Integration

An integral is an infinite sum of infinitesimal terms.

An integral is an operator. I want to convey what I mean by the term "operator" by first discussing an operator with which you are already familiar, namely the derivative operator. An operator is just a mathematical agent that acts on a function to yield another function. The derivative operator $\frac{d}{dx}$ (the derivative with respect to *x*) acts on a function of *x* to yield another function of *x*. So, for instance, if *y* is a function of *x* then we can apply the derivative operator to it. We write the derivative of *y*(*x*) as $\frac{d}{dx}y$ or $\frac{dy}{dx}$. I now write the derivative operator as $\frac{d}{dx}\square$ to emphasize the fact that the derivative operator operates on a function and that function goes where the box is. The integral operator is an operator too. Using the "box" notation the integral operator can be written $\int \square \, dx$. (But see footnote 1.) The integral operator operates on a function and that function goes where the box is. The point here, is not about the order. (In fact, it is perfectly okay to write the integral operator as $\int dx \square$.) The point is that the integration operator includes both the *integral sign* "$\int$" **and** the *differential* "*dx*". Remember, an integral is an infinite sum of terms. An infinite sum of finite terms is infinite. For the integral to be finite (and thus meaningful), the terms themselves must be infinitesimal[2]. So, given the function $f(x)$, $\int f(x)$ is nonsense, whereas, $\int f(x)\, dx$ is the integral of f with respect to *x*. Let's make that a little more concrete. Consider the function $f(x) = x^2$. The expression $\int x^2$ is nonsense, whereas, $\int x^2 dx$ is the integral of x^2 with respect to *x*. The bottom line is, if you want to stick an integral sign "$\int$" in front of something, you better make sure that something has a differential in it, otherwise, your final expression is nonsense.

Enough about notation and jargon, let's explore the big idea behind integration by means of an example. Consider an object whose velocity as a function of time is given by $v(t) = 1.5\frac{m}{s^3}t^2$. Assume you need to know how far it goes during the first 4.0 seconds of its motion. At first you might be tempted to use your junior high school formula "distance is speed times time". The "time" is clearly 4.0 seconds but what are you going to use for the speed. If you evaluate $v(t)$ at $t = 4.0$ seconds you get the speed at the 4 second mark but the object is not going that fast for the whole four-second time interval from 0 to 4.0 s. Its speed is 0 at time zero, 1.5 m/s at 1 s, 6.0 m/s at 2 seconds, 13.5 m/s at 3 seconds, and 24 m/s at 4 seconds. The object is clearly speeding up during the entire 4 seconds. You might try using 12 m/s for the average speed calculated by adding the initial velocity and the final velocity and dividing by two. But that only works if the

[1] I am using the generic variable name *x*. The variable does not have to be *x*. It could be *t* or *z* or anything else.
[2] Infinite is unimaginably large. Infinitesimal is vanishingly small. Finite is of ordinary size.

acceleration is constant. Note that you can solve $v(t)=1.5\frac{m}{s^3}t^2$ to find that the velocity is 12 m/s at time t = 2.83 s. So, in the case at hand, the object goes slower than 12 m/s for the first 2.83 s which is more than half of the 4-second time interval. So the average speed has to be less than 12 m/s. Again that business of the average velocity over a time interval being half the sum of the velocity at the start of the time interval and the velocity at the end of the time interval only works in the case of constant acceleration, and you should be able to discern that if the velocity of an object is given by $v(t)=1.5\frac{m}{s^3}t^2$, then the acceleration of the object is not constant. This rules out any constant acceleration equations implementation that we might have had in mind. So how do we go about finding out how far the object goes in the first 4.0 seconds?

If we had an easy way to get the average speed, knowing what we know, we could just use that and multiply the average speed by the four seconds to get the distance. But, in general, there is no easy way to get the average speed. So again I ask, how *do* we go about finding out how far the object goes in the first 4.0 seconds?

Newton is the one that came up with the idea. Essentially, what he said was that one should use an approximation scheme which, taken to extremes, yields the correct answer. Here's the idea. Divide the time interval up into several parts. Let's try four parts at first. Then, each time interval is 1 second. The distance will be the average speed during the 1st second, times one second; plus the average speed during the 2nd second, times one second; plus; the average speed during the 3rd second, times one second; plus the average speed during the 4th second, times one second. Again, the difficulty is, I don't know what the average speed during any one of the four one second time intervals is. Still, whether I use, as my estimated speed for the average speed during the time interval, in each case, the speed at the start of the one-second time interval, the speed at the end of the one second time interval, or the sum of the two divided by 2, the end result will be closer to the actual value than I'd get by using the corresponding value for the entire four-second time interval. Why? Simply because the speed doesn't change as much during one second as it does during four seconds. So, using any reasonable method to "guess" the average speed during each one second time interval will yield a result closer to the actual value than using the same method to arrive at a guess for the entire four second time interval.

