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, but incompletely and only in one minor application. It then uses differentials again much later (mostly in material for Calculus 2 and 3), but they are useful much earlier. So I will make heavy use of them.

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 general or useful way that I know.)

Fundamental theorem

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

$$\mathrm{d}f(u) = f'(u)\,\mathrm{d}u.$$

That is, the differential of $f(u)$ equals $f'(u)$ times the differential of u, where f' is the derivative of f (as a function). 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

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{f'(x)\,\mathrm{d}x}{\mathrm{d}x} = f'(x),$$

so $\mathrm{d}y$ divided by $\mathrm{d}x$ really is the derivative.

For 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:

$$\mathrm{d}(x^2) = 2x\,\mathrm{d}x. \tag{*}$$

Conversely, if (by performing a calculation with differentials) you discover the equation (*) above, then you know the derivative of f as well:

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

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

$$\mathrm{d}(u^2) = 2u\,\mathrm{d}u.$$

(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

$$\mathrm{d}(x^4) = \mathrm{d}\big((x^2)^2\big) = 2(x^2)\,\mathrm{d}(x^2) = 2x^2(2x\,\mathrm{d}x) = 4x^3\,\mathrm{d}x.$$

So now you have learnt a new derivative.

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: $\mathrm{d}k = 0$ 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 Root Rule: $\mathrm{d}(\sqrt[k]{u}) = \frac{\sqrt[k]{u}\,\mathrm{d}u}{ku}$ if k 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, even the Root Rule (which is not in the textbook). It is up to you how many of these rules to learn.

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. We have already seen an example of this: applying the Chain Rule to the function $f(x) = x^2$ gives the special case of the Power Rule for $k = 2$. Here are a few other functions whose derivatives you will learn, expressed as rules for differentials:

- $\mathrm{d}(\exp u) = \exp u\,\mathrm{d}u$, or $\mathrm{d}(\mathrm{e}^u) = \mathrm{e}^u\,\mathrm{d}u$.
- $\mathrm{d}(\ln u) = \dfrac{\mathrm{d}u}{u}$.
- $\mathrm{d}(\sin u) = \cos u\,\mathrm{d}u$.
- $\mathrm{d}(\cos u) = -\sin u\,\mathrm{d}u$.
- $\mathrm{d}(\arctan u) = \dfrac{\mathrm{d}u}{u^2+1}$.

And more!

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). 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!

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 $y = \mathrm{e}^{x^2}$, when we want the derivative of y with respect to x. So:

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

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

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

Now we have the second derivative (also written $\mathrm{d}^2y/\mathrm{d}x^2$).

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:

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

This process is called **implicit differentiation**.

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

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

(which could be simplified further). Notice that I substitute the known expression for $\mathrm{d}y/\mathrm{d}x$ 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. If we start with the same equation as above, then this will give us an equation relating the derivatives with respect to t:

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

Then if we have information about one or both of these derivatives, then this 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 be time).

Appendix: Definitions and proofs

Since d is an operator, it must be applied to a function. If f is a function, we define the **differential** of f to be a function of *two* variables; but we write $\mathrm{d}f(x)h$ instead of $\mathrm{d}f(x,h)$ for the value of this function at x and h. Its definition is:

$$\mathrm{d}f(x)h = f'(x)h.$$

Notice that the right-hand side is a product, the result of multiplying $f'(x)$ and h, while the left-hand side is *not* a product; it only looks like a product because of the unusual way of writing $\mathrm{d}f$ as a function of two variables. Also notice that $\mathrm{d}f(x)h$ is defined if and only if f is differentiable at x, regardless of the value of h.

Now, we 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, we must be tacitly assuming that these expressions are functions of some quantity. If all of the quantities in an application of calculus may be expressed as functions of one quantity x, then we identify such a quantity x and call it the **independent variable**. (In multivariable calculus, studied in Calculus 3, you may need several independent variables.) So, if y is any of these quantities, then we have $y = f(x)$ for some function f (possibly a constant function or an unknown function, but still in principle some function).

If $y = f(x)$, then when we write $\mathrm{d}y$, we simply mean $\mathrm{d}f(x)$, where $\mathrm{d}f$ is the function of two variables defined above. This leaves $\mathrm{d}y$ as a function of one variable, the variable written as h above. In all of our applications of differentials, when we write equations between differentials, we are really writing equations between functions of h, which we never bother to apply to any particular value. (Any expression that does *not* contain a differential is constant as a function of h.)

In particular, if I is the identity function $I(x) = x$, then $I'(x) = 1$ for all x, so we have $\mathrm{d}I(x)h = h$. So,

$$\mathrm{d}f(x)h = f'(x)h = f'(x)\,\mathrm{d}I(x)h;$$

substituting y for $f(x)$ and x for $I(x)$, we have

$$\mathrm{d}y = f'(x)\,\mathrm{d}x$$

as an equation between functions of h, so

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f'(x)$$

as expected. (That is, $\mathrm{d}y/\mathrm{d}x$ is a constant function of h, whose value is $f'(x)$.)

More generally, if $u = g(x)$, so that $(f \circ g)(x) = f\big(g(x)\big) = f(u)$, then

$$\mathrm{d}(f \circ g)(x)h = (f \circ g)'(x)h = f'\big(g(x)\big)g'(x)h = f'\big(g(x)\big)\,\mathrm{d}g(x)h,$$

where I have used the Chain Rule to find $(f \circ g)'$. Substituting $f(u)$ for $(f \circ g)(x)$ and u for $g(x)$, we have

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

as an equation between functions of h. This is the Chain Rule in differential form. Another way to look at this is that the choice of independent variable is arbitrary; we could use u just as well as x (at least if we work only with quantities that may be expressed as functions of u).

Every other rule for differentials follows from the corresponding rule for derivatives by substituting appropriate expressions and multiplying by $\mathrm{d}x$. For example, from

$$(f + g)'(x) = f'(x) + g'(x),$$

we write u for $f(x)$ and v for $g(x)$, so that $u + v = (f + g)(x)$. Then this rule for derivatives becomes

$$\frac{\mathrm{d}(u + v)}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}x} + \frac{\mathrm{d}v}{\mathrm{d}x},$$

which gives the corresponding rule for differentials:

$$\mathrm{d}(u + v) = \mathrm{d}u + \mathrm{d}v.$$

All of the others may be proved similarly.