Module M5.1 Introducing integration

1 Opening items

1.1 Module introduction

The notion of a *limit of a sum* arises in many applications of mathematics to physics. For example, suppose we want to find the total mass of a cylindrical column of air, the density of which decreases with height in a known way. One approach is to divide the air column into a number of horizontal discs — like a stack of coins — each sufficiently thin that we may regard it as having constant density. We can then work out the mass of each disc by multiplying its volume by its (constant) density, and add all those disc masses together to find an approximate value for the mass of the entire column. Our final answer will only be an approximation to the correct value because our assumption that each disc has constant density is only approximately true. Nonetheless, we may obtain a good estimate in this way and we may improve its accuracy by increasing the number of discs in the sum. Indeed, if we consider what happens as the thickness of the discs approaches zero (and the number of discs approaches infinity) then we see that the limit of the sum of disc masses is the exact value of the air column's total mass. This is typical of the way in which the limit of a sum may arise in a physical problem. Limits of sums of this kind are referred to by physicists and mathematicians as *definite integrals*. The *formulation* of physical problems in terms of definite integrals is a major theme of this module.

Another major theme is the *evaluation* of definite integrals. Sometimes the only way of evaluating the limit of a sum is to perform the sum for a very large number of terms and then try to work out what would happen 'in the limit' as the number of terms approached infinity. This 'head-on' approach called *numerical integration* can provide accurate answers, but it can also be time consuming and very hard work. Fortunately there are many situations in which the evaluation of definite integrals can be performed more easily thanks to an important mathematical result known as the *fundamental theorem of calculus*. According to the fundamental theorem there is a remarkable link between the evaluation of the limit of a sum and the process of differentiation. In fact, many definite integrals may be evaluated by a procedure in which the key step amounts to little more than the reversal of the usual process of differentiation. This process of *inverse differentiation,* or *indefinite integration* as it is more commonly called, is a vital part of the mathematical tool kit of every physicist and engineer.

In Section 2 of this module we introduce the basic concepts of definite and indefinite integration. We start by recalling the definition of a derivative and then go on to define the *inverse derivative* that emerges from the process of inverse differentiation. Using the physical concepts of position, velocity and acceleration we show how inverse derivatives may be used to evaluate quantities such as the distance travelled in a given time by an object moving with varying velocity. This discussion leads us to formulate the concept of a definite integral defined as the limit of a sum. We then state (and prove) the fundamental theorem of calculus which provides the formal link between the evaluation of limits of sums and inverse differentiation. Finally, we round off the section by explaining why it makes sense to refer to inverse derivatives as indefinite integrals.

Section 3 is concerned with the formulation of problems in terms of definite integrals and their evaluation using indefinite integrals. Section 4 attempts to put the topic into its scientific context and provides some exercises for you to do.

1.2 Fast track questions

Study comment   Can you answer the following *Fast track questions*? If you answer the questions successfully you need only glance through the module before looking at the *Module summary* (Subsection 5.1) and the *Achievements* listed in Subsection 5.2. If you are sure that you can meet each of these achievements, try the *Exit test* in Subsection 5.3. If you have difficulty with only one or two of the questions you should follow the guidance given in the answers and read the relevant parts of the module. *However, if you have difficulty with more than two of the Exit questions, you are strongly advised to study the whole module.*

Question F1

Evaluate the following indefinite integrals:

(a) $\int (1 + 2x + 3x^2)\,dx$ (b) $\int [\sin(2x) + \cos(3x)]\,dx$

(c) $\int \left(e^t + \dfrac{1}{t^2} \right) dt$ (d) $\int dw$

Question F2

Evaluate the following definite integrals:

(a) $\int_{-1/4}^{1/4} \cos(2\pi x)\,dx$

(b) $\int_{0}^{3} (2t - 1)^2\,dt$

(c) $\int_{1}^{2} \frac{(1 + e^t)^2}{e^t}\,dt$

(d) $\int_{4}^{9} \sqrt{x}\left(x - \frac{1}{x}\right)dx$

Question F3

An object starts from rest at time $t = 0$, and moves along the x-axis with acceleration $a_x(t)$ given, at time t, by

$a_x(t) = a - bt$ where $a = 5 \, \text{m s}^{-2}$ and $b = 3 \, \text{m s}^{-3}$

State whether each of the following statements is true or false and explain why.

(a) The object comes to rest when $t = (5/3)$ s

(b) The displacement of the object from its initial position is zero when $t = 5$ s.

(c) The magnitude of the area bounded by the velocity–time graph and the t-axis, from $t = 0$ to $t = 1$ s, represents the distance the object travels in the first second.

(d) The magnitude of the area bounded by the graph of $y = a_x(t)$ and the t-axis between $t = 0$ s and $t = 5/3$ s represents the maximum speed of the object in the first ten seconds.

Study comment

Having seen the *Fast track questions* you may feel that it would be wiser to follow the normal route through the module and to proceed directly to *Ready to study?* in Subsection 1.3.

Alternatively, you may still be sufficiently comfortable with the material covered by the module to proceed directly to *Closing items*.

1.3 Ready to study?

Study comment   In order to study this module you will need to be familiar with the following terms: *Cartesian coordinate system*, *derivative* (or *derived function*), *function*, *graph*, *limit*, *magnitude* (of a vector, as in *magnitude of a force*), *summation symbol* (i.e. Σ). You should be able to *differentiate* a range of functions, and be able to *expand*, *simplify* and *evaluate* expressions, including those that involve *exponential* and *trigonometric functions.* If you are uncertain of any of these terms, you can review them by referring to the *Glossary* which will indicate where in *FLAP* they are developed.

In particular you will need to know the following derivatives:

$$\frac{d}{dt}(t^n) = nt^{n-1}, \ \frac{d}{dt}[\sin(at)] = a\cos at, \ \frac{d}{dt}[\cos(at)] = -a\sin at, \ \frac{d}{dt}(e^{at}) = ae^{at}$$

(where *n* and *a* are constants); and to remember that a *function* $f(x)$ is often given by a formula such as $y = f(x)$, and is a rule that assigns a unique value of *y* to each value of *x* (see the note below). The following *Ready to study questions* will allow you to establish whether you need to review some of the topics before embarking on this module. Throughout this module $\sqrt{x}$ will represent the *positive* square root of *x*.

A special note on functions   The function

$$f(x) = 1 + x + x^2$$

is an example of a function that is defined for all values of x. One may think of the function as a sort of machine with x as the input and $1 + x + x^2$ as the output. The input x is known as the *independent variable* and, if we write $y = f(x)$, the output y is known as the *dependent variable*, since the function $f(x)$ determines the way in which y depends on x.

The same function f could equally well be defined using some other symbol, such as t, to represent the independent variable:

$$f(t) = 1 + t + t^2$$

This freedom to relabel the independent variable is often of great use, though it is vital that such changes are made consistently throughout an equation.

We may *evaluate* this function, whether we call it $f(x)$ or $f(t)$, for any value of the independent variable; for example, if we choose to use x to denote the independent variable, and set $x = 1$, we have

$$f(1) = 1 + 1 + 1^2 = 3$$

Similarly, if $x = \pi$ $f(\pi) = 1 + \pi + \pi^2$

and, if $x = 2a$ $f(2a) = 1 + 2a + (2a)^2 = 1 + 2a + 4a^2$

When we write expressions such as $f(\pi)$ or $f(2a)$ whatever appears within the brackets is called the *argument* of the function. The value of $f(x)$ is determined by the *value* of its argument, irrespective of what we call the argument.

Question R1

For each of the following functions $f(t)$ find the <u>derived function</u> or <u>derivative</u>, $f'(t) = \dfrac{df}{dt}$:

(a) $f(t) = t^{2/3}$,   (b) $f(t) = \sin\left(\dfrac{\pi t}{2}\right)$,   (c) $f(t) = 5\mathrm{e}^{at}$.

Question R2

Expand the following expressions:

(a) $\dfrac{(1 + \mathrm{e}^x)^2}{\mathrm{e}^x}$ and   (b) $\sqrt{x}\left(x - \dfrac{1}{x}\right)$.

Question R3

If $F(x) = x\left[\sin\left(\dfrac{\pi x}{2}\right) - 1\right]$, evaluate the following:

(a) $F(-1) - F(-2)$,   (b) $F(1) - F(-1)$,   (c) $F(1) - F(-2)$.

Question R4

(a) Evaluate $\displaystyle\sum_{i=1}^{4} i^2$ where i stands for a real integer.

(b) Given $t_i = 1 + (i/4)$, write down the values of t_1, t_2, t_3 and t_4 and then evaluate $\displaystyle\sum_{i=1}^{3}(t_{i+1} - t_i)$.

(c) If we are also told that $v(t) = t^2$, calculate the value of $\displaystyle\sum_{i=1}^{3} v(t_i)(t_{i+1} - t_i)$.

