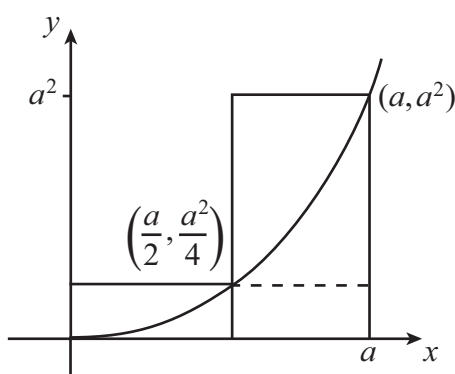
However, we now have two limits on the area of the circle, an upper one and a lower one. To ‘trap’ the area of the circle ever tighter between an upper limit and a lower limit, we would need to make the squares in the grid smaller and smaller. Then the parts that were included or discarded would be, in total, smaller and smaller.

3. The area under the graph of a quadratic function

We are now going to apply this process to an area, one edge of which is defined by a curve whose equation, $y = f(x)$, is known. We begin by looking at a specific case, the area contained by the x -axis from $x = 0$ to $x = a$, the curve $y = x^2$, and the line $x = a$. The area A we want to find is shown on the following diagram.



How can this area be calculated? A first attempt might be made by taking two rectangles, of equal width $a/2$, and heights corresponding to the y -values of the x co-ordinates that extend above the curve. So the height of the first rectangle will be $a^2/4$, and the height of the second rectangle will be a^2 . These two rectangles include more area than that contained between the curve and x -axis, and so we have an over-estimate of the area.



The area under the curve, A , is less than the total area of the two rectangles. So

$$A < \left(\frac{a}{2} \times \frac{a^2}{4}\right) + \left(\frac{a}{2} \times a^2\right).$$

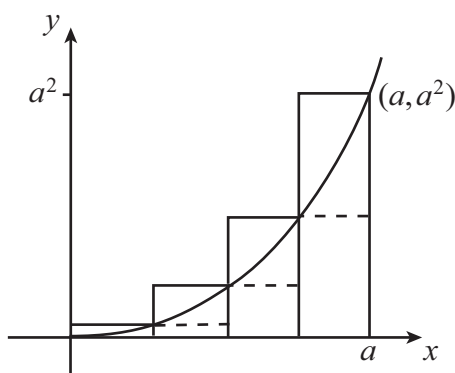
We now take a rectangle which contains only area between the curve and the x -axis. This is the rectangle whose upper boundary is marked by a dashed line, and comes below the curve. Its height is $a^2/4$. Its area provides an under-estimate of the area under the curve. Hence

$$\left(\frac{a}{2} \times \frac{a^2}{4}\right) < A.$$

We can combine these two inequalities into a single formula by writing

$$\left(\frac{a}{2} \times \frac{a^2}{4}\right) < A < \left(\frac{a}{2} \times \frac{a^2}{4}\right) + \left(\frac{a}{2} \times a^2\right).$$

In keeping with our ideas about the circle, let us now reduce the size of the grid. In this case, we reduce the width of the rectangles. So taking the areas of the rectangles which come ‘above’ and ‘below’ the curve we have the following diagram.



In each case, we can find the height of a rectangle by looking at the corner where the rectangle meets the curve, and taking the y co-ordinate. Let us first take the rectangles ‘above’ the curve. The width of each rectangle is $a/4$. The height of the first rectangle is $a^2/16$, the height of the second is $2^2a^2/16$, and so on. Adding together the areas of the four rectangles, we see that the area A satisfies the inequality

$$A < \left(\frac{a}{4} \times \frac{a^2}{16}\right) + \left(\frac{a}{4} \times \frac{2^2a^2}{16}\right) + \left(\frac{a}{4} \times \frac{3^2a^2}{16}\right) + \left(\frac{a}{4} \times \frac{4^2a^2}{16}\right),$$

and we can take out factors of $a/4$ and $a^2/16$ on the right-hand side to re-write this as

$$A < \frac{a}{4} \times \frac{a^2}{16} \times (1^2 + 2^2 + 3^2 + 4^2).$$

We can do the same for the rectangles below the curve, and this gives

$$\left(\frac{a}{4} \times \frac{a^2}{16}\right) + \left(\frac{a}{4} \times \frac{2^2a^2}{16}\right) + \left(\frac{a}{4} \times \frac{3^2a^2}{16}\right) < A.$$

