INTEGRALS

In the page on derivatives, we see that the derivative of a function $f(x)$ gives the function's rate of change at each point over which the function is defined. Thus every function satisfying some basic conditions (which means basically almost every function that is used in physics) has another function that is its derivative. A natural question to ask, therefore, is - is every function also the derivative of some other function? In other words, is there a way of inverting the process of differentiation (as the act of calculating the derivative is known) to go from a derivative back to the function that was differentiated to get that derivative?

It won't surprise you to learn that integration is the process by which this inversion of differentiation is done. However, if we just define the integral of a function to be its 'anti-derivative' and leave it at that, we miss out on a very important property of integration. In fact, the usual way of studying integration is to *start* with this property and derive from it the fact that an integral is an antiderivative.

So what is this vital property of integration? It is that the integral of a function, calculated between two different points, gives the area under the curve of that function. If we adopt this as the definition of an integral, we can then derive the antiderivative property from it. This importance of this derivation is clear from its popular name: it is known as the *fundamental theorem of calculus*.

Since the main goal of this page is to give the reader a feel for how integration works so that it can be applied to physics, we won't discuss things with iron-clad rigour. With that caveat in mind, let's see how an integral can be defined and then see how the fundamental theorem of calculus works.

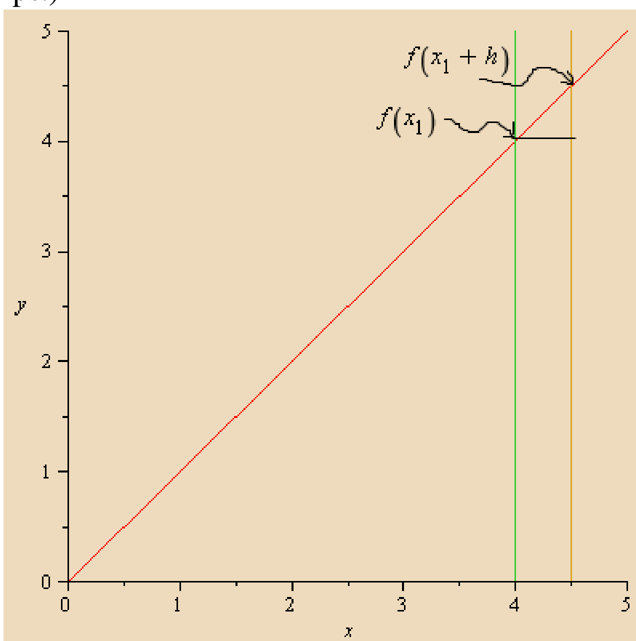
We'll begin with a very simple curve, and you should keep this curve in mind as you work through what follows. Drawing pictures as you go along can also be quite helpful.

$$f(x) = x \tag{0.1}$$

The graph of this function is the 45-degree straight line that passes through the origin. If we want to find the area under this curve between the points $x = 0$ and $x = a$ for some constant a, we note that this area is just a right-angled triangle with its base and height both equal to a. Thus we can use the standard formula for the area of a triangle to see that the area is

(0.2)
$$A = \frac{1}{2}a^2$$

Now let's see how we can find the area using an integral. The way an integral calculates an area is by dividing the area under a curve up into a number of vertical strips, each of width h. (If you are following along using the curve above as an example, you can draw the curve, and then draw a number of equally spaced vertical lines to divide the triangle up into vertical strips.)



In general, of course, the curve that provides the upper bound to the area we're trying to find can have any shape. It could be a straight line as in our example with the triangle, or it could have some other, in general, curved shape. In the diagram, we have drawn the function $f(x) = x$ and taken $x_1 = 4$ and the width of the vertical strip is $h = 0.5$.

So to begin with, the integral approximates the area of each vertical strip by finding the point at which the left edge of the strip intersects the bounding curve, and then just drawing a horizontal line from that point across to the right edge of the strip, thus creating a rectangle. If we're considering the strip whose left edge coincides with $x = x_1$ then the height of the rectangle is $f(x_1)$. Since the width of all the strips is h, the area of that rectangular strip is $f(x_1)h$.