2 The concept of integration

2.1 Inverse derivatives: reversing differentiation

From a mathematical point of view, *differentiation* is a rather straightforward procedure. Given a function $f(x)$, its *derivative* is another function that may be denoted by $f'(x)$ or $\dfrac{df}{dx}$ and which is defined by

$$\frac{df}{dx} = \lim_{h \to 0}\left(\frac{f(x+h) - f(x)}{h} \right) \qquad (1)$$

for all values of x at which a unique limit exists. This simple rule makes it possible to find derivatives for an enormous range of functions. ☞

The importance of differentiation in the study of physics would be hard to overestimate. Not only is it used in the analysis and solution of a great many problems, it even plays a fundamental role in the definition of many basic concepts. Nowhere is this more obvious than in **kinematics**, the study of motion.

For example, consider an object moving in a straight line along the *x*-axis of a *Cartesian coordinate system*. Such an object is said to be undergoing **linear motion**, and its **position** at any time *t* is determined by the single position coordinate *x*. We can represent the position by $x(t)$ to emphasize that it is a *function* of *t*, since its value changes with time as the object moves along the line. Now, for any moving object the two important physical quantities that help to characterize the motion are **velocity** and **acceleration**, and these may both be defined as derivatives. In the case of linear motion along the *x*-axis we have:

velocity $\quad v_x = \dfrac{dx}{dt}$ ☞

acceleration $\quad a_x = \dfrac{dv_x}{dt}$

Thus, given the position of a moving object as a function of time, i.e. given the explicit form of the function $x(t)$, it is usually fairly easy to determine the corresponding velocity $v_x(t)$ and acceleration $a_x(t)$, and hence obtain some insight into the nature of the motion.

◆ The position of a certain object undergoing linear motion is specified by $x(t) = at^3 + bt + c$, where *a*, *b* and *c* are constants. What are the functions $v_x(t)$ and $a_x(t)$ that describe the velocity and acceleration of the object? By considering the values of the functions $x(t)$ and $v_x(t)$ when $t = 0$, explain the physical significance of constants *b* and *c*.

It is comforting to know that it is relatively easy to determine the velocity and acceleration that correspond to a given form for $x(t)$, but unfortunately this is not a problem that often confronts a physicist. More common is the problem that arises if we turn the previous discussion on its head and ask how we might determine the position of an object if we are told its velocity, or perhaps its acceleration, as a function of time? The latter problem is particularly common since, according to **Newton's second law** of motion, the acceleration of a body of fixed **mass** is proportional to the total **force** acting on that body, and we often start the analysis of a problem knowing only the forces that are involved. ☞ Clearly, in order to deal with such problems and determine $x(t)$ from $v_x(t)$ or $a_x(t)$ we need to reverse the process of differentiation. This reverse process is known as **inverse differentiation** and is generally much harder than ordinary differentiation. ☞

As an example of inverse differentiation, and its pitfalls, let us try inverting the problem we considered above. Specifically, let us suppose we are told that the velocity of a particular object moving in one dimension is given by $v_x(t) = 3at^2 + b$ and that we want to find its position as a function of time. How should we do this?

In view of the earlier discussion, your first thought might be the right one that inverse differentiation would give $x(t) = at^3 + bt + c$. However, you will not usually be in possession of such 'privileged' information. Normally the best that you could do would be to try to think of the most general function with a derivative equal to the given form of $v_x(t)$. Looking at the form of $v_x(t)$ it is pretty clear that its inverse derivative must include the expression $at^3 + bt$, but there is no evidence whatsoever that a constant c should be added to this.

However, since the derivative of *any* constant is zero, you might well say that it is quite *possible* that there is an additional constant since we may add *any* constant to $at^3 + bt$ and the resulting expression will still have the property that its derivative is identical to $v_x(t)$. Thus, if we only know that

$$v_x(t) = \frac{dx}{dt} = 3at^2 + b \qquad (2)$$

then the most we can say about $x(t)$ is that it is of the general form

$$x(t) = at^3 + bt + \text{arbitrary constant} \qquad (3)$$

The point here is that a knowledge of the velocity alone is not sufficient to completely determine the position of the object as a function of time because we do not know its initial position on the *x*-axis (i.e. the value of the constant). Of course, if we are given some extra information, such as the location of the object at $t = 0$, then we may be able to determine the value of the arbitrary constant, but, in the absence of such additional information, Equation 3 is the most complete answer we can find.

As you can see, in the case of inverse differentiation there is no simple rule to apply, no reliable formula to use, all you can do is to 'inspect' the given function and to use your knowledge of differentiation to 'suggest' the form of the inverse derivative. You can check your answer by differentiating it to make sure you recover the function you started with, but you must always remember to add an arbitrary constant to your answer since the derivative of any constant is zero.

We can sum up the process of inverse differentiation in the following way.

Given a function $f(x)$, its **<u>inverse derivative</u>** is any function $F(x)$ such that

$$\frac{dF}{dx} = f(x)$$

An *inverse derivative* of a given function is also known as a **primitive** of that function, though for reasons that will become clear later, physicists usually refer to them as *indefinite integrals*. Note that a function may (and generally does) have infinitely many inverse derivatives, corresponding to the infinitely many possible choices for the value of the arbitrary constant. It therefore makes sense to speak of *an* inverse derivative rather than *the* inverse derivative of a given function.

❖   Write down three different inverse derivatives of $f(x) = 2x$.

In view of the many possible answers to the question 'what is an inverse derivative of $f(x) = 2x$?' it is customary to present the answer in the 'general' form $F(x) = x^2 + C$, where C is an arbitrary constant. This leads to the following observation:

If $F_1(x)$ and $F_2(x)$ are both inverse derivatives of the same function $f(x)$, then there exists a constant K such that

$\quad F_1(x) = F_2(x) + K$ ☞

If you feel that you have grasped the principle of inverse differentiation, try the following question. If you're not so confident, treat the first two parts as a worked example and then try to answer the remaining part.

❖ Write down an inverse derivative $F(x)$ for each of the following functions:

(a) $f(x) = x^2$, (b) $f(x) = 3x^2 - 2x$, (c) $f(x) = ax/2$, where a is a constant.

❖ Write down an inverse derivative $F(t)$ for each of the following functions:

(a) $f(t) = t^{-3}$ (b) $f(t) = -3t^{-1.5} + 5t^{1.5}$

✦ At time t, an object undergoing linear motion has acceleration $a_x(t) = 4At^3$, where $A = 1\,\mathrm{m\,s^{-5}}$. Given that the object is at rest at $t = 0$, find the velocity of the object at $t = 3\,\mathrm{s}$.

Motion with uniform acceleration under gravity

To stress the importance of inverse differentiation (i.e. indefinite integration) let us use it to derive some well-known equations — those that describe an object moving vertically near the surface of the Earth, with constant <u>acceleration due to gravity</u>. In this case, if we let the x-axis point vertically downwards the object will experience a constant acceleration given by

$$a_x(t) = g \qquad \qquad (4)$$

where g is the <u>magnitude of the acceleration due to gravity</u>, which has a value of about $9.81\,\mathrm{m\,s^{-2}}$. ☞

At $t = 0$, when the object has position coordinate $x(0)$, its initial velocity is $v_x(0) = u_x$, where u_x is a constant. Applying inverse differentiation to Equation 4 we see that $v_x(t) = gt + C$ for some constant C. It follows that $v_x(0) = C$, but we know that $v_x(0) = u_x$, so $C = u_x$, and therefore

$$v_x(t) = u_x + gt \qquad \qquad (5)$$

We have now found the velocity, but we can apply the same argument again to find how the position varies with time. This time we require a function of t that when differentiated produces $u_x + gt$. It is not difficult to see that the appropriate function is

$$x(t) = u_x t + \frac{gt^2}{2} + D$$

for some constant D. Now since the initial position of the object is $x(0)$ we see that $D = x(0)$ and hence

$$x(t) = u_x t + \frac{gt^2}{2} + x(0) \tag{6}$$

In fact, when dealing with problems of this kind we usually want to know how far the object is from its initial position and in which direction (i.e. upwards or downwards). This information is given by the **displacement** of the object from its initial position, ☞ which in this case is defined by

$$s_x(t) = x(t) - x(0) \tag{7}$$

Using this definition to eliminate $x(t)$ and $x(0)$ from Equation 6 we have

$$s_x(t) = u_x t + \frac{gt^2}{2} \tag{8}$$

and this equation gives us the displacement of the object at any time t.

The **distance** $s(t)$ of the object from its initial position, at time t, is a positive quantity given by the magnitude
☞ of its displacement, so $s(t) = |s_x(t)|$, and it too can be found from Equation 8.

$$s_x(t) = u_x t + \frac{gt^2}{2} \qquad \text{(Eqn 8)}$$

For the sake of completeness we note that Equation 5

$$v_x(t) = u_x + gt \qquad \text{(Eqn 5)}$$

may be rearranged to give $t = (v_x - u_x)/g$ and that upon substituting this into Equation 8 and rearranging we
obtain

$$v_x^2 = u_x^2 + 2gs_x \qquad (9)$$

Equations 5, 8 and 9 are a particular case of the *uniform acceleration equations*, which are derived elsewhere in
FLAP without using inverse differentiation, but at the price of much greater labour. The simplicity of the above
derivation of these fundamentally important kinematic relationships is ample proof of the power of inverse
differentiation.

