Calculus for the Unconvinced

Programmed Workbook

James Pickering

Contents

Introduction

If you're reading this, you're probably a student who is learning calculus for the first time as part of a science degree, but that didn't do much maths before this, and are pretty unsure and unconvinced about this whole thing[1]. It seems like a lot of maths and you're not sure you signed up for that. After all – the adverts for chemistry were all about saving the world with new drugs and energy technology, not tedious algebra. Maybe you've even felt that the amount of maths means that this choice of subject wasn't for you. Hopefully we can change that view – there's no reason why anyone can't learn this stuff.

It's true that maths plays a central role in science, and especially in chemistry, so we do need to learn it. However, it's also true that in schools maths is almost universally taught in a terrible way that makes it seem way more boring and hard than it actually is. They normally focus on the *machinery* behind the maths (like, how to solve quadratic equations), but don't say much about *why* you would actually want to do that (because lots of equations that describe the world around us are quadratic).

In my opinion, once you understand the *substance* behind the equations and the machinery, it all gets easier. Blindly learning how to manipulate algebra isn't useful for anyone. There are times when we will focus on the machinery, especially at first, but we will always try to make it clear *why* this stuff is useful. Understanding that is what will make you an excellent scientist. It's more important to be able to translate between the physical world around us and the mathematical world on paper than it is to be able to do fiddly algebra.

In this workbook, we are going to learn the machinery behind **calculus**, and why it is useful. We're not going to just learn a load of rules (though there will be a bit of this), we're going to keep coming to back to real examples where this is useful to show you how it's a beautiful and elegant way to understand the world around us.

The workbook is structured as a series of steps in each chapter. Each step asks a question that you should answer before the next step. The answer will be given at the top of the next step, so scroll through it step by step and work at your own pace.

[1] Or you're an undergraduate at the University of Leicester who I am using to test this idea for a book. Though that doesn't mean you're not also unconvinced.

Chapter 1

Intuition

To start with, we're going to think about something called a *derivative*. Before we do the algebra-crunching, it's helpful to have an intuition of what a *derivative* physically means and why we should give a shit about calculating it - so let's do that.

1.1 The *derivative* of y with respect to x is written as:

$$\frac{\mathrm{d}y}{\mathrm{d}x}$$

We'll discuss what this means shortly, but firstly - what would the derivative of f with respect to t be written as?

1.2

$$\frac{\mathrm{d}f}{\mathrm{d}t}$$

Easy, right? It doesn't matter what letters we use for our variables. But we'll stick with y and x for a bit.

Anyway, the derivative $\frac{\mathrm{d}y}{\mathrm{d}x}$ tells us about how the variable y changes as we vary x. The $\mathrm{d}x$ represents a small change in x.

What does $\mathrm{d}y$ represent?

1.3 A small change in y. The small d here means *'an infinitesimally small change in'*, if we are being strict. We will just think of it as meaning a small change.

The overall derivative is basically the ratio of these, i.e:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\text{Small change in } y}{\text{Small change in } x}$$

If $\frac{\mathrm{d}y}{\mathrm{d}x} = 1$, what does that imply about dx and dy? Try to multiply both sides of that equation by dx and see what happens.

1.4 We find that $\mathrm{d}y = \mathrm{d}x$. So it implies that a change in x, dx, produces an equal change in y, dy.

If the derivative, $\frac{\mathrm{d}y}{\mathrm{d}x}$, is a big positive number (say, 100), what does that imply about the change in y dy produced by a given change in x, dx?

1.5 It implies that there's a big change in y when we change x, i.e:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 100 \rightarrow \mathrm{d}y = 100\mathrm{d}x$$

So the change in y is 100 times bigger than the change in x. So if I increased x by 2, y would increase by 200.

If the derivative is a small number, (say, 0.01), what does that imply about the change in y dy produced by a given change in x, dx?

1.6 It implies that there'll be a small change in y when we change x, i.e:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 0.01 \rightarrow \mathrm{d}y = 0.01\mathrm{d}x$$

So the change in y is 0.01 times bigger than (or, 100 times *smaller than*) the change in x. So if I increased x by 2, y would only increase by 0.02. Note that y still increases though.

What about if the derivative is zero? What does that imply about the change in y for a given change in x?

1.7 It implies that $\mathrm{d}y = 0$ for a given change in x. So, y doesn't change as we change x. Another way of saying this is that y *is independent of* x.

What about if the derivative, $\frac{\mathrm{d}y}{\mathrm{d}x}$, is a negative number, like -5?