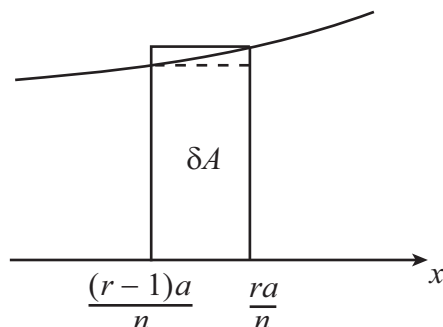
Once again we can take out factors of $a/4$ and $a^2/16$ on the right-hand side to give

$$\frac{a}{4} \times \frac{a^2}{16} \times (1^2 + 2^2 + 3^2) < A.$$

Again, we can combine these two inequalities to obtain

$$\frac{a}{4} \times \frac{a^2}{16} \times (1^2 + 2^2 + 3^2) < A < \frac{a}{4} \times \frac{a^2}{16} \times (1^2 + 2^2 + 3^2 + 4^2).$$

Now suppose we divide the x -axis between $x = 0$ and $x = a$ into n equal strips, each of width a/n , and look at the area of the r -th rectangle which comes ‘above’ the curve.



The height of the r -th rectangle ‘above’ the curve is the y co-ordinate of the top right-hand corner, and this is $(\frac{ra}{n})^2$. So the area of the r -th rectangle is

$$\frac{a}{n} \times \frac{r^2 a^2}{n^2},$$

and we can tidy this up to give

$$\frac{a^3}{n^3} r^2.$$

So if we write δA for the area below the curve in this little strip, we will have

$$\delta A < \frac{a^3}{n^3} r^2.$$

Adding up the areas of all these outer rectangles will give an over-estimate of the total area A under the curve, so that

$$A < \sum_{r=1}^n \frac{a^3}{n^3} r^2.$$

Now let us look at the r -th rectangle ‘below’ the curve. The height of the r -th rectangle is the y co-ordinate of the top left-hand corner, and this is $(\frac{(r-1)a}{n})^2$. So the area of the $(r-1)$ -th rectangle is

$$\frac{a}{n} \times \frac{(r-1)^2 a^2}{n^2}$$

and we can tidy this up to give

$$\frac{a^3}{n^3} (r-1)^2.$$

The area of this rectangle is less than δA , so that

$$\frac{a^3}{n^3} (r-1)^2 < \delta A.$$

Adding up the areas of all these inner rectangles will give an under-estimate of the total area A under the curve, so that

$$\sum_{r=1}^n \frac{a^3}{n^3} (r-1)^2 < A.$$

Thus the area A is contained between the two estimates,

$$\frac{a^3}{n^3} \sum_{r=1}^n (r-1)^2 < A < \frac{a^3}{n^3} \sum_{r=1}^n r^2.$$

Now there is a formula for the sum of the squares of the first n integers. The sum is given by the formula

$$\sum_{r=1}^n r^2 = 1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

We can use this formula directly for the sum on the right-hand side of the inequality. We can also use it for the sum on the left-hand side, because

$$\begin{aligned} \sum_{r=1}^n (r-1)^2 &= 0^2 + 1^2 + \dots + (n-1)^2 \\ &= 1^2 + \dots + (n-1)^2 \\ &= \sum_{r=1}^{n-1} r^2, \end{aligned}$$

so that we just have to replace n by $n-1$ in the formula. We get

$$\begin{aligned} \sum_{r=1}^n (r-1)^2 &= \frac{(n-1)[(n-1)+1][2(n-1)+1]}{6} \\ &= \frac{(n-1)n(2n-1)}{6}. \end{aligned}$$