Motion with non-uniform acceleration

It is important to appreciate that the above method can be applied to any motion in which the acceleration is a known function of time — the acceleration does not have to be uniform (i.e. constant). We have already seen some examples of this kind, but to emphasize the point, and to provide a further illustration, we now consider **simple harmonic motion**, ☞ a form of linear motion in which the acceleration of an object is given by:

$$a_x(t) = -A\omega^2 \sin(\omega t) \qquad\qquad (10) \quad ☞$$

where A and ω are positive constants, known respectively as the **amplitude** and **angular frequency** of the motion. Suitable SI units for these quantities would be m for the amplitude and s^{-1} for the angular frequency, so the combination $A\omega^2$ can be expressed in units of m s^{-2}, the same units as acceleration.

Now, the general properties of the sine function are such that the term $\sin(\omega t)$ in Equation 10 varies repeatedly between +1 and −1 as t increases, so both the magnitude and the direction of the acceleration fluctuate with time. As you might expect, these regular changes of sign result in back and forth (oscillatory) motion; a fact we shall now demonstrate using inverse differentiation.

The first step is to find the velocity $v_x(t)$, an inverse derivative of $a_x(t)$. In this case it must involve a function whose derivative is $-A\omega^2 \sin(\omega t)$. The only obvious choices involve $\cos(\omega t)$, the derivative of which is $-\omega \sin(\omega t)$.

Bearing this in mind, and remembering the need to include an arbitrary additive constant C in the answer, it is not too difficult to see that the required inverse derivative has the general form

$$v_x(t) = A\omega \cos(\omega t) + C \qquad (11) \quad ☞$$

In order to find the position as a function of time, $x(t)$, we must find the appropriate inverse derivative of $v_x(t)$. This is one of those cases where 'we' means 'you'.

Question T1

Write down the inverse derivative of $v_x(t)$ in this case, and then use the conditions $x = 0$ when $t = 0$, and $v_x = 0$ when $t = \pi/(2\omega)$ to show that

$$x(t) = A \sin(\omega t) \qquad (12) \quad ❑$$

As t increases and the function $\sin(\omega t)$ varies smoothly between $+1$ and -1, the value of $x(t)$ varies smoothly between $+A$ and $-A$. Moreover due to the periodic (repeating) nature of the sine function, the motion is repeated every time t increases by $2\pi/\omega$. Thus Equation 12 does indeed describe oscillations, as promised, and the quantity $T = 2\pi/\omega$ is known as the **period** of the motion. From this you can see that the conditions $x = 0$ when $t = 0$, and $v_x = 0$ when $t = \pi/(2\omega)$ given in Question T1 are equivalent to saying that the oscillating object passes through the origin at $t = 0$, and is momentarily at rest a quarter of a period later (at $t = T/4 = \pi/(2\omega)$) when its displacement from the origin attains its maximum value, A. ☞

Question T2

Write down inverse derivatives of x^2, x^3 and x^4. Write down a rule for determining an inverse derivative of x^p, where the power p may have any real value except -1. ❏

Question T3

Write down an inverse derivative of (a) e^{-x}, and (b) $x + e^x$. ❏

2.2 Inverse derivatives and the area under a graph

We now extend our study of inverse derivatives by exposing an important relationship between an inverse derivative of a function and the area under the graph of that function. The relationship is a general one, but we will introduce it in the physical context of kinematics.

Consider an object moving along a straight line with velocity $v_x(t)$. Provided we know the explicit form of the function $v_x(t)$ it is a fairly straightforward matter to plot the graph of v_x against t. Such a graph is called the **velocity–time graph** of the motion.

Figure 1a shows a particularly simple velocity–time graph, that of an object released from rest at time $t = 0$ and falling vertically downwards with constant acceleration under the influence of gravity. As we saw in the last subsection, if we take vertically downwards to be the positive x-direction then the velocity of such an object, at time t, is

$$v_x = gt \qquad (13)$$

where the constant g is the magnitude of the acceleration due to gravity. This simple relationship is reflected in the velocity–time graph where the value of v_x increases in proportion to t, i.e. it is a straight-line graph.
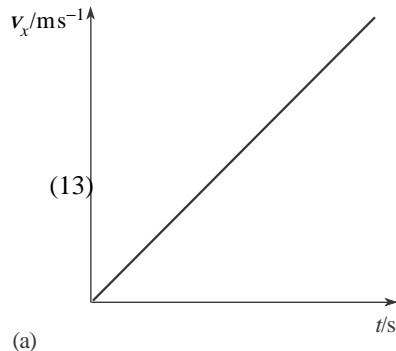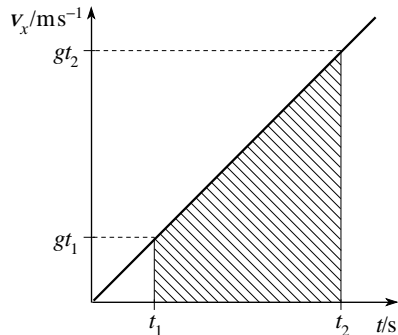

(a)

Figure 1 (a) A simple velocity–time graph.

If we choose two particular times, t_1 and t_2 say, where $t_2 > t_1$, then we can see from Equation 13

$$v_x = gt \qquad \text{(Eqn 13)}$$

(or from Figure 1b) that the corresponding values of the velocity are $v_x(t_1) = gt_1$ and $v_x(t_2) = gt_2$.

(b)

Figure 1 (b) The area under the graph between $t = t_1$ and $t = t_2$.

It follows that the average velocity of the falling object over the period from t_1 to t_2 is

$$\frac{1}{2}[v_x(t_1) + v_x(t_2)] = \frac{1}{2}(gt_1 + gt_2)$$

and that the distance travelled during that time is
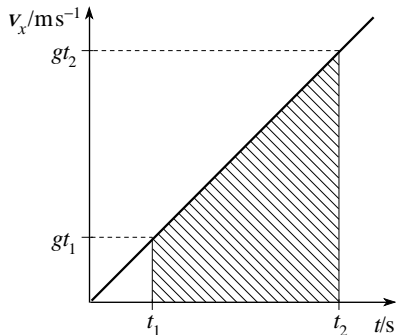
$$\frac{1}{2}(gt_1 + gt_2)(t_2 - t_1)$$

Now, this distance represents the change in the position coordinate of the object over the interval $t_2 - t_1$ so we can write

$$x(t_2) - x(t_1) = \frac{1}{2}(gt_1 + gt_2)(t_2 - t_1) \qquad (14)$$

However, the right-hand side of this equation

$$x(t_2) - x(t_1) = \frac{1}{2}(gt_1 + gt_2)(t_2 - t_1) \qquad \text{(Eqn 14)}$$

can be interpreted graphically in terms of Figure 1b as the shaded area under the velocity–time graph between $t = t_1$ and $t = t_2$. ☞
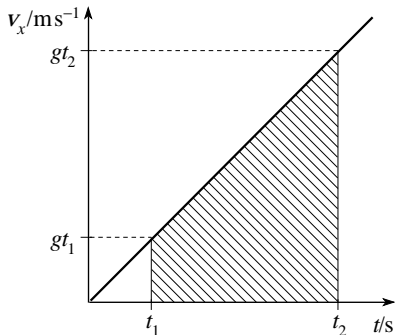

Figure 1 (b) The area under the graph between $t = t_1$ and $t = t_2$.

Note that the 'area' referred to here must be expressed in terms of the scale units that appear on the graph axes, *not in terms of the actual area* of paper in mm^2 or any other unit. In view of this interpretation we can say that

$x(t_2) - x(t_1)$ = the area under the graph of $v_x(t)$ between t_1 and t_2 (15)

◆ If you take g to be $10 \, \text{m s}^{-2}$, what is the hatched area in Figure 1b if $t_1 = 5 \, \text{s}$ and $t_2 = 10 \, \text{s}$? What is the corresponding change in position $x(t_2) - x(t_1)$?

(b)

Figure 1 (b) The area under the graph between $t = t_1$ and $t = t_2$.

If we recall that $x(t)$ is an *inverse derivative* of $v_x(t)$, we can express Equation 15

$x(t_2) - x(t_1) =$ the area under the graph of $v_x(t)$ between t_1 and t_2 (Eqn 15)

in another way:

The area under the graph of $v_x(t)$ between $t = t_1$ and $t = t_2$ is equal to the corresponding change in its inverse derivative $x(t_2) - x(t_1)$.