1.8 This implies that y gets smaller as x increases. You can see this as:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -5 \rightarrow \mathrm{d}y = -5\mathrm{d}x$$

So the change in y is -5 times the change in x. So if I increased x by 2, y would only 'increase' by -10, or in other words, y would *decrease* by 10. You can also see this generally by the argument:

$$\frac{\mathrm{d}y}{\mathrm{d}x} < 0 \rightarrow \mathrm{d}y < 0 \text{ if } \mathrm{d}x \text{ is positive}$$

So our change in y must be less than zero, so y get smaller as x gets bigger.

Enough abstract maths. Imagine we have a chemical reaction where some reactant, R, is consumed, and product, P, is produced. Something like:

$$\mathrm{R} \longrightarrow \mathrm{P}$$

If we measure the concentration of P, [P], as the time, t, passes, would the derivative

$$\frac{\mathrm{d[P]}}{\mathrm{d}t}$$

Be positive or negative? Think about what must happen to the concentration of P over time.

1.9 It would be positive, because P is produced as time, t progresses. So both [P] and t get bigger together.

What if we looked at the amount of reactant, R? Does the reactant concentration [R] increase or decrease as time progresses?

1.10 It decreases, because reactant is used up.

So, as time passes, what is the sign of the derivative:

$$\frac{\mathrm{d[R]}}{\mathrm{d}t}$$

Is it positive ([R] increases as t increases), or negative ([R] decreases as t increases)?

1.11 It's negative, because reactant gets used up, so the concentration of reactant, [R], decreases as t gets bigger. These derivatives are ways we measure *rates of chemical reaction*, but more about that later.

This idea is used in many other areas than chemistry. If we are measuring the distance travelled by an object, s, as time t passes, we can think about the derivative:

$$\frac{\mathrm{d}s}{\mathrm{d}t}$$

If the object moves a very large distance in a short time, is this derivative big or small?

1.12 It would be big, because ds is big (large distance) and dt is small (short time).

Is that object moving quickly, or slowly?

1.13 Quickly, right? because it's moving a large distance in a short time.

So, is the speed of the object high, or low?

1.14 High, of course.

Actually, we *define* the speed of the object as the derivative of the distance, s, with respect to time, t:

$$\frac{\mathrm{d}s}{\mathrm{d}t}$$

If the object moves slowly (so only covers a small distance in a certain time), is this derivative big or small?

1.15 Small – because ds would be small for a given change in time dt.

Hopefully you're starting to see that:

a These derivatives tell us about real physical stuff that might be useful.

b It would be handy to have a way to calculate these derivatives, as often we are interested in thinking about how things change in response to other things.

But if you're unconvinced, these aren't even only useful in science. What if we were running a business and wanted to look at the prices of things over time? We could imagine wanting to know about the change of price with respect to time, or:

$$\frac{\mathrm{d}\,\mathrm{Price}}{\mathrm{d}t}$$

If prices increase over time, is this derivative positive or negative?

1.16 Positive, because if prices increase, then d Price is positive, and as time passes then dt is positive – so the derivative:

$$\frac{\mathrm{d}\,\mathrm{Price}}{\mathrm{d}t}$$

is also positive.

This derivative is actually one way of measuring *inflation*, which is a metric that captures how prices change over time. This derivative stuff is everywhere.

1.17 Maybe you're not interested in chemistry, physics, or business (but hopefully you're at least interested in the first one of these). So, here's another example.

Imagine we are doctors monitoring the spread of disease in a society over time. We might be interested in the number of people, N, are infected as time, t, changes.

What derivative could be define to help us think about this change?

1.18

$$\frac{\mathrm{d}N}{\mathrm{d}t}$$

The derivative $\frac{\mathrm{d}N}{\mathrm{d}t}$ would tell us how the number of infected people changes over time. Later on we will talk about *differential equations* and see how we could start from that derivative and then work out the expected number of infected people in the future. Hopefully you can see that would be useful!

As a final example, let's go back to chemistry. Sometimes we'll think about something called the *heat capacity* of an object or material – you probably remember it from A level.

If an object requires a lot of heat (Q) to be added to raise the temperature (T), then it has a *high* heat capacity. What heat capacity does it have if only a small amount of heat is needed to raise the temperature?

1.19 A low heat capacity – i.e. it can't 'absorb' much heat before the temperature starts to rise. It doesn't have much 'capacity' to store that heat before getting hotter.