Substituting this into the inequality above gives

$$\frac{a^3}{n^3} \times \frac{(n-1)n(2n-1)}{6} < A < \frac{a^3}{n^3} \times \frac{n(n+1)(2n+1)}{6}.$$

On each side, we can cancel an n and multiply out the two remaining brackets to give

$$\frac{a^3}{6} \times \frac{1}{n^2} \times (2n^2 - 3n + 1) < A < \frac{a^3}{6} \times \frac{1}{n^2} \times (2n^2 + 3n + 1).$$

We can also divide each bracket through by the remaining n^2 in the denominator, to give

$$\frac{a^3}{6} \times \left(2 - \frac{3}{n} + \frac{1}{n^2}\right) < A < \frac{a^3}{6} \times \left(2 + \frac{3}{n} + \frac{1}{n^2}\right).$$

Now as we increase the number of strips, their width becomes smaller and smaller, and our estimate of the area becomes more and more accurate. If we take this process to the ‘limit’ and let n tend to infinity, then $3/n$ and $1/n^2$ tend to zero. Thus the area under the curve A is trapped between $2a^3/6$ and $2a^3/6$, so that

$$A = \frac{a^3}{3}.$$

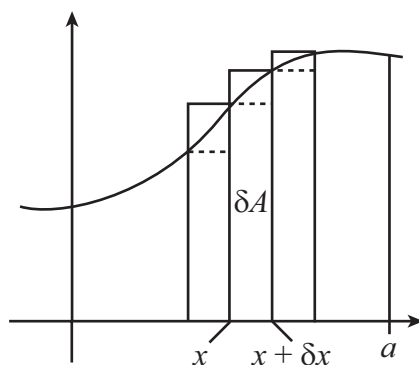


Key Point

To find the area, we divide up the area under the curve into two series of rectangles, all with the same width. One set of rectangles contains elements of area above the curve and provides an over-estimate of the area. The other set of rectangles is completely within the area under the curve and omits some of that area, thus providing an under-estimate of the area. We then let the width of these rectangles tend to zero by letting n tend to infinity, trapping the value for the area under the curve between two equal limits.

4. The area under a general graph

We have just seen how to find the area under a very particular curve, namely the graph of $y = x^2$. We are now going to extend this process to cover all curves which have similar properties to $y = x^2$. One particular property is that we shall want the curve to lie above the x -axis. We shall then have a general process for finding the area contained by a curve, the x -axis, and specified limits.



We shall look at the area contained by the curve, the x -axis, the y -axis, and the ordinate $x = a$. We shall divide this area into thin strips, each of the same width δx . Let us look at a typical strip, with area δA .

In this strip, there is a point P on the curve with co-ordinates (x, y) . The y co-ordinate of P gives the height of the inner rectangle. As we increase x by a small positive amount δx , we move along the curve to the point Q , with co-ordinates $(x + \delta x, y + \delta y)$. Here, δy is the change in y resulting from the small change in x . The y co-ordinate of Q gives the height of the outer rectangle.

Now the area δA is caught between two rectangles. Both rectangles have width δx . So

$$y \delta x < \delta A < (y + \delta y) \delta x .$$

Now the total area A between the curve and the x -axis from $x = 0$ to $x = a$ is given by the sum of all the little areas δA ,

$$A = \sum_{x=0}^{x=a} \delta A ,$$

and hence

$$\sum_{x=0}^{x=a} y \delta x < A < \sum_{x=0}^{x=a} (y + \delta y) \delta x ,$$

or

$$\sum_{x=0}^{x=a} y \delta x < A < \sum_{x=0}^{x=a} (y \delta x + \delta y \delta x) .$$

We now want to let δx tend to zero, so that the vertical strips get thinner and thinner. As the change in x tends to zero, the change in y , δy , also tends to zero. And the product of two small quantities, $\delta y \delta x$, is very small indeed. In the limit, the area A is given by

$$A = \lim_{\delta x \rightarrow 0} \sum_{x=0}^{x=a} y \delta x .$$

(There are various assumptions which have been made in arriving at this result. One is that the graph of the function, as we have drawn it, is increasing. This means that height of each inner rectangle is given by the y -value at its left hand edge, and the height of each outer rectangle is given by the y -value at its right-hand edge. But the same result is valid for functions which are decreasing, or which oscillate up and down. There are some technical mathematical arguments, including some conditions on the types of function which can be considered, that ensure the validity of the formula we have given.)