Although we have only deduced this relationship for one particular form of $v_x(t)$ it is actually a general relationship. In fact, it is a very general relationship since it not only applies to any form of $v_x(t)$ but to any function at all, provided that function has an inverse derivative and provided we take due care over signs (see later). Subject to these provisos we can say quite generally that:

If $F(x)$ is any one of the inverse derivatives (i.e. indefinite integrals) of $f(x)$, then the **area under the graph** of $f(x)$, between a and b where $b > a$, is given by $F(b) - F(a)$.   ☞

In case you are worried that this result, even if true for one inverse derivative of $f(x)$, might not be true for *all* the inverse derivatives of $f(x)$, just take note of the following argument.

If $F_1(x)$ and $F_2(x)$ are both inverse derivatives of $f(x)$ then we know that $F_1(x) = F_2(x) + K$ for some constant K. It therefore follows that

$$F_1(b) - F_1(a) = [F_2(b) + K] - [F_2(a) + K] = F_2(b) - F_2(a)$$

So if the boxed result is true for one inverse derivative of $f(x)$ it is true for *any* inverse derivative of $f(x)$.

A note on notation   As you can see, expressions such as $F(b) - F(a)$ and $x(t_2) - x(t_1)$ are very common in these discussions of inverse derivatives and the area under a graph. In order to simplify the process of writing such expressions we will henceforth indicate them by means of square brackets, as follows

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

$$\text{and } [x(t)]_{t_1}^{t_2} = x(t_2) - x(t_1) \quad \square$$

The boxed statement given above is very important, but so far it is little more than an assertion. Before applying it to any physical or mathematical problems we should really give you a good reason for believing that it is true for *any* $f(x)$.

Figure 2 represents the graph of some general function $f(x)$, the variable x here need not represent position, it could be any quantity. If we let a and b represent particular values of x then the shaded region in Figure 2 represents the area under the graph between $x = a$ and some general value x that is less than or equal to b. This area (measured in the appropriate scale units) will be denoted by $A(x)$.

If we now consider what happens to $A(x)$ when x increases by some small amount Δx, we can see that $A(x)$ will also increase by an amount ΔA. We can approximate this increase in area by the area of the small rectangle shown in Figure 2, so

$$\Delta A \approx f(x)\Delta x \qquad ☞$$

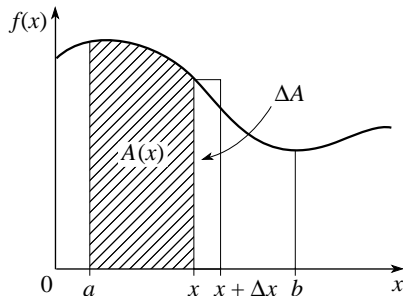and therefore $\qquad \dfrac{\Delta A}{\Delta x} \approx f(x)$

Figure 2 The graph of a general function $f(x)$.

Now, this approximation will become increasingly accurate as Δx becomes smaller, so in the limit as Δx tends to zero, we can say ☞

$$\frac{dA}{dx} = f(x)$$

In other words, the area function $A(x)$ is an *inverse derivative* of the function $f(x)$ whose graph has been drawn. But we have already seen that if the boxed statement is true for one inverse derivative of $f(x)$ it is true for *any* inverse derivative of $f(x)$. Since it is certainly true that the area under the graph between $x = a$ and $x = b$ is equal to $A(b) - A(a)$, it follows that it is also equal to $F(b) - F(a)$ where $F(x)$ is *any* inverse derivative of $f(x)$, as claimed.

Before leaving the general principles and looking at some applications, there is one more point that must be made about the area under a graph and that concerns its sign. So far, both the graphs we have considered have had the property of being positive at all points. However, as you are well aware, it is perfectly possible for a velocity $v_x(t)$, or a general function $f(x)$, to be negative over all or part of its domain of definition.

Would our results about inverse derivatives and areas under graphs still be true if all or part of the graph had been below the horizontal axis, as in Figure 3?

The answer is yes, though in such cases the 'area' below the horizontal axis (such as A_2 in Figure 3) must be treated as a negative quantity. This is a sufficiently important point that it deserves a box of its own.

> When calculating the area under a graph, any regions that are below the horizontal axis should be regarded as having negative areas.

The above results can be used in two ways. If we are given, or can easily find, the area under the graph of a function, then we can determine the corresponding change in the inverse derivative. On the other hand, if we know or can easily find the inverse derivative of a function, then we can determine the area under the graph of that function. Both these applications are illustrated in the following questions which you should now attempt.
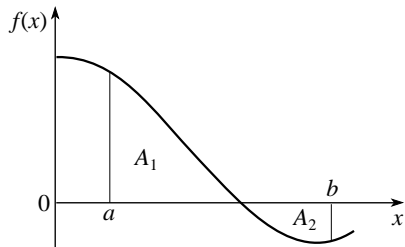

Figure 3 The area under a graph in any region where the graph is below the horizontal axis, such as A_2, will be negative. According to the convention used in *FLAP*, the total area under the graph in Figure 3 is $A_1 + A_2$ where A_2 is a negative quantity. Other authors sometimes adopt a different convention in which the area of each region is defined to be positive and the area under the graph is then given by $A_1 - A_2$.

◆ Figure 4 shows the velocity–time graph of a moving object. Determine the displacement of the object from its initial position after 7 seconds. How far does the object travel during the final second of its journey? (You will need to know that the area of a triangle of base length a and height h is $\frac{1}{2}ha$.)
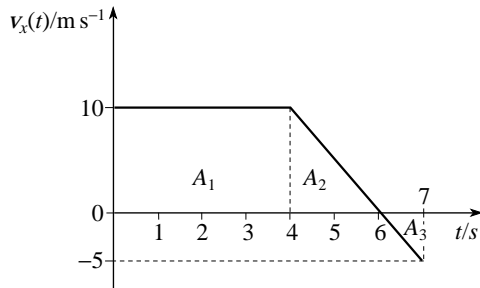

Figure 4 A velocity–time graph in which $v_x(t)$ changes sign.

◈ Find the area under the graph of $f(x) = 3x^2$ between $x = 1$ and $x = 3$ (see Figure 5).

🔲

Note that in answering this last question the value of C makes no difference to the final result. Since $F(x)$ could be *any* inverse derivative of $f(x)$ it makes sense to choose the simplest and take $C = 0$ in such cases.

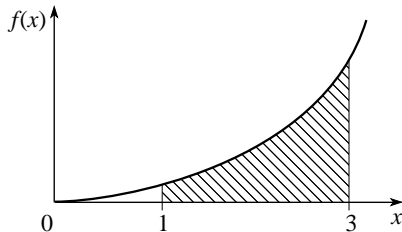◈ Find the magnitude of the area under the graph of $y = 9 - x^2$ between $x = 4$ and $x = 7$.

🔲

Figure 5 The graph of $f(x) = 3x^2$.

Question T4

(a) An object moves along a fixed x-axis at time t with velocity $v_x(t) = v_0 - bt^2$ where $v_0 = 2\,\text{m s}^{-1}$ and $b = 1\,\text{m s}^{-3}$. Determine the displacement of the object from its initial position (i.e. its position at $t = 0$) after two seconds, and explain how that displacement is related to a graph of v_x against t.

(b) By noting that the velocity changes sign at $t = \sqrt{2}$ s, find the total distance travelled by the object in the first two seconds and explain how this relates to a graph of v_x against t. ❏

2.3 Definite integrals: the limit of a sum

We have seen how inverse differentiation can be used to find the area under the graph of a function; however there is an alternative method of evaluating areas which does not require any knowledge of inverse derivatives. This alternative technique is based on the idea that the area under a graph can be estimated by breaking it up into a large number of very thin rectangles with easily calculable areas that can be added together. This idea is illustrated in Figure 7.

