How to Use Calculus Like a Physicist

Physics A300*

Fall 2003

The purpose of these notes is to make contact between the abstract descriptions you may have seen in your calculus classes and the applications of mathematical methods to problems in Physics.

1 Derivatives

1.1 Leibniz Notation

Many math books define the derivative of a function $f(x)$ with the notation

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \tag{1.1}$$

This is perfectly unambiguous, but conceals somewhat the physical nature of a derivative as a ratio of small changes. Thus it's sometimes useful to work with alternative notation, invented by Leibniz, which writes

$$\frac{df}{dx} = f'(x) = \lim_{\Delta x\to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{\Delta x\to 0} \frac{\Delta f}{\Delta x} \tag{1.2}$$

where $\Delta f = f(x+\Delta x) - f(x)$ is the change in $f(x)$ associated with a change in x.

1.2 Working With Differentials

The notation $\frac{df}{dx}$ represents the derivative, *not* the ratio of two numbers df and dx. However, it is often useful to manipulate it as though it were. In particular, we refer to dx as an *infinitesimal* change in x, and $df = f(x+dx)-f(x)$ is the corresponding infinitesimal change in $f(x)$. This is really just a notational convenience, but it's something that makes it a lot easier to do implicit differentiation, and, in multi-variable calculus, to obtain relationships involving partial derivatives.

1.2.1 Example: Differentialting $\ln x$ by Implicit Differentiation

To give an example of how using differentials simplifies things, consider the function

$$y = \ln x \tag{1.3}$$

which is implicitly defined by

$$x = e^y \tag{1.4}$$

If we know that

$$\frac{d}{dy} e^y = e^y \tag{1.5}$$

we can find $\frac{d}{dx} \ln x$ by implicit differentiation. First, we take the differential of both sides of (1.4), and obtain

$$dx = d(e^y) = e^y dy = x dy \tag{1.6}$$

If we allow ourselves manipulate the differentials algebraically as though they were numbers, we can solve for

$$\frac{dy}{dx} = \frac{1}{x} \tag{1.7}$$

which is indeed the derivative of $\ln x$. There is some non-trivial mathematics behind showing that you can treat differentials in this way, but the important thing for us as Physicists is that it not only seems like a reasonable thing to do, it works.

1.3 The Chain Rule

The chain rule as expressed by a mathematician is typically written as follows:

If $f = g \circ h$ is the function defined by the composition of g and h, i.e.,

$$f(x) = g(h(x)) \tag{1.8}$$

then the derivative of f is given in terms of the derivatives of g and h by

$$f'(x) = g'(h(x))\, h'(x) \tag{1.9}$$

Written that way, it's not exactly easy to remember. But let's rewrite it in Leibniz notation. First, for notational simplicity, let's define

$$y = h(x) \tag{1.10a}$$

and

$$z = f(x) = g(h(x)) = g(y) \tag{1.10b}$$

Now in Leibniz notation,

$$\frac{dz}{dx} = f'(x) \tag{1.11a}$$

$$\frac{dy}{dx} = h'(x) \tag{1.11b}$$

$$\frac{dz}{dy} = g'(y) \tag{1.11c}$$

and the chain rule (1.9) becomes

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} \tag{1.12}$$

Again, the calculus-through-algebra outlook tells us that this is really obvious, since one can just cancel out the dys.

Note that our notation as Physicists is also simpler in that we don't need to define separate functions $f(x)$ and $g(y)$; we're usually interested more in the quantity z than in the functional form of its relationship to x or y.

1.4 The Chain Rule in Multi-Variable Calculus

Note that the chain rule is not always as simple as straightforward algebra when functions can depend on more than one variable. Consider the function $w = F(t) = f(x, y, z) = f(g(t), h(t), k(t))$. Written as the function $F(t)$, w depends only on one variable, so we can define an ordinary derivative

$$\frac{dw}{dt} = F'(t) \tag{1.13}$$

while, if we see it as a function $f(x, y, z)$ of three variables, we have to define partial derivatives with respect to all three individually:

$$\frac{\partial w}{\partial x} = f_x(x, y, z) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y, z)}{\Delta x} \tag{1.14a}$$

$$\frac{\partial w}{\partial y} = f_y(x, y, z) = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y, z)}{\Delta y} \tag{1.14b}$$

$$\frac{\partial w}{\partial z} = f_z(x, y, z) = \lim_{\Delta z \to 0} \frac{f(x, y, z + \Delta z)}{\Delta z} \tag{1.14c}$$

Of course $x = g(t)$, $y = h(t)$, and $z = k(t)$ are all functions of one variable, so they have ordinary derivatives

$$\frac{dx}{dt} = g'(t) \tag{1.15a}$$

$$\frac{dy}{dt} = h'(t) \tag{1.15b}$$

$$\frac{dz}{dt} = k'(t) \tag{1.15c}$$

[The rest of this section still needs to be written, but the expression for the chain rule is

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} \tag{1.16}$$

]

2 Solving Ordinary First-Order Differential Equations By Integration

Often in Physics, one has an expression for the first time derivative of a quantity, along with its value at some initial time, and needs to find its value at all times. So for example, you may know

$$\dot{x}(t) = at \tag{2.1a}$$

$$x(0) = x_0 \tag{2.1b}$$

One standard way to solve this problem is to take the indefinite integral of $\dot{x}(t)$ to find the value of $x(t)$ up to an additive constant, then use the initial condition to set that constant. So for example in the problem stated in (2.1), we would find

$$x(t) = \int \dot{x}(t)\, dt = \frac{1}{2}at^2 + C \tag{2.2}$$

and then set the constant C via

$$x_0 = x(0) = C \tag{2.3}$$

which tells us

$$x(t) = x_0 + \frac{1}{2}at^2 \tag{2.4}$$

On the other hand, we could also solve the problem in one step with a definite integral, as follows. The Fundamental Theorems of Calculus tell us that for any t_1 and t_2,

$$\int_{t_1}^{t_2} \dot{x}(t)\, dt = x(t_2) - x(t_1) \tag{2.5}$$

or

$$x(t_2) = x(t_1) + \int_{t_1}^{t_2} \dot{x}(t)\, dt \tag{2.6}$$

If we set t_1 to 0 (the time at which the initial condition is known), everything on the right-hand side is given in the statement of the problem, e.g., (2.1), so we can integrate to find $x(t_2)$ for any t_2. We'd like to rename t_2 to t so that we can state the answer as $x(t)$, but we should be a little careful, because the integration variable is already called t, and it's a no-no to have the same variable appear in the same expression as the integration variable of a definite integral and also outside the integral (or in its limits). So we rename the integration variable from t to t', which lets us rename t_2 to t, and obtain the general expression

$$x(t) = x(0) + \int_0^t \dot{x}(t')\, dt' \tag{2.7}$$

in the example problem, this becomes

$$x(t) = x_0 + \int_0^t at'\, dt' = x_0 + \frac{1}{2}at'^2 \Big|_0^t = x_0 + \frac{1}{2}at^2 \tag{2.8}$$

as before.

Both methods are valid, but the indefinite integral method saves you from having to define a lot of temporary integration constants.