My inclination in refining a method like this would be to work real hard at getting the best estimate for the average speed during a particular time interval. But Newton was smarter than that. He said that we shouldn't waste any time getting a better estimate of the average speed. Rather, that we should use shorter time intervals. He took it to extremes. Suppose we divide the original four-second time interval up into four million equal time intervals. Each one is a millionth of a second. For each millionth of a second, from zero to four seconds, if we multiply the average speed during that millionth of a second, times one millionth of a second, and add all the results together, we get the distance traveled during the first four seconds. Of course, we don't know the average speed during any given millionth of a second, but, we're not going to be far off whether we use the value at the start of that millionth of a second, the value at the end of that millionth of a second, or the sum of the two divided by 2.

Check it out. Suppose we always used the value at the start of the time interval. Every term in our sum will be a little bit on the low side because the actual average speed during the millionth of a second time interval will be a little greater than the speed at the start of the time interval. Now do the whole procedure using the speed at the end of each time interval. This results in a distance that is greater than the actual distance. But, by how much. Your second sum of four million terms will look just like the first one, but, it will be missing the first term (0 m/s times 0.000001s) and it will include a term, its last term (24 m/s times 0.000001 s), that does not appear in the first sum. So, the second attempt at calculating the distance will yield a value that is 0.000024 m greater than the first attempt. The first attempt is too low, the second too high. The actual distance is somewhere in between. But the two results differ from each other by only 0.000024 m. So the discrepancy between either distance and the actual distance is even less than that.

You'd think that would be good enough for Newton. Dividing a four-second time interval up into 4 million 0.000001 s time intervals is already pretty extreme in my book. But he carried it even further. He divided it up into an infinite number of infinitesimal time intervals. This drops the discrepancy between the low estimate and the high estimate down to zero meaning that (since the actual distance is between the two, either one is the actual distance. So he winds up with an infinite sum of infinitesimal elements, something we now call an integral.

For the case at hand, we write the integral as:

$$\Delta x = \int_0^{4.0\text{s}} v(t)\, dt$$

It represents the infinite sum of terms

$$v(0)dt + v(dt)dt + v(2dt)dt + v(3dt)dt + ... + v(4\text{s} - 3dt)dt + v(4\text{s} - 2dt)dt + v(4\text{s} - dt)dt + v(4\text{s})dt$$

For each value of t, starting at $t = 0$, and working our way up to $t = 4.0$ s, we evaluate v at t and multiply the result by dt. We add that product to an accumulating sum. Then we increase t by dt and repeat until t reaches the value of 4.0 s. The end result of the sum is the distance traveled by the object from time 0 to time 4.0 s.

Some remarks on the jargon. In the expression $\Delta x = \int_0^{4.0\text{s}} v(t)\, dt$, the variable t is referred to as the variable of integration. $v(t)$ is the function that is being integrated. The 0 is the lower limit of integration and the 4.0 s is the upper limit of integration. The equation reads, "Delta x is equal to the integral of $v(t)$ from zero to four seconds". The result of the integration can be expressed as a function evaluated at the upper limit of integration minus the same function evaluated at the lower limit of integration. The function is called the antiderivative of the function that is being integrated. If we name the antiderivative $x(t)$, what we are saying can be written

$$\int_0^{4.0\text{s}} v(t)\, dt = x(4.0s) - x(0)$$

Determining the antiderivative of a function is a matter of evaluating the infinite sum of infinitesimal terms, the infinite series, that an integral represents. Fortunately, all of the hard work has been done for us my the mathematicians, so here, in your physics course, we can provide you with a few simple rules for arriving at the results. If you haven't already done so, you will learn where these simple rules come from in your calculus class. Here, we simply present them to you, along with some information on notation and usage, without proof.

First a few comments on the relation between the derivative operator and the integral operator. We again rely on the example of the object whose velocity is given by $v(t)=1.5\frac{\text{m}}{\text{s}^3}t^2$ to make our points. First off, as you know, the velocity $v(t)$ is just the time derivative of the position variable *x*: $v(t)=\frac{dx}{dt}$. This means that our integral $\int_0^{4.0\text{s}} v(t)\,dt$ is the same thing as $\int_0^{4.0\text{s}} \frac{dx}{dt}\,dt$. We can treat this expression as if the *dt*'s cancel and write it as $\int_0^{4.0\text{s}} dx$. We don't need any fancy rules to interpret this. It represents the sum of all the infinitesimal changes in position *dx* that the object experiences from time $t = 0$ to time $t = 4.0\text{s}$. This is nothing but the total change in position of the object, which we can express as $x(4.0\text{s}) - x(0)$. So the antiderivative function is just our position x. (But see footnote 3.) Check it out. Start with *x*. Take the derivative of *x* with respect to *t*, $\frac{dx}{dt}$. Now integrate that. You get *x* back. Integration is the inverse operation to taking the derivative. It works the other way too.