In Figure 7 the area under the graph of $f(x)$ from $x_1 = a$ to $x_{n+1} = b$ has been divided into n vertical strips, each of which can be approximated by rectangles. The first rectangle is of height $f(x_1)$ and width $x_2 - x_1$, and the last is of height $f(x_n)$ and width $x_{n+1} - x_n$.
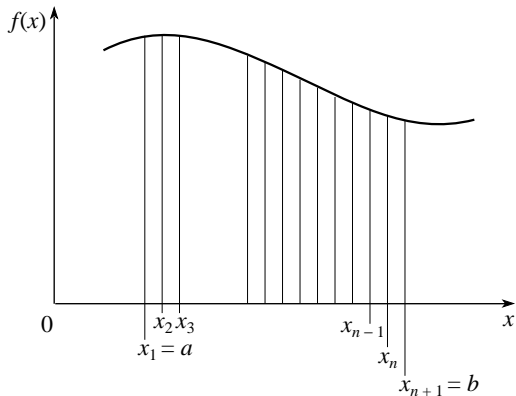

Figure 7 The area under the graph of an arbitrary function may be approximated by a sum of rectangles. ($x_1 = a$, $x_{n+1} = b$)

If we let $\Delta x_i = x_{i+1} - x_i$ where i may be any whole number in the range $1 \le i \le n$, ☞ then we can say that the area of the ith rectangle is $f(x_i)\Delta x_i$, and that the total area under the graph between a and b is approximately

$$f(x_1)\Delta x_1 + f(x_2)\Delta x_2 + f(x_3)\Delta x_3 + \ldots + f(x_n)\Delta x_n = \sum_{i=1}^{n} f(x_i)\Delta x_i \qquad ☞$$

This approximation to the area under the graph will become increasingly accurate if the number of rectangles increases *and* the width of the rectangles decreases. We can ensure that both requirements are met by letting Δx denote the width of the widest rectangle, and then taking the limit of the above sum as Δx tends to zero. Thus we have

the area under the graph of $f(x)$ from a to $b = \lim_{\Delta x \to 0} \left(\sum_{i=1}^{n} f(x_i)\Delta x_i \right)$

Now this limit of a sum is very important, but also very cumbersome, so it is given a special name, it is called a **definite integral**, and it is denoted by a special symbol, a sort of distorted 'S' (for summation). Thus,

$$\int_{a}^{b} f(x)\,dx = \lim_{\Delta x \to 0} \left(\sum_{i=1}^{n} f(x_i)\Delta x_i \right) \qquad\qquad (16)$$

A note on terminology   Definite integrals are of such great importance that each part of the symbol has its own name. The distorted S is called an **integral sign**, the values *a* and *b* are known as the lower and upper **limits of integration**, respectively, the function that is being integrated ($f(x)$ in this case) is called the **integrand** and the *dx* which indicates that the integration is to be performed 'with respect to the variable *x*' is called the **element of integration** or **integration element**. The use of the terms *integral* and *integration* in this context makes good sense, since they imply bringing things together to make a whole and that is exactly what a summation does.   ❑

The process of evaluating a sum and taking its limit can be very tedious, so you will be pleased to learn that there are easier ways of evaluating integrals that will be explained in the next subsection (though you might already guess what they are). However, the idea of a limit of a sum is one that arises quite naturally in many scientific problems and that is why definite integrals are an essential part of the physicist's tool kit. In the remainder of this subsection we will show you how a finite sum can be used to estimate the value of a definite integral, and then invite you to express various physical problems in terms of definite integrals, without asking you to evaluate them.

Example 1   (a) Given that $f(x) = x^2 + 1$, write down the values of $f(x)$ at the following values of x: $x_1 = 1.00$, $x_2 = 1.50$, $x_3 = 2.00$, $x_4 = 2.50$, $x_5 = 3.00$, $x_6 = 3.50$ and $x_7 = 4.00$.

(b) By letting $\Delta x_i = x_{i+1} - x_i$, write down the value of Δx_i for $i = 1$ to 6.

(c) Use your answers to (a) and (b) to work out the corresponding values of $f(x_i)\Delta x_i$ and hence evaluate $\sum_{i=1}^{6} f(x_i)\Delta x_i$.

(d) Explain the relationship between your answer to (c) and the definite integral $\int_{1}^{4}(x^2 + 1)\,dx$.

Solution   The answers to parts (a), (b) and (c) are given in Table 1. ☞

Table 1   Values of $f(x) = x^2 + 1$.

i	x_i	$f(x_i)$	Δx_i	$f(x_i)\Delta x_i$
1	1.00	2.00	0.50	1.00
2	1.50	3.25	0.50	1.63
3	2.00	5.00	0.50	2.50
4	2.50	7.25	0.50	3.63
5	3.00	10.00	0.50	5.00
6	3.50	13.25	0.50	6.63
7	4.00	17.00		

$$\sum_{i=1}^{6} f(x_i)\,\Delta x_i = 20.39$$

(d) The sum that has been evaluated is a crude approximation ☞ to the given definite integral. By dividing the interval from $x = 1.0$ to $x = 4.0$ into finer parts (i.e. by using more values of x that are more narrowly separated), the quality of the approximation can be improved. ❑

The method used to estimate the value of the definite integral in Example 1 is a crude example of a process known as **numerical integration**. We will not pursue that topic in this module except to note that there are far more sophisticated methods of using a finite sum to estimate the value of a definite integral, many of which can be very precise but require a great deal of repetitive labour and are therefore usually implemented on a computer.

Now try the following questions.

◆ (a) A car moves along a straight road with constant speed V for a time $T = t_2 - t_1$. How far does it travel in that time?

(b) At time $t = t_2$ the car starts to slow down, and it eventually comes to rest at time $t = t_3$. Throughout this period its speed is given by

$$v(t) = at - bt^2$$

where a and b are constants. By considering how far the car travels during a short time Δt_i, deduce the form of the definite integral that represents the total distance travelled between t_2 and t_3.

◆ Figure 8 shows a column of atmospheric air of cross-sectional area A and total height H. At a distance h above the ground, the <u>density</u> of the air in the column is given by

$$\rho(h) = \rho_0 e^{-h/\lambda} \qquad \text{where } \rho_0 \text{ and } \lambda \text{ are constants}$$

Use the fact that any small part of the column extending from h to $h + \Delta h$ has volume $A\Delta h$ and a mass approximately given by

$$\rho(h)A\,\Delta h$$

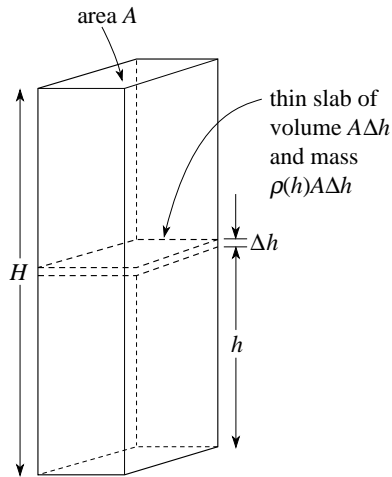to deduce an expression (involving a definite integral) for the total mass of the column.

Figure 8 A column of air of height H, cross-sectional area A and variable density $\rho(h)$.

✦ An elastic string of natural length L can be maintained at a stretched length $L + x$ by applying a <u>force</u>

$$F_x(x) = \frac{\lambda x}{L} \qquad \text{where } \lambda \text{ is a constant}$$

In stretching the string the force is said to do *work*, and the work done in increasing the extension x by a small amount Δx is approximately

$$F_x(x)\,\Delta x$$

If this approximation becomes more accurate as Δx becomes smaller, write down a definite integral for the work done in stretching the string from its natural length to twice its natural length.

So much for the formulation of definite integrals by considering the limits of sums. Let us now turn to their evaluation.

2.4 The fundamental theorem of calculus

In the last subsection we were mainly concerned with the definition of a definite integral as the limit of a sum. But we started the last subsection by showing that the limit of a sum could be used to describe the area under a graph. Thus, a definite integral can be interpreted in terms of an area under a graph. However, we already know (from Subsection 2.2) that the area under the graph of a function $f(x)$ between $x = a$ and $x = b$ is equal to the corresponding change in the value of its inverse derivative $F(b) - F(a)$, where $\dfrac{dF}{dx} = f(x)$.

It follows that we can write:

If $F(x)$ is any inverse derivative of a given function $f(x)$, so that $\dfrac{dF}{dx} = f(x)$, then

$$\int_a^b f(x)\,dx = \left[F(x)\right]_a^b = F(b) - F(a) \tag{17}$$

This remarkable result, relating the limit of a sum to a difference in inverse derivatives is known as the **fundamental theorem of calculus**. It provides the key to evaluating limits of sums and provides another reason for our interest in inverse derivatives. ☞

Example 2

Use the fundamental theorem of calculus to find the value of $\int\limits_{1}^{3} t^4 \, dt$.

Solution   Since $\dfrac{d}{dt}\left(\dfrac{t^5}{5}\right) = t^4$

$$\int\limits_{1}^{3} t^4 \, dt = \left[\dfrac{t^5}{5}\right]_{1}^{3} = \left(\dfrac{3^5}{5}\right) - \left(\dfrac{1^5}{5}\right) = \dfrac{243}{5} - \dfrac{1}{5} = 48.4 \quad \text{☞}$$

Note that the fundamental theorem applies to *any* inverse derivative of the integrand, so we have deliberately chosen to use the simplest in which the arbitrary constant *C* is zero. ❏

✦ Use Equation 17

$$\int_a^b f(x)\,dx = \left[F(x)\right]_a^b = F(b) - F(a) \qquad \text{(Eqn 17)}$$

to evaluate the following three integrals that were obtained at the end of Subsection 2.3

(a) $\int_{t_2}^{t_3} (at - bt^2)\,dt$, (b) $\int_0^H \rho_0 e^{-h/\lambda} A\,dh$, (c) $\int_0^L \frac{\lambda x}{L}\,dx$.

There are two important points to note about the answers to this question.

1   In every case the arbitrary constant C that was included in the inverse derivative played no part in the final answer since the answer only concerned a *difference* of the form $\left[F(x)\right]_a^b$. That is why you were able to put $C = 0$ in each case without any fear of error.