We need a notation for the right-hand side of this expression. The notation we use is

$$A = \lim_{\delta x \rightarrow 0} \sum_{x=0}^{x=a} y \delta x = \int_0^a y \, dx .$$

We say that this is the ‘the integral of y with respect to x , between the limits 0 and a ’.



Key Point

The area A between the graph of a positive function, the x -axis, the y -axis and the vertical line $x = a$ is given by

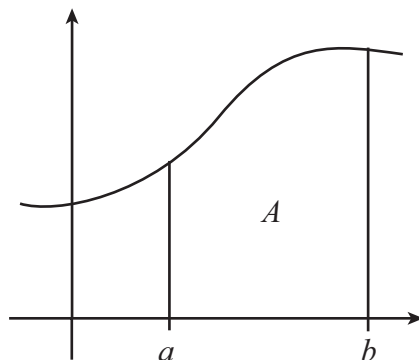
$$A = \lim_{\delta x \rightarrow 0} \sum_{x=0}^{x=a} y \delta x .$$

This expression is also written as

$$A = \int_0^a y \, dx$$

and is called the integral of y with respect to x , between the limits 0 and a .

We can also find the area between a curve, the x -axis, and two vertical lines $x = a$ and $x = b$, where $a < b$.



Now we can think of this area as the difference between two areas, so that

$$A = \int_0^b y \, dx - \int_0^a y \, dx .$$

We normally write this as

$$\int_a^b y \, dx .$$



Key Point

The area A between the graph of a positive function, the x -axis and the vertical lines $x = a$ and $x = b$ is given by

$$A = \int_0^b y \, dx - \int_0^a y \, dx .$$

This expression is also written as

$$\int_a^b y \, dx$$

and is called the integral of y with respect to x , between the limits a and b .

One final property of integrals which we can derive tells us about the case where we have three vertical lines $x = a$, $x = b$ and $x = c$, with $a < b < c$. We can find a relationship between the three integrals

$$\int_a^b y \, dx , \quad \int_a^c y \, dx , \quad \int_b^c y \, dx .$$

We know that

$$\int_a^c y \, dx = \int_a^b y \, dx + \int_b^c y \, dx .$$

We can add and subtract the same quantity on the right-hand side, so we have

$$\int_a^c y \, dx = \int_0^c y \, dx - \int_0^b y \, dx + \int_0^b y \, dx - \int_0^a y \, dx .$$

Now we can re-write the right-hand side as

$$\int_a^c y \, dx = \int_b^c y \, dx + \int_a^b y \, dx ,$$

writing the integral between the limits a and c as the sum of two integrals, the first between the limits b and c , and the second between the limits a and b . In terms of areas, this is just what we would expect.

Exercises

1. Calculate the area contained by the ordinates $x = 0$, $x = 2$, the x -axis and the curve defined by the equation $y = x$ by dividing the area into n rectangular strips that (a) form an upper limit to the area, and (b) form a lower limit, as in the example of the area contained by the curve $y = x^2$, the x -axis and the ordinates $x = 0$ and $x = a$. Check your answer by calculating the area of the triangle in the usual way.
2. Calculate the area contained by the ordinates $x = 0$, $x = a$, the x -axis and the curve defined by the equation $y = e^x$ by dividing the area into n rectangular strips and (a) form an upper limit to the area and (b) form a lower limit. Show that this area is $e^a - 1$.