If you integrate a function, and then take the derivative of the result, you get the original function back. So when you integrate a function, what you are doing is finding a function whose derivative is equal to the original function. Here, when we say "integrate a function" we really mean "find the antiderivative of a function." Such an integral is called an indefinite integral and is written without limits of integration. (As you might guess, an integral that includes the limits of integration is called a definite integral.) Let's use this information to arrive at an answer for the example we have been talking about.

We need to calculate $\int_0^{4.0\text{s}} v(t)\,dt$ for the case in which $v(t)=1.5\frac{\text{m}}{\text{s}^3}t^2$. That is, we need to calculate $\int_0^{4.0\text{s}} 1.5\frac{\text{m}}{\text{s}^3}t^2\,dt$. Now this represents an infinite sum in which every term is being multiplied by the constant $1.5\frac{\text{m}}{\text{s}^3}$. We can factor that constant out of the sum and write the

---

[3] The position function is actually the initial position plus the antiderivative. Here we specialize to the case of an initial position of zero. In general, when you find an antiderivative of $f(x)$ you are finding a function $g(x)$ whose derivative is $f(x)$. Add any constant to $g(x)$ that you want. Call the result $h(x) = g(x) + \text{constant}$. $h(x)$ must also be an antiderivative of $f(x)$ because the derivative of $h(x)$ is the derivative of g(x) plus the derivative of the constant (which is of course 0). So if the derivative of $g(x)$ is $f(x)$ then the derivative of $h(x) = g(x) + \text{constant}$ is also $f(x)$. That means that $f(x)$ has an infinite set of antiderivatives, one for each of the infinite number of possible values of the constant.

integral as $1.5\frac{\text{m}}{\text{s}^3}\int_0^{4.0\text{s}} t^2\, dt$. To evaluate the $\int_0^{4.0\text{s}} t^2\, dt$ part we need to find a function whose derivative is t^2. To wind up with the power of your variable when taking derivatives, you just start with the next higher power. The derivative of t^3 is $3t^2$. We can compensate for that 3 that comes down when you take the derivative of t^3 by including a factor of $\frac{1}{3}$ in our guess for the antiderivative of t^2. That makes $\frac{1}{3}t^3$ our current candidate for the antiderivative of t^2. Indeed, if you take the derivative of it with respect to t you get t^2 so $\frac{1}{3}t^3$ is indeed an antiderivative of t^2. Now we introduce a bit more notation. Having established that $\frac{1}{3}t^3$ is the antiderivative of t^2, we write

$$1.5\frac{m}{s^3}\int_0^{4.0\text{s}} t^2\, dt = 1.5\frac{m}{s^3}\left(\left. \frac{1}{3}t^3 \right|_0^{4.0\text{s}} \right)$$

That $\Big|_0^{4.0\text{s}}$ part is to be read "evaluated from 0 to 4.0 s" and we implement it on the next line by plugging the upper value in for *t*, writing a minus sign, and then copying the expression with 0 plugged in for *t*. So the next line reads

$$1.5\frac{\text{m}}{\text{s}^3}\int_0^{4.0\text{s}} t^2\, dt = 1.5\frac{\text{m}}{\text{s}^3}\left(\frac{1}{3}(4.0s)^3 - \frac{1}{3}(0)^3\right)$$

$$1.5\frac{\text{m}}{\text{s}^3}\int_0^{4.0\text{s}} t^2\, dt = 32\ \text{m}$$

Hey, this is the integral $\Delta x = \int_0^{4.0\text{s}} v(t)\, dt$ that answers the question: How far does an object whose velocity $v(t) = 1.5\frac{\text{m}}{\text{s}^3}t^2$ go during the first four seconds of its motion. In the process we have come up with a rule for the integral of a power.

Let's write that power rule as a function of *x*. Consider the function $f(x) = cx^{\text{n}}$ where *c* is a constant and n is a constant that is not equal to −1. An antiderivative of x^{n} can be arrived at by incrementing the power by 1 and dividing by the new power. So, the integral of $f(x)$ from *a* to *b* can be expressed as:

$$\int_a^b f(x)dx = \int_a^b cx^n dx = c\int_a^b x^n dx = c\left(\left. \frac{x^{n+1}}{n+1} \right|_z^b \right) = c\left(\frac{b^{n+1}}{n+1} - \frac{a^{n+1}}{n+1}\right)$$