2   In each case the final answer only involved constants, not variables. Thus, even though we may not know the numerical values of all these constants, the result of each definite integral is itself a constant.

Question T5

Evaluate the following definite integrals.

(a) $\int\limits_{1}^{2} x^{1/2}\, dx$,   (b) $\int\limits_{-2}^{0} 4x^2\, dx$,   (c) $\int\limits_{-\pi/2}^{\pi/2} \sin x\, dx$.   ❏

Study comment   In view of its importance it is appropriate to provide some further justification for the fundamental theorem of calculus. However such justification will not actually help you to evaluate integrals. We therefore present this justification in the form of an aside that you may regard as optional reading unless instructed otherwise by your tutor.

Aside — justifying the fundamental theorem of calculus

An informal justification for the fundamental theorem of calculus, without reference to the area under a graph, can be obtained as follows:

Since $\qquad f(x) = \dfrac{dF}{dx} = \lim_{\Delta x \to 0}\left[\dfrac{F(x + \Delta x) - F(x)}{\Delta x}\right]$

for each value of x we have

$\qquad f(x)\Delta x \approx F(x + \Delta x) - F(x)$

Now, the definite integral from a to b is obtained as the limit of a sum of terms of the form $f(x_i)\,\Delta x_i$ where x_i takes values

$\qquad x_1, x_2, \ldots x_n$ with $a = x_1 < x_2 < x_3 < \ldots < x_n < x_{n+1} = b$

and $\Delta x_i = x_{i+1} - x_i$

so that

$$\int_a^b f(x)\,dx \approx \sum_{i=1}^n f(x_i)\,\Delta x_i \approx \sum_{i=1}^n \left[F(x_{i+1}) - F(x_i)\right]$$

$$= \left[F(x_{n+1}) - F(x_n)\right] + \left[F(x_n) - F(x_{n-1})\right] + \ldots + \left[F(x_3) - F(x_2)\right] + \left[F(x_2) - F(x_1)\right]$$

But this last expression is equal to $F(x_{n+1}) - F(x_1)$, since all the intermediate terms, such as $F(x_n)$ and $F(x_2)$ cancel.

However, $x_1 = a$ and $x_{n+1} = b$, so $F(x_{n+1}) - F(x_1) = F(b) - F(a)$.

So, in the limit we have $\int_a^b f(x)\,dx = F(b) - F(a)$

thus justifying the fundamental theorem of calculus. ❏

2.5 Indefinite integrals

Indefinite integrals are nothing new, in fact we have already remarked that 'indefinite integral' is an alternative term for *inverse derivative*; only the notation used is new.

If $F(x)$ is an inverse derivative of $f(x)$, so that $F'(x) = f(x)$, then we write

$\int f(x)\,dx = F(x)$, and we call $\int f(x)\,dx$ an **indefinite integral** of $f(x)$. ☞

This is the notation and terminology that physicists generally use when dealing with inverse derivatives. All of our earlier results involving inverse derivatives can be expressed in terms of *indefinite integrals* and henceforth that is what we will do.

Thus, using C to represent an arbitrary constant, we can now write:

$$\int 3x^2\, dx = x^3 + C \qquad\qquad \text{since} \quad \frac{d}{dx}\left(x^3 + C\right) = 3x^2$$

$$\int e^{5x}\, dx = \frac{e^{5x}}{5} + C \qquad\qquad \text{since} \quad \frac{d}{dx}\left(\frac{e^{5x}}{5} + C\right) = e^{5x}$$

$$\int \sin\left(\frac{x}{2}\right) dx = -2\cos\left(\frac{x}{2}\right) + C \qquad \text{since} \quad \frac{d}{dx}\left[-2\cos\left(\frac{x}{2}\right) + C\right] = \sin\left(\frac{x}{2}\right)$$

In this context it clearly makes good sense to call C the **constant of integration**.

Notice that the indefinite integral of a function is another function. This should be contrasted with the fact that the definite integral of a function between given limits of integration is a constant.

Question T6

Find the following indefinite integrals:

(a) $\int t^3 \, dt$, (b) $\int \cos x \, dx$, (c) $\int e^{2x} \, dx$.

Hint: Use your knowledge of differentiation to guess the answer and then check the result. ❑

Question T7

At time t the acceleration of an object moving along a straight line may be represented by $a_x(t) = a_0(e^{kt} - e^{-kt})$, where a_0 and k are constants.

(a) Evaluate $\int a_x(t) \, dt$ and explain its physical interpretation.

(b) If $k = 1 \text{ s}^{-1}$, and if the position of the object, $x(t)$ at time t, satisfies $x(0) = 0$ and $x(1 \text{ s}) = 1$ m, find the constant of integration in $\int a_x(t) \, dt$, in terms of a_0. ❑

2.6 Summary of Section 2

We now have two types of integral:

(a) The *definite integral*: $\int_a^b f(x)\,dx$, a specific value, defined in terms of the limit of a sum and interpreted as the area under the graph of $f(x)$ between $x = a$ and $x = b$ (with regions below the x-axis regarded as having negative area.)

(b) The *indefinite integral*: $\int f(x)\,dx$, a function $F(x)$, defined by the requirement that $\dfrac{dF}{dx} = f(x)$ and interpreted as the result of reverse differentiation. A given function $f(x)$ generally has infinitely many indefinite integrals any two of which will differ by a constant.

These two types of integral are related by the *fundamental theorem of calculus* which says that if $F(x)$ is any indefinite integral of $f(x)$, then

$$\int_a^b f(x)\,dx = \left[F(x)\right]_a^b = F(b) - F(a)$$

3 The applications of integration

3.1 Some uses of integration

Integration, the analysis and evaluation of definite and indefinite integrals, is of great importance in mathematics, science and technology. We have already seen three ways in which integration can arise in physical problems:

1 as a way of reversing the effect of differentiation;
2 as a way of determining the limit of a sum;
3 as a way of determining the area under a graph (a special case of (2)).

In this subsection we look at more examples of integration to illustrate its power. It is not necessary for you to fully understand all of the physical concepts that are used in these examples in order to appreciate the mathematical points that are being made.

The mass of an object of variable density

Suppose we need to determine the mass of a solid metal bar, of length L, whose density $\rho(x)$ at a distance x from one end is given by $\rho(x) = k(x + a)^2$ for some particular constants k and a. We suppose that the bar has a uniform square cross section of side b.

The mass of the bar between x and $x + \Delta x$ is approximately $b^2\rho(x)\Delta x$ and the total mass is approximately obtained by adding together the masses of all these small pieces. The exact mass is given by the definite integral

$$\int_0^L b^2\rho(x)\,dx = \int_0^L kb^2(x + a)^2\,dx$$

Since $\quad \dfrac{d}{dx}\left(kb^2\,\dfrac{(x + a)^3}{3}\right) = kb^2(x + a)^2 \quad$ it follows that

$$\int_0^L b^2\rho(x)\,dx = \left[kb^2\,\frac{(x + a)^3}{3}\right]_0^L = kb^2\,\frac{(L + a)^3}{3} - kb^2\,\frac{a^3}{3} = \frac{kb^2 L}{3}\left(L^2 + 3La + 3a^2\right)$$

Question T8

A bar of length $2L$ has a circular cross section of radius R. Its <u>density</u>, $\rho(x)$, ☞ when x is measured from the mid-point of the bar, is $\rho(x) = Ae^{2Bx}$. Find an expression for the mass of the bar. (*Hint*: The bar is not symmetrical.) ❑

Summing a force and calculating a total moment

Consider a simplified mathematical model of a dam in which we assume that it is rectangular in shape, of width L and depth d and placed with its plane vertical, as in Figure 9.

The *pressure* at a depth x below the surface of the water is $P(x) = kx$ for some constant k, and so the magnitude of the force acting on a rectangular element of the dam wall of width L and height Δx (and consequently area $\Delta A = L\Delta x$) at a depth x is approximately

$$\Delta F \approx P(x)\Delta A = Lkx\Delta x \qquad ☞$$

since this approximation becomes more accurate as Δx becomes smaller,

the total force on the dam is of magnitude

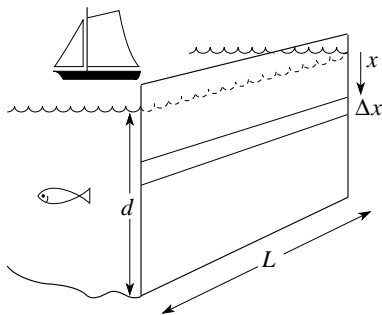$$\int_0^d Lkx \, dx = \left[\frac{Lkx^2}{2} \right]_0^d = \frac{Lkd^2}{2}$$

Figure 9 A rectangular dam.