That description above was in words, but words are long. What if we defined a derivative:

$$\frac{\mathrm{d}Q}{\mathrm{d}T}$$

If we add a fixed amount of heat, Q, to an object and then measure a large increase in temperature, T, is this derivative positive or negative, and large or small?

1.20 Positive and small. If there's a large increase in temperature ($\mathrm{d}T$ is big) for a given amount of heat added $\mathrm{d}Q$, then $\frac{\mathrm{d}Q}{\mathrm{d}T}$ is going to be small (large denominator in the fraction). But it will be still be positive, the temperature doesn't go *down* as we add the heat.

What if we add a fixed amount of heat to an object and measure a *small* increase in temperature? Is the derivative $\frac{\mathrm{d}Q}{\mathrm{d}T}$ large or small?

1.21 Large – because now $\mathrm{d}T$ is small for a given $\mathrm{d}Q$. Dividing by a smaller number in the derivative gives a larger number.

You may have noticed that this is similar to our definition of the heat capacity. In fact, *the* definition of the heat capacity, C, is:

$$C = \frac{\mathrm{d}Q}{\mathrm{d}T}$$

If an object has a high heat capacity, is the derivative above large or small?

1.22 Large.

If the heat capacity is high, and we add a certain amount of heat $\mathrm{d}Q$, is the change in temperature $\mathrm{d}T$ large or small?

1.23 Small. We can see this mathematically, if we don't like words:

$$C = \frac{\mathrm{d}Q}{\mathrm{d}T} \rightarrow C\mathrm{d}T = \mathrm{d}Q \rightarrow \mathrm{d}T = \frac{\mathrm{d}Q}{C}$$

So if the amount of heat added dQ is fixed, and the heat capacity C is high, then $\frac{\mathrm{d}Q}{C}$ is small and so dT is small. Strictly, multiplying and dividing by derivatives like this isn't really allowed, but it's useful and we're not anally retentive mathematicians.

Hopefully you are seeing that these *derivatives* are useful things that tell us about how things change. In this course we are going to learn how to calculate them, and, as we go along, learn how we can use them in our science.

To finish, if the derivative $\frac{\mathrm{d}y}{\mathrm{d}x}$ is large, what does that mean about the relationship between x and y again?

1.24 It means that a given change in x, dx produces a big change in y, dy.

1.25 Summary:

- The derivative of y with respect to x is written as $\frac{\mathrm{d}y}{\mathrm{d}x}$.
- If the derivative is big, it means that there's a large change in y for a given change in x and vice versa.
- Physically, derivatives represent the *change* in one quantity as another quantity is varied. They can be used to represent reaction rates, motion of objects, inflation, and many other useful concepts.

Next time we will start learning how to calculate derivatives.

Chapter 2

Basics

Now let's learn how to calculate some derivatives of basic functions. At first we are just going to learn the rules so we can get to using our derivatives for things, but later on we will talk about where these rules come from. Think of it as learning how to drive the car first, then worrying about how it all works.

The process of calculating derivatives of functions is called *differentiation*. So we *differentiate* functions to find their *derivatives*.

Remember that when we differentiate one variable with respect to another, what you are doing is finding out the *rate of change of the first variable with respect to the other*. Always remember that the maths we are doing has a physical meaning – it is not maths for the sake of maths!

2.26 Let's start with simple functions like:

$$y = x^2$$

Functions like this, that just contain powers of the independent variable (here x), are called *polynomials*.

The derivative of the function above is:

$$\frac{dy}{dx} = 2x$$

The steps for calculating this are:

1. Look at the power (in this case we had x^2 so the power is 2). Multiply the function by the power (so we get $2x^2$)
2. Then, reduce the power by 1, so we end up with $2x^{2-1} = 2x^1 = 2x$.

What would the derivative of $y = x^3$ be?

2.27

$$\frac{dy}{dx} = 3x^2$$

Remember we multiply our function by the power (3), and then reduce the power by one (from 3 to 2).

What if I wanted to find out the derivative of the function:

$$f = t^2$$

with respect to t? What would the derivative $\frac{df}{dt}$ be?

2.28

$$\frac{df}{dt} = 2t$$

Can you see that this was just the same as the example in the first box, but using different letters? It doesn't matter what letters we use to represent the variables – the rules are the same.

What would the derivative of $y = 2x^2$ with respect to x be? Remember the same rule applies.

2.29

$$\frac{dy}{dx} = 4x$$

We get there by multiplying by the power (to give $2x^2 \times 2 = 4x^2$), and then reducing the power by one (to give $4x$).

