

Differentials

A word about notation: As I remarked earlier (on page 3 of the main notes), when $y = f(x)$, we can write $dy/dx = f'(x)$; both sides of the latter equation are notation for a *derivative*, which is one of the things that differentiation produces. The left-hand side means the derivative of y with respect to x , while f' in the right-hand side is a function which is the derivative of the original function f . To say that the derivative of f is f' suggests that the derivative is a basic concept, not a combination of anything more complicated, and that is how the textbook approaches derivatives. But the left-hand side suggests that a derivative is a ratio, the result of dividing dy by dx , and this is how they were originally used. As for dy and dx themselves, they are the *differentials* of y and x ; a differential is another thing that differentiation produces.

1 Differentials

Many calculations in calculus are easier to do using *differentials*. Furthermore, differentials and the related *differential forms* are often used in applications, especially (but not only) to physics. The official textbook covers differentials (in Section 3.11), but incompletely and only in one minor application. It then uses differentials again later (mostly in material for Calculus 2 and 3), but they are useful much earlier. So I will make heavy use of them.

If x is a variable quantity, then dx is the **differential** of x . You can think of dx as indicating an infinitely small (infinitesimal) change in the value of x , or (better) the amount by which x changes when an infinitesimal change is made (an infinitely small change in the value of the independent variable t). A precise definition is in the next section, but you will *not* be tested directly on that; what you need to know is how to *use* differentials.

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

We often divide one differential by another; for example, dy/dx 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 dy/dx literally is dy divided by dx . Unfortunately, d^2y/dx^2 , the second derivative of y with respect to x , is *not* literally $d^2y = d(dy)$ divided by $dx^2 = (dx)^2$; for this reason, I prefer the notation $(d/dx)^2y$, meaning $(d/dx)(d/dx)y = (d/dx)(dy/dx) = d(dy/dx)/dx$ for the second derivative.

2 The Chain Rule

The most important fact about differentials is this: If f is a differentiable function, then

$$d(f(u)) = f'(u) du.$$

That is, the differential of $f(u)$ equals $f'(u)$ times the differential of u , where f' is the derivative of the function f . This fact not only shows the relationship between differentials and derivatives, but also (because u could be any quantity) it encapsulates the **Chain Rule** in differential form. The Chain Rule is an important principle in calculus, which is often difficult to learn how to use; but with differentials it is easy.

In particular, if $y = f(x)$, then

$$\frac{dy}{dx} = \frac{d(f(x))}{dx} = \frac{f'(x) dx}{dx} = f'(x),$$

so dy divided by dx really is the derivative.

For a better example, suppose that you have discovered (say from the definition as a limit) that the derivative of $f(x) = x^2$ is $f'(x) = 2x$. Then this fact can be expressed in differential form:

$$d(x^2) = 2x dx \tag{*}$$

(because $d(x^2) = d(f(x)) = f'(x) dx = 2x dx$). Conversely, if (by performing a calculation with differentials) you discover the equation (*) above, then you know the derivative of f as well:

$$f'(x) = \frac{d(f(x))}{dx} = \frac{d(x^2)}{dx} = \frac{2x dx}{dx} = 2x.$$

Whichever of these facts you discover first, once you know them, you know something even more general:

$$d(u^2) = 2u du.$$

(The power to derive this from equation (*) is the Chain Rule.) The value of this is that u can be any expression whatsoever; for example, if $u = x^2$ again, then

$$d(x^4) = d((x^2)^2) = 2(x^2) d(x^2) = 2x^2(2x dx) = 4x^3 dx.$$

So now you have learnt a new derivative, without having to calculate it from scratch.

3 Rules of differentiation

Every theorem about derivatives of functions may also be expressed as a theorem about differentials. Here are the most common rules:

- The Constant Rule: $d(K) = 0$ if K is constant.
- The Sum Rule: $d(u + v) = du + dv$.
- The Translate Rule: $d(u + C) = du$ if C is constant.
- The Difference Rule: $d(u - v) = du - dv$.
- The Product Rule: $d(uv) = v du + u dv$.
- The Multiple Rule: $d(ku) = k du$ if k is constant.
- The Quotient Rule: $d\left(\frac{u}{v}\right) = \frac{v du - u dv}{v^2}$.
- The Power Rule: $d(u^n) = nu^{n-1} du$ if n is constant.
- The Root Rule: $d(\sqrt[m]{u}) = \frac{\sqrt[m]{u} du}{mu}$ if m is constant.

Of these, only the Constant Rule, the Sum Rule, the Product Rule, and the Power Rule are absolutely necessary, since every other expression built out of the operations in the rules above can be built out of the operations in these four rules. However, it is often handy to use all of these rules; it is up to you how many of these rules to learn. (The Power Rule given here really corresponds to the *Generalized* Power Rule in the textbook, because it incorporates the Chain Rule within it. The Root Rule is not in the textbook, because a root can be algebraically transformed into a power; but the version here rationalizes the denominator, which can be convenient.)