The designers of such a dam have many other factors to take into account in addition to the total force that it must resist. For example, the pressure of the water also has a twisting effect on the structure. This twisting effect is known as the *torque* of the force, and for the element of area discussed above this *torque* about the line where the dam emerges from the water has a magnitude given by

$$x \times Lk\,x\Delta x = Lk\,x^2\Delta x$$

The total torque about the waterline is given by

$$\int_0^d Lkx^2\,dx = \left[\frac{Lkx^3}{3}\right]_0^d = \frac{Lkd^3}{3}$$

Finding an average

Suppose that a car completes a journey in such a way that it covers a total <u>distance</u> S in a time T, but at a <u>speed</u> $v(t)$ that varies throughout the journey. (Remember the speed of an object is defined as the <u>*magnitude*</u> of its <u>velocity</u>, so although the speed may vary it may never be negative.) Since the car covers a distance S in a time T, it is clear that its average speed is given by

$$\langle v \rangle = S/T \qquad (18) \quad ☞$$

However, since the distance covered during any small time interval Δt will be $v(t)\Delta t$, it is also clear that the total distance covered during the journey will be

$$S = \int_{t_1}^{t_1+T} v(t)\,dt \qquad (19)$$

where t_1 denotes the time at which the journey started. It follows from Equations 18 and 19 that

$$\langle v \rangle = \frac{1}{T} \int_{t_1}^{t_1+T} v(t)\,dt \qquad (20)$$

Thus, the integral provides a way of *defining* the average of a quantity that varies continuously. This definition has many applications.

Question T9

A bar of length $2L$ has a <u>temperature</u> $T(x)$ that varies along its length, when x is measured from the mid-point of the bar. If $T(x) = A \cos(Bx)$, find an expression for the average temperature of the bar. What is the value of this average temperature if $L = 2.0\,\text{m}$, $A = 30\,\text{K}$ and $B = 0.4\,\text{m}^{-1}$? ❏

3.2 The stability of a satellite

To end this section we consider one final example, concerning the stability of an orbiting satellite, that brings together some of the ideas that have been introduced earlier.

The motion of a satellite in orbit about the Earth is generally very complicated, so we will consider a highly idealized case of such motion. First we will suppose that the satellite can be modelled by a uniform thin rod, of length $2L$ and uniform mass per unit length ρ, which means that its total mass is $2L\rho$ and that on the Earth's surface it would balance about its mid-point (its *centre of mass*). We will also suppose that the satellite is placed in a circular *geostationary orbit* above the equator, in which it takes 24 hours to complete an orbit and thus appears fixed in the sky from any point on the Earth's surface.

The distance from the centre of the Earth to the satellite will be denoted by *H*.

Finally, we will assume that when placed in its orbit the satellite is made to rotate about an axis through its centre at a rate of exactly 360° in 24 hours in such a way that it remains parallel to the Earth's surface throughout its orbit (as indicated in Figure 10a). ☞

Once the satellite is placed in orbit (and given its initial rotation) all the stabilizers are switched off, and it is subject only to the effects of the Earth's gravity. The question we wish to address is this: 'Will the satellite remain with its orientation parallel to the Earth's surface (as in Figure 10a), or will it begin to tumble?'

To answer the question we need to know a little about the force due to gravity. If two particles of mass *M* and *m* are a distance *r* apart, then each is attracted towards the other with a force of magnitude GmM/r^2 (where *G* is known as *Newton's gravitational constant*).
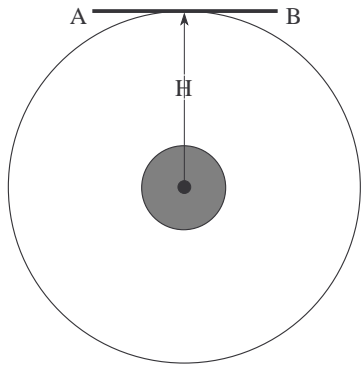
(a)

Figure 10 A satellite (AB) in geostationary orbit.

A large spherical object, such as the Earth, exerts the same gravitational force as a particle of equal mass placed at its centre, and, since the satellite is small compared to the size of its orbit, we can assume that all the forces act vertically downwards, i.e. parallel to the line through the centre of the rod (satellite) and the centre of the Earth.

We will now suppose that the satellite is disturbed very slightly, so that it is at an angle α to the horizontal (as in Figure 10b), and we will attempt to find the torque, or turning effect, of the forces due to gravity.
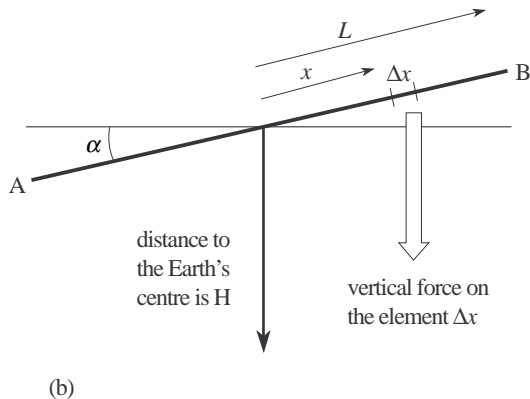

(b)

Figure 10   A satellite (AB) in geostationary orbit.

If we find that the torque tends to restore the satellite to its horizontal position we can conclude that the satellite is stable, but if we find that the torque tends to increase the angle α we must conclude that the satellite is unstable since a small disturbance may send it into a spin from which it will not recover.

If we consider a small element of the satellite of length Δx with position coordinate x (measured along the satellite, from its centre), then the mass of that element will be

(density per unit length) $\times$ (length of element) $= \rho \Delta x$

and the distance of the element from the centre of the Earth will be $H + x \sin \alpha$.

Now the downward vertical force acting on the element will have a magnitude given approximately by

$$\frac{(\text{gravitational constant}) \times (\text{mass of Earth}) \times (\text{mass of element})}{(\text{distance of element from centre of Earth})^2}$$

So, if we represent this magnitude by ΔF and let M represent the mass of the Earth we can write

$$\Delta F \approx \frac{GM\rho\Delta x}{(H + x \sin \alpha)^2}$$

The torque, or turning effect, of the vertical force on the element, about a line that passes through the centre of the satellite and is perpendicular to the plane of Figure 10, is $x \Delta F$.

In this case ΔF (being a magnitude) must be positive, but x (and consequently $x \Delta F$) may be positive or negative. A positive torque will tend to reduce α and restore stability, a negative torque will tend to increase α and lead to instability.

To find the total torque on the satellite we need to consider the definite integral

$$\int_{-L}^{L} \frac{GM\rho x}{(H + x \sin \alpha)^2} \, dx$$

This integral can certainly be evaluated using an appropriate indefinite integral, but the techniques required to deduce that indefinite integral are beyond the scope of this module.
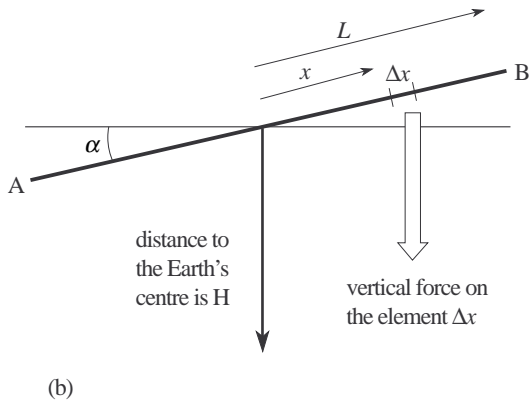

(b)

Figure 10 A satellite (AB) in geostationary orbit.

However, we can use the interpretation of a definite integral as an area under a graph to obtain the result we need. If we put

$$f(x) = \frac{GM\rho x}{(H + x\sin\alpha)^2}$$

then, for suitably chosen values of the constants, we may plot the graph of $f(x)$, as in Figure 11.

The required integral is the area under the graph of $y = f(x)$ between $-L$ and L. From the graph it is easy to see that the negative contribution, from the interval $-L \le x \le 0$, is greater in magnitude than the positive contribution, from the interval $0 \le x \le L$. This means that the net result must be *negative* and therefore the total moment acts so as to increase the angle α. So our conclusion is that any small disturbance of the satellite will lead to an even greater disturbance, in other words, it is *unstable*. ☞
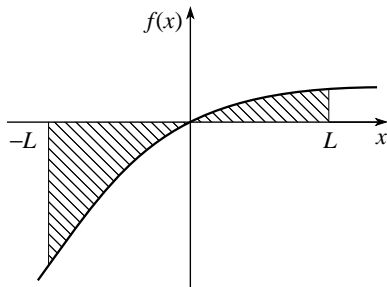

Figure 11 Areas under the graph representing the positive and negative torques acting on the satellite.

We have only dealt with the simplest of cases. In practice one would need to consider more complicated structures, and the effects of other factors such as expansion due to solar heating, but nevertheless the example illustrates the importance of integration and the value of the methods that we have introduced in this module in evaluating physical problems.

4 Conclusion

4.1 The importance of integration

The concept of a *definite integral* (the *limit of a sum*) is fundamental to much of applicable mathematics, and is essential to a proper study of physics.