What about the derivative of $y = 3x^{-1}$? Remember it's the same rule, even with a negative power.

2.30

$$\frac{dy}{dx} = -3x^{-2}$$

Same rules:

1. Multiply by the power: $3x^{-1} \times -1 = -3x^{-1}$
2. Reduce the power by one: $-3x^{-1-1} = -3x^{-2}$

We just need to remember that rule and we can differentiate any polynomial.

What would the derivative of $y = 8x$ be?

2.31

$$\frac{dy}{dx} = 8$$

To see why this works, we just follow the rules again:

1. Multiply by the power (which is one, $x^1 = x$): $8x \times 1 = 8x$
2. Reduce the power by one: $8x^{1-1} = 8x^0 = 8 \times 1 = 8$

Because anything to the power of zero is one.

How about the derivative of $y = 4$? Remember that because $x^0 = 1$ you can think of this as $y = 4x^0$ and then just apply the normal rule.

2.32

$$\frac{dy}{dx} = 0$$

Because in the first step we multiply by the power, which is zero, so: $4x^0 \times 0 = 0$. In fact, the derivative of any *constant term* (anything that doesn't depend on the variable being differentiated by), will always be zero.

Before we continue, evaluate derivatives of the following functions with respect to x:

$y = 4x^5$ $\qquad$ $y = 8\pi$ $\qquad$ $y = -6x^{-0.5}$

2.33

$\frac{dy}{dx} = 20x^4$ $\frac{dy}{dx} = 0$ $\frac{dy}{dx} = 3x^{-1.5}$

If you didn't get all of those, then have a look back at the previous frames before continuing.

We've done a lot of abstract 'machinery' learning so far, so let's apply it to something tangible – an example from sport. You might be aware of a phenomenon where a ball that is kicked or thrown over a distance (L) actually moves sideways as it travels. It's called 'swing' in cricket, and 'curl' or 'bend' in football. The amount that the ball swings or bends, S, is related to L by:

$$S = AL^2$$

Where A is a bundle of constants that relate to things like ball size and air density. What is the derivative of S with respect to L?

2.34

$$\frac{dS}{dL} = 2AL$$

Just by applying our normal rule: $AL^2 \to 2AL^{2-1} \to 2AL$.

So, if the ball travels for a longer distance (larger L), will there more or less swing, S?

2.35 It will swing more, because $\frac{dS}{dL}$ is proportional to L, so if L gets larger, the derivative gets larger, and we get a bigger change in swing dS for a given travel distance dL.

Anyway – that's all fun, but let's keep working to expand the number of types of functions we can differentiate. If we have to differentiate a polynomial containing a few terms like this:

$$y = x^2 + 2x + 1$$

Then it's as easy as just differentiating each term one-by-one, so:

$$\frac{dy}{dx} = 2x + 2 + 0 = 2x + 2$$

What is the derivative of $y = x^3 + 2x^4 + 3x$?

2.36

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 3x^2 + 8x^3 + 3$$

We just apply the rule to each term in the sequence, easy!

What is the derivative of $f = At + B + t^8$, where A and B are constants? Note now we will differentiate with respect to t.

2.37

$$\frac{\mathrm{d}f}{\mathrm{d}t} = A + 0 + 8t^7 = A + 8t^7$$

Same rules – just different symbols. Remember that the exact symbols we use don't matter, the process is always the same.

Of course, there are many other functions in mathematics that we'll use often when describing the world around us. It'd be helpful to be able to differentiate those. Let's start with one very useful one, the derivative of:

$$y = \mathrm{e}^x$$

The derivative of this is actually:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^x$$

So, e^x is the derivative of itself. That's neat, and makes it a very useful function for later on once we start talking about *differential equations*.

Can you work out what the derivative of $y = \mathrm{e}^x + x$ would be?

2.38

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^x + 1$$

Again, we just differentiate every term of the expression one-by-one.

If we have a *constant coefficient* in front of our equation, like:

$$y = A\mathrm{e}^x$$

Where A is a constant, then that constant will just sit there while we do the derivative, so:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = A\mathrm{e}^x$$

What is the derivative of $y = 5e^x$?

2.39

$$\frac{dy}{dx} = 5e^x$$

Other functions we might want to differentiate are the trigonometric functions, $\sin(x)$ and $\cos(x)$:

$$y = \sin(x) \rightarrow \frac{dy}{dx} = \cos(x)$$
$$y = \cos(x) \