In addition, every time that you learn the derivative of a new function, you learn a new rule for differentials, by applying the Chain Rule to that function. I already showed you an example of this earlier in these notes: applying the Chain Rule to the function $f(x) = x^2$ gives the special case of the Power Rule for $n = 2$. Here are a few other functions whose derivatives you will learn, expressed as rules for differentials:

- $d(\exp u) = \exp u du$.
- $d(\ln u) = \frac{du}{u}$.
- $d(\sin u) = \cos u du$.
- $d(\cos u) = -\sin u du$.
- $d(\operatorname{atan} u) = \frac{du}{u^2 + 1}$.

And more! (To be clear, $\exp u$ means e^u , $\ln u = \log_e u$, u is in radians in $\sin u$ and $\cos u$, and $\operatorname{atan} u$ is what is also written $\arctan u$, $\operatorname{Tan}^{-1} u$, or $\tan^{-1} u$ and gives a result in radians.)

Notice that every one of these rules 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). An expression like this is called a **differential form** (although actually there are more general sorts of differential forms). 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!

4 Defining differentials

To formally define what differentials are and prove their properties, I'll make the same assumption that I made at the beginning of these notes, that there is an independent variable t that every other variable is a function of. Then, I said that if $u = f(t)$, then $u|_{t=c} = f(c)$. Now I'll say that, if $u = f(t)$ and the function f is differentiable, then

$$du|_{t=c, dt=h} = f'(c) h.$$

More generally, if $u = f(t)$ and $v = g(t)$, then

$$(u dv)|_{t=c, dt=h} = f(c) g'(c) h.$$

Again, this is abstract, but the concrete application is straightforward; for example:

$$\begin{aligned} (2x dx + 3 dx)|_{x=4, dx=0.05} &= 2(4)(0.05) + 3(0.05) = 0.55, \\ (2x dx + 3y dy)|_{x=4, y=5, dx=0.05, dy=0.02} &= 2(4)(0.05) + 3(5)(0.02) = 0.7. \end{aligned}$$

(I've put small numbers in for dx and dy , because this is most often what comes up in practice, although for theoretical purposes it doesn't matter.) It's now more common to be given only partial information; for example:

$$\begin{aligned} (2x dx + 3 dx)|_{x=4} &= 2(4) dx + 3 dx = 11 dx, \\ (2x dx + 3y dy)|_{x=4, y=5} &= 2(4) dx + 3(5) dy = 8 dx + 15 dy. \end{aligned}$$

Notice that you *don't* plug in the values of x and y inside the differential operator d ; if you're not given values of dx and dy , then those differentials must remain in the answer.

While expressions like the above come up occasionally (see the discussion of linear approximation later on), the main purpose of a precise definition is to prove theorems. (That's how we can be sure that the rules of Calculus will always work, at least when the definitions that prove them can be made to apply.) Earlier I gave a list of rules for differentials; we can prove these using the precise definition of differential and the known rules for derivatives of functions. For example, if $u = f(t)$ and $v = g(t)$, then $uv = f(t)g(t) = (fg)(t)$. Therefore,

$$d(uv)|_{t=c, dt=h} = (fg)'(c) h = (f'(c)g(c) + f(c)g'(c)) h = g(c)f'(c)h + f(c)g'(c)h = (v du + u dv)|_{t=c, dt=h}.$$

Here, I've used the formal definition of differential along with the Product Rule for derivatives of functions. The conclusion is that $d(uv)$ and $v du + u dv$ always evaluate to the same result, so

$$d(uv) = v du + u dv,$$

which is the Product Rule for differentials. In the same way, all of the rules for differentials follow from rules for derivatives of functions.

The Chain Rule is an important special case, so I'll prove it too. If $u = g(t)$ and f is any function, then $f(u) = f(g(t)) = (f \circ g)(t)$, so if f is differentiable, then

$$d(f(u)) \Big|_{\substack{t=c, \\ dt=h}} = d((f \circ g)(t)) \Big|_{\substack{t=c, \\ dt=h}} = (f \circ g)'(c) h = f'(g(c)) g'(c) h = (f'(u) du) \Big|_{\substack{t=c, \\ dt=h}}.$$

Again, I used the definition of differential and the Chain Rule for functions, and my conclusion is the Chain Rule for differentials:

$$d(f(u)) = f'(u) du$$

whenever f is a differentiable function.

It's not really essential to assume that there exists a *single* independent variable that every other variable is a function of, and we'll stop making that assumption in Calculus 3 (if you stick around that long). Then the formal definition will become a little trickier, but all of the rules for differentials will continue to apply exactly as I stated them above.