In this module we have seen how a variety of physical problems can be formulated in terms of definite integrals and we have also seen some of the ways in which such integrals can be evaluated. Finite sums and areas under graphs can both provide estimates of the values of definite integrals, but normally the most convenient method is to use the fundamental theorem of calculus to relate the definite integral to a difference in values of an appropriate indefinite integral, if such a function can be found. Almost all the definite integrals considered in this module have corresponded to indefinite integrals that it has been easy to express in terms of elementary functions (such as e^x, $\sin x$, x^n, and so on). However, this is not always so. As your studies continue it is quite certain that you will encounter integrals, both definite and indefinite, of a more challenging kind. Some of the methods that can be used to deal with these integrals are explored in other *FLAP* modules.

The following questions are intended to improve your basic skill in evaluating integrals, and to remind you of some of the important ideas introduced in this module.

Question T10

Evaluate the following indefinite integrals:

(a) $\int (1 + x + x^2)\,dx$ (b) $\int (\sin x + \cos x)\,dx$

(c) $\int \left(e^{-x} + \dfrac{1}{x^3} \right) dx$ (d) $\int 2\,dw$ ❏

Question T11

Evaluate the following definite integrals:

(a) $\displaystyle\int_{-1/2}^{1/2} \sin(\pi x)\,dx$,   b) $\displaystyle\int_{0}^{3} (3t + 2)^2\,dt$,   (c) $\displaystyle\int_{0}^{1} \frac{(1 + e^x)}{e^x}\,dx$,   d) $\displaystyle\int_{1}^{4} \sqrt{x}\,(x + 1)\,dx$.   ❏

Question T12

Find the sum of the magnitudes of the two shaded areas in Figure 12, i.e. find the magnitude of the area enclosed by the graph of the function $y = e^x - 1$, the x-axis and the lines $x = -1$ and $x = 1$. Note that in this case you are not being asked for 'the area under the graph' which may be negative, but for the *magnitude* of the area *enclosed*, which must be positive. ❏

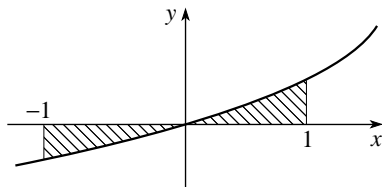**Figure 12** See Question T12.

Question T13

The acceleration, velocity and position coordinate of an object moving along a straight line are represented by $a_x(t)$, $v_x(t)$ and $x(t)$, respectively. What physical meaning, if any, can you associate with the following?

$$\int_{t_1}^{t_2} a_x(t)\,dt, \quad \int_{t_1}^{t_2} v_x(t)\,dt \quad \text{and} \quad \int_{t_1}^{t_2} x(t)\,dt \quad ❏$$

5 Closing items

5.1 Module summary

1   If $f(x)$ is a given function and $F(x)$ is any function such that $\dfrac{dF}{dx} = f(x)$ then we write $F(x) = \int f(x)\,dx$ and we call $F(x)$ an *indefinite integral* (or *inverse derivative* or *primitive*) of $f(x)$.

2   If $F_1(x)$ and $F_2(x)$ are both indefinite integrals of the same function $f(x)$, then there exists a constant K such that

$$F_1(x) = F_2(x) + K$$

3   For motion in a straight line, if the *position*, *velocity* and *acceleration* at time t are respectively $x(t)$, $v_x(t)$ and $a_x(t)$ then

$$v_x(t) = \frac{dx}{dt} \quad \text{and} \quad a_x(t) = \frac{dv_x}{dt}$$

so, $x(t) = \int v_x(t)\,dt \quad \text{and} \quad v_x(t) = \int a_x(t)\,dt$

4   The *definite integral* of a function $f(x)$ from $x = a$ to $x = b$ is defined by the limit of a sum and may be written as

$$\int_a^b f(x)\,dx = \lim_{\Delta x \to 0}\left(\sum_{i=1}^n f(x_i)\,\Delta x_i\right) \qquad \text{(Eqn 16)}$$

where Δx is the width of the largest subinterval, and where $x_1 = a$ and $x_{n+1} = b$.

5   According to the *fundamental theorem of calculus*, if $F(x)$ is any indefinite integral of a given function $f(x)$, so that $\dfrac{dF}{dx} = f(x)$, then

$$\int_a^b f(x)\,dx = \left[F(x)\right]_a^b = F(b) - F(a)$$

This often provides a convenient way of evaluating definite integrals.

6   The definite integral $\int_a^b f(x)\,dx$, with $b > a$, may be interpreted as *the area under the graph* of $f(x)$ between $x = a$ and $x = b$. ☞ It follows from the fundamental theorem of calculus that the area under the graph of $f(x)$ between $x = a$ and $x = b$ is equal to the corresponding change in the indefinite integral, $F(b) - F(a)$. When calculating the area under a graph, any regions that are below the horizontal axis should be regarded as having negative areas.

5.2 Achievements

Having completed this module, you should be able to:

A1 Define the terms that are emboldened and flagged in the margins of the module.

A2 Define an indefinite integral of a given function in terms of a process that reverses the effect of differentiation, and use that definition to express physical quantities (such as position and velocity, in the case of linear motion) as indefinite integrals of other physical quantities (such as velocity and acceleration).

A3 Determine (by inspection) indefinite integrals of a range of simple functions in various mathematical and physical contexts, recognizing that the process of indefinite integration generally introduces an arbitrary constant.

A4 Define the definite integral of a given function between given limits in terms of the limit of a sum, and use that definition to formulate various physical problems in terms of definite integrals.

A5 Recognize that the definition of a definite integral as the limit of a sum can be used to justify the use of a finite sum to estimate the value of a given definite integral.

A6 Interpret the definite integral of a given function between given limits in terms of the area under the graph of that function between those limits (with due regard to signs), and use this identification to determine areas and/or evaluate definite integrals in sufficiently simple cases.

A7 State the fundamental theorem of calculus that relates definite and indefinite integrals, and use it to evaluate definite integrals in sufficiently simple cases.

Study comment   You may now wish to take the *Exit test* for this module which tests these Achievements. If you prefer to study the module further before taking this test then return to the *Module contents* to review some of the topics.

5.3 Exit test

Study comment   Having completed this module, you should be able to answer the following questions, each of which tests one or more of the *Achievements.*

Question E1

(*A2*, *A3* and *A5*)   The acceleration of an object moving along the *x*-axis is given at time *t* by

$$a_x(t) = a_0(e^{t/k} + e^{-t/k})$$

where a_0 and k are constants. If $a_0 = 2.00 \text{ m s}^{-2}$ and $k = 1.00 \text{ s}$, obtain expressions for the velocity $v_x(t)$ and the position $x(t)$ given that $v_x(0) = x(0) = 0$.

What is the area under the graph of $a_x(t)$ against *t* between $t = 1 \text{ s}$ and $t = 2 \text{ s}$, and what does this area represent?

Question E2

(*A5* and *A6*)   The indefinite integral $\int \exp(x^2)\,dx$ ☞ cannot be expressed in terms of elementary functions, so it is not possible to evaluate the definite integral $\int_0^1 \exp(x^2)\,dx$ in a straightforward way using the fundamental theorem of calculus. Describe two other ways in which you might estimate the value of this definite integral. (You are not required to perform these evaluations.)

[?]

Question E3

(*A3*)   Evaluate the following indefinite integrals.

(a) $\int x^{1/2}\,dx$,   (b) $\int (x+3)(x+2)\,dx$,   (c) $\int 6\cos(3x)\,dx$,

(d) $\int (e^x + e^{-x})(e^x - e^{-x})\,dx$.

(*Hint*: By expanding where appropriate, you should be able to determine by inspection (i.e. guess) a function which, when differentiated, produces the given integrand in each case.)

[?]

Question E4

(*A7*) Use the fundamental theorem of calculus to evaluate each of the integrals in Question E3

(a) $\int x^{1/2}\, dx$, (b) $\int (x+3)(x+2)\, dx$, (c) $\int 6\cos(3x)\, dx$,

(d) $\int (e^x + e^{-x})(e^x - e^{-x})\, dx$.

between 0 and 1, i.e. with lower limit 0 and upper limit 1.

Question E5

(*A3*, *A6* and *A7*)   Find the sum of the magnitudes of the shaded areas in Figure 13, i.e. find the area enclosed by the graph of the function $y = x^3 + 1$, the x-axis and the lines $x = -2$ and $x = 2$.

(Note that in this case you are not being asked for 'the area under the graph,' which would be the difference of these magnitudes.)
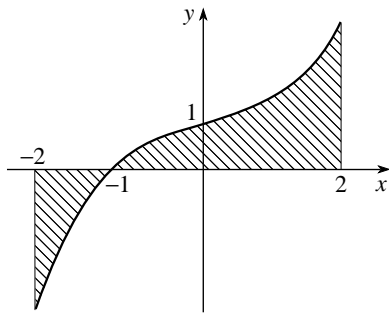

Figure 13   See Question E5.