You have probably noticed that I like to do calculus using *differentials*. Differentials and the related *differential forms* are often used in applications, especially (but not only) to physics. The official textbook originally covers differentials incompletely and only in one minor application; it then uses them again for integration, at least as a notational convenience. But they are useful for much more. Now is the time to explain what they are.

Notation and terminology

If y is a variable quantity, then $\mathrm{d}y$ is the **differential** of y. You can think of $\mathrm{d}y$ as indicating an infinitely small (infinitesimal) change in the value of y, or the amount by which y changes when an infinitesimal change is made. A precise definition is at the end of these notes, but you are *not* responsible for knowing that; what you need to know is how to *use* differentials.

Note that $\mathrm{d}y$ is *not* d times y, and d is also *not* exactly a function of y. Rather, y (being a *variable* quantity) should itself be a function of some other quantity x, and $\mathrm{d}y$ is also a function of a sort; so d is an *operator*: something that turns one function into another function. (However, an expression like $A\,\mathrm{d}y$ does involve multiplication: it is A times the differential of y.)

We often divide one differential by another; for example, $\mathrm{d}y/\mathrm{d}x$ is the result of dividing the differential of y by the differential of x. The textbook introduces this notation early to stand for the *derivative* of y with respect to x, and indeed it is that; but what the book doesn't tell you is that $\mathrm{d}y/\mathrm{d}x$ literally is $\mathrm{d}y$ divided by $\mathrm{d}x$. (Unfortunately, $\mathrm{d}^2y/\mathrm{d}x^2$, the second derivative, is *not* literally d^2y divided by $\mathrm{d}x^2$, at least not in any generally useful way that I know.)

Differentials and the Chain Rule

One sometimes sees the Chain Rule expressed as

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u} \cdot \frac{\mathrm{d}u}{\mathrm{d}x}, \tag{1}$$

but the Chain Rule is a nontrivial fact that cannot be proved by simply cancelling factors. I prefer to state the Chain Rule as

$$\mathrm{d}f(u) = f'(u)\,\mathrm{d}u; \tag{2}$$

the point is that the *same* function f' appears regardless of which argument u we use.

Even this is more abstract than how the Chain Rule is applied. For example, suppose that you have discovered (say from the definition as a limit) that the derivative of $f(x) = \sin x$ is $f'(x) = \cos x$. Since $f'(x)$ may be defined as $\mathrm{d}f(x)/\mathrm{d}x$, this derivative can be expressed in differential form without even bothering to name the functions involved:

$$\mathrm{d}(\sin x) = \cos x\,\mathrm{d}x. \tag{3}$$

Once you know this, you know something even more general:

$$\mathrm{d}(\sin u) = \cos u \,\mathrm{d}u \tag{4}$$

for any other differentiable quantity u; the power to derive equation (4) from equation (3) is the Chain Rule. Thus, using $u = x^2$ (to continue the example),

$$\mathrm{d}\big(\sin(x^2)\big) = \cos(x^2)\,\mathrm{d}(x^2) = \cos(x^2)(2x\,\mathrm{d}x) = 2x\cos(x^2)\,\mathrm{d}x.$$

You may divide both sides of this equation by $\mathrm{d}x$ if you wish, but the basic calculation involves only rules for differentials.

For the record, here are the rules for differentiation using differentials that you should know:

- The Constant Rule: $\mathrm{d}k = 0$ if k is constant.
- The Sum Rule: $\mathrm{d}(u+v) = \mathrm{d}u + \mathrm{d}v$.
- The Translate Rule: $\mathrm{d}(u+k) = \mathrm{d}u$ if k is constant.
- The Difference Rule: $\mathrm{d}(u-v) = \mathrm{d}u - \mathrm{d}v$.
- The Product Rule: $\mathrm{d}(uv) = v\,\mathrm{d}u + u\,\mathrm{d}v$.
- The Multiple Rule: $\mathrm{d}(ku) = k\,\mathrm{d}u$ if k is constant.
- The Quotient Rule: $\mathrm{d}\left(\frac{u}{v}\right) = \frac{v\,\mathrm{d}u - u\,\mathrm{d}v}{v^2}$.
- The Power Rule: $\mathrm{d}\big(u^k\big) = ku^{k-1}\,\mathrm{d}u$ if k is constant.
- The Exponentiation Rule: $\mathrm{d}(k^u) = k^u \ln k \,\mathrm{d}u$ if k is constant.
- The Logarithm Rule: $\mathrm{d}(\log_k u) = \frac{\mathrm{d}u}{u \ln k}$ if k is constant.
- The Sine Rule: $\mathrm{d}(\sin u) = \cos u\,\mathrm{d}u$.
- The Cosine Rule: $\mathrm{d}(\cos u) = -\sin u\,\mathrm{d}u$.
- The Tangent Rule: $\mathrm{d}(\tan u) = \sec^2 u\,\mathrm{d}u$.
- The Cotangent Rule: $\mathrm{d}(\cot u) = -\csc^2 u\,\mathrm{d}u$.
- The Secant Rule: $\mathrm{d}(\sec u) = \tan u \sec u\,\mathrm{d}u$.
- The Cosecant Rule: $\mathrm{d}(\csc u) = -\cot u \csc u\,\mathrm{d}u$.
- The Arcsine Rule: $\mathrm{d}(\arcsin u) = \frac{\mathrm{d}u}{\sqrt{1-u^2}}$.
- The Arccosine Rule: $\mathrm{d}(\arccos u) = -\frac{\mathrm{d}u}{\sqrt{1-u^2}}$.
- The Arctangent Rule: $\mathrm{d}(\arctan u) = \frac{\mathrm{d}u}{u^2+1}$.
- The Arccotangent Rule: $\mathrm{d}(\operatorname{arccot} u) = -\frac{\mathrm{d}u}{u^2+1}$.
- The Arcsecant Rule: $\mathrm{d}(\operatorname{arcsec} u) = \frac{\mathrm{d}u}{|u|\sqrt{u^2-1}}$.
- The Arccosecant Rule: $\mathrm{d}(\operatorname{arccsc} u) = -\frac{\mathrm{d}u}{|u|\sqrt{u^2-1}}$.
- The First Fundamental Theorem of Calculus: $\mathrm{d}\Big(\int_u^v f(t)\,\mathrm{d}t\Big) = f(v)\,\mathrm{d}v - f(u)\,\mathrm{d}u$.

The last one might not be familiar to you in that form, but it's also very handy.

Partial derivatives

Notice that every one of the rules above turns the differential on the left into a sum of terms (possibly only one term, or none in the case of the Constant Rule), each of which is an ordinary expression multiplied by a differential (or something algebraically equivalent to this). Such an expression is called a differential *form*, or more precisely a **differential 1-form**. (If, when you are calculating the differential of an expression, your result at any stage is *not* like this, then you have made a mistake!)

Now apply this to a function of several variables. If $f(x, y, z)$ can be expressed using the operations in the list above (and possibly even if it cannot), then its differential will come out as

$$\mathrm{d}f(x,y,z) = f_1(x,y,z)\,\mathrm{d}x + f_2(x,y,z)\,\mathrm{d}y + f_3(x,y,z)\,\mathrm{d}z \tag{5}$$

for some functions f_1, f_2, and f_3. These functions are the **partial derivatives** of f. (Since subscripts can be used for many things, another notation for these is $D_i f$ for f_i.) For example, if $f(x, y, z) = x^2y + \sin(z^2)$, then

$$\mathrm{d}f(x,y,z) = 2xy\,\mathrm{d}x + x^2\,\mathrm{d}y + 2z\cos(z^2)\,\mathrm{d}z,$$

so

$$\begin{aligned} f_1(x,y,z) &= 2xy, \\ f_2(x,y,z) &= x^2, \text{ and} \\ f_3(x,y,z) &= 2z\cos(z^2). \end{aligned}$$

If instead we write u for $f(x, y, z)$, then we have a different notation for the coefficients on the differentials:

$$\mathrm{d}u = \left(\frac{\partial u}{\partial x}\right)_{y,z}\mathrm{d}x + \left(\frac{\partial u}{\partial y}\right)_{x,z}\mathrm{d}y + \left(\frac{\partial u}{\partial z}\right)_{x,y}\mathrm{d}z.$$