5 Using differentials

The main technique for using differentials is simply to take the differential of both sides of an equation. However, you may only do this to an equation that holds *generally*, but *not* to an equation that holds only for *particular* values of the variables. (Ultimately, this is because d is an operator, not a function, so it must be applied to entire functions, not only to particular values of those functions.)

The simplest case is an equation such as $y = \exp(x^2)$, when we want the derivative of y with respect to x . So:

$$\begin{aligned} y &= \exp(x^2); \\ dy &= d(\exp(x^2)) = \exp(x^2) d(x^2) = \exp(x^2) \cdot 2x dx = 2x \exp(x^2) dx; \\ \frac{dy}{dx} &= 2x \exp(x^2). \end{aligned}$$

Now we have the derivative. If we want the second derivative, then we do this again:

$$\begin{aligned} dy/dx &= 2x \exp(x^2); \\ d(dy/dx) &= d(2x \exp(x^2)) = \exp(x^2) d(2x) + 2x d(\exp(x^2)) \\ &= \exp(x^2) \cdot 2 dx + 2x \cdot 2x \exp(x^2) dx = (2 \exp(x^2) + 4x^2 \exp(x^2)) dx; \\ (d/dx)^2 y &= \frac{d(dy/dx)}{dx} = 2 \exp(x^2) + 4x^2 \exp(x^2). \end{aligned}$$

Now we have the second derivative (also written d^2y/dx^2).

The previous example began with an equation solved for y . But we don't need this; suppose instead that we have $y^5 + x^2 = x^5 + y$ (which *cannot* be solved for either variable using the usual algebraic operations of addition, subtraction, multiplication, division, powers, and roots). Undaunted, we forge ahead anyway:

$$\begin{aligned} y^5 + x^2 &= x^5 + y; \\ d(y^5 + x^2) &= d(x^5 + y); \\ d(y^5) + d(x^2) &= d(x^5) + dy; \\ 5y^{5-1} dy + 2x^{2-1} dx &= 5x^{5-1} dx + dy; \\ 5y^4 dy - dy &= 5x^4 dx - 2x dx; \\ (5y^4 - 1) dy &= (5x^4 - 2x) dx; \\ \frac{dy}{dx} &= \frac{5x^4 - 2x}{5y^4 - 1}. \end{aligned}$$

This process is called **implicit differentiation**.

The second derivative is a little more straightforward at first (or it would be if we didn't have to use the Quotient Rule), but there is a twist at the end:

$$\begin{aligned}
 dy/dx &= \frac{5x^4 - 2x}{5y^4 - 1}; \\
 d(dy/dx) &= d\left(\frac{5x^4 - 2x}{5y^4 - 1}\right) = \frac{(5y^4 - 1)d(5x^4 - 2x) - (5x^4 - 2x)d(5y^4 - 1)}{(5y^4 - 1)^2} \\
 &= \frac{(5y^4 - 1)(20x^3 - 2)dx - (5x^4 - 2x)(20y^3)dy}{(5y^4 - 1)^2} \\
 &= \frac{20x^3 - 2}{5y^4 - 1}dx - \frac{20y^3(5x^4 - 2x)}{(5y^4 - 1)^2}dy; \\
 (d/dx)^2y &= \frac{d(dy/dx)}{dx} = \frac{20x^3 - 2}{5y^4 - 1} - \frac{20y^3(5x^4 - 2x)}{(5y^4 - 1)^2} \frac{dy}{dx} \\
 &= \frac{20x^3 - 2}{5y^4 - 1} - \frac{20y^3(5x^4 - 2x)}{(5y^4 - 1)^2} \frac{5x^4 - 2x}{5y^4 - 1}
 \end{aligned}$$

(which could be simplified further). Notice that I substitute the known expression for dy/dx in the last step.

Another handy application of differentials is the case where both quantities x and y may be expressed as functions of some other quantity t . (For the purposes of formal definitions, we always assume that this is possible, but now we're really going to use it.) If we start with the same equation as above, then this will give us an equation relating the derivatives with respect to t :

$$\begin{aligned}
 y^5 + x^2 &= x^5 + y; \\
 d(y^5 + x^2) &= d(x^5 + y); \\
 d(y^5) + d(x^2) &= d(x^5) + dy; \\
 5y^{5-1}dy + 2x^{2-1}dx &= 5x^{5-1}dx + dy; \\
 5y^4 \frac{dy}{dt} + 2x \frac{dx}{dt} &= 5x^4 \frac{dx}{dt} + \frac{dy}{dt}.
 \end{aligned}$$

If we have information about one or both of these derivatives, then this equation will often give us useful information to solve a problem. This situation is called **related rates**, since derivatives can be viewed as rates of change (especially derivatives with respect to time t , although the t in the equation above doesn't have to stand for time).

When we get to integrals, differentials become so useful that even the textbook starts using them, but I'll save that for later.