Now suppose we define another function $F(x_1)$ to be the area under the curve between $x = 0$ and $x = x_1$. If we extend the area beyond x_1 by considering the next strip, then the area of that strip is $f(x_1 + h)h$ so

$$F(x_1+h) = F(x_1) + f(x_1+h)h \tag{0.3}$$

That is, the total area under the curve between $x = 0$ and $x = x_1 + h$ is the area between $x = 0$ and $x = x_1$ plus the strip between $x = x_1$and $x = x_1 + h$.

If we rearrange this equation a bit, we get

$$f(x_1+h) = \frac{F(x_1+h) - F(x_1)}{h} \tag{0.4}$$

Readers of the page on derivatives may notice the expression on the right has a familiar form. It looks a lot like the formula for the derivative of $F(x)$; all that's missing is the limit as $h \to 0$. However, the use of a non-zero value for h was just to get the idea of how the area can be calculated. Clearly if $h > 0$, any such calculation of the area will be approximate (unless the curve bounding the area is horizontal, but that's a bit too simple), so to get the exact area, we will have to take a limit (well this is calculus, after all; we have to have a limit in there somewhere). So (dropping the subscript from x_1 and writing the formula in terms of x:

$$f(x) = \lim_{h \to 0} \frac{F(x_1+h) - F(x_1)}{h} \tag{0.5}$$

$$= \frac{dF(x)}{dx} \tag{0.6}$$

That is, the curve $f(x)$ that defines the upper bound of an area is the derivative of the function $F(x)$ that specifies the area itself.

If you've been watching critically, you might justifiably complain about something at this point. We defined $F(a)$ to be the area under the curve between $x = 0$ and $x = a$, but in fact the same argument works no matter what endpoints we choose. That is, we could equally well have defined $F(a)$ to be the area between $x = 1$ and $x = a$ or between any point and $x = a$. What gives?

The point is that the formula $f(x) = dF(x)/dx$ says that the curve $f(x)$ says how fast the area is *changing* at each point, not what the area actually is at that point. So it doesn't matter what we choose as the starting point for measuring the area; all that matters is how fast the area is increasing or decreasing (areas below the x-axis count as negative). Changing the starting point at which we measure the area is equivalent to adding a constant to $F(x)$ and this constant disappears when we take its derivative.

We can invert the relation above to specify $F(x)$ in terms of $f(x)$, and thus introduce the integral sign:

(0.7) $$F(x) = \int f(x)\, dx$$

For a simple example, we return to our curve above: $f(x) = x$. According to the notion of integration, the bounding curve must be derivative of the function that defines the area under the curve. We therefore have

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

from which we get

(0.9) $$F(x) = \frac{1}{2}x^2$$

Well, that's *one* solution anyway, and corresponds to the case where we measure the area starting from $x = 0$. If we started measuring the area from some other point such as $x = x_0$ then we would have

(0.10) $$F(x) = \frac{1}{2}x^2 - \frac{1}{2}x_0^2$$

Since the second term is a constant, its derivative is zero, so the relation $dF/dx = x$ still holds true.

That's the essence of what integration is. As with derivatives, there are a number of formulas (mainly obtained by inverting the corresponding formulas for derivatives) that allow all the simple functions to be integrated. There are also a few techniques, such as integration by parts, that allow more complex functions to be integrated. However, unlike derivatives, there is no general catch-all method by which any function can be integrated. This fact gives mathematics teachers unlimited scope for devising exam questions, but on a broader scale, can be real barrier to progress in theoretical physics. There are many weighty tomes giving tables of integrals for various types of functions, and symbolic algebra packages such as Maple, Mathematica and Macsyma allow the calculation of integrals, but if you do much work in physics you will encounter some integrals that are very difficult to work out.