So for example, if $u = x^2y + \sin(z^2)$, then

$$\mathrm{d}u = 2xy\,\mathrm{d}x + x^2\,\mathrm{d}y + 2z\cos(z^2)\,\mathrm{d}z,$$

so

$$\begin{aligned} \left(\frac{\partial u}{\partial x}\right)_{y,z} &= 2xy, \\ \left(\frac{\partial u}{\partial y}\right)_{x,z} &= x^2, \text{ and} \\ \left(\frac{\partial u}{\partial z}\right)_{x,y} &= 2z\cos(z^2). \end{aligned}$$

You read $(\partial u/\partial x)_{y,z}$ as 'the partial derivative of u with respect to x, fixing y and z', which means ‹whatever comes before $\mathrm{d}x$ in an expansion of $\mathrm{d}u$ in the variables x, y, and z›. All of this information is necessary to avoid ambiguity, although in practice people usually write simply $\partial u/\partial x$ and expect you to guess from context what the other variables are.

Of course, people also mix notation for f with notation for u, writing $D_x f$, f_x, $\partial f/\partial x$, and so on.

Differential forms

Besides the differential 1-forms discussed above, there are more general differential forms with terms involving products of differentials. There are several kinds, but the simplest are the **exterior differential forms**; here is a typical one:

$$x^2\,\mathrm{d}x \wedge \mathrm{d}y + xy\,\mathrm{d}x \wedge \mathrm{d}z - 3xyz\,\mathrm{d}y \wedge \mathrm{d}z.$$

The wedge '$\wedge$' for multiplication is there to indicate that these are *exterior* differential forms; and this form of multipliciation is like the cross product of vectors in that it is skew-commutative:

$$\mathrm{d}y \wedge \mathrm{d}x = -\mathrm{d}x \wedge \mathrm{d}y.$$

This is a more advanced subject, but it explains many otherwise strange formulas that we will need towards the end of this course.

Appendix: Definitions

Since d is an operator, it must be applied to a function. Suppose that f is a function of n variables; I'll write $f(\vec{x})$ instead of $f(x, y, \ldots)$, since n could be any number (well, any whole number). Then define the **differential** $\mathrm{d}f$ of f to be a function of $2n$ variables; but write $\left\langle \mathrm{d}f(\vec{x}) \middle| \vec{h} \right\rangle$ instead of $\mathrm{d}f\left(\vec{x}, \vec{h}\right)$ for the value of this function at $\vec{x}$ and $\vec{h}$. Its definition is this:

$$\left\langle \mathrm{d}f(\vec{x}) \middle| \vec{h} \right\rangle = \lim_{t \to 0} \frac{f\left(\vec{x} + t\vec{h}\right) - f(\vec{x})}{t}.$$

We say that f is **differentiable** at $\vec{x}$ if $\left\langle \mathrm{d}f(\vec{x}) \middle| \vec{h} \right\rangle$ is defined for every possible value of h.

Now, I have been applying d to variables like x and y and to expressions built out of them. So in order to make sense of this, I must have been tacitly assuming that these expressions are functions of some quantity or quantities. If all of the quantities in an application of calculus may be expressed as functions of n quantities $\vec{x}$, then we pick these and call them the **independent variables**. So, if u is any of these quantities, then we have $u = f(\vec{x})$ for some function f (possibly a constant function or an unknown function, but still in principle some function). These independent variables do *not* have to be anything that appears directly in any calculation (because of the Chain Rule); for these definitions to work, it is only necessary that *some* choice of indepedent variables is possible.

If $u = f(\vec{x})$, then when we write $\mathrm{d}u$, we simply mean $\mathrm{d}f(\vec{x})$, where $\mathrm{d}f$ is the function of $2n$ variables defined above. This leaves $\mathrm{d}u$ as a function of only n variables, the variables written as $\vec{h}$ above. Technically, a differential form such as $3\,\mathrm{d}x + 2y\,\mathrm{d}z$ is also a function of the n variables $\vec{h}$, defined by

$$\langle 3\,\mathrm{d}x + 2y\,\mathrm{d}z | h \rangle = 3\langle \mathrm{d}x | h \rangle + 2y\langle \mathrm{d}z | h \rangle.$$

In all of our applications of differentials, when we write equations between differential forms, we are really writing equations between functions of $\vec{h}$. However, we never bother to apply these to any particular argument. It is in this way that all of the rules for manipulating differentials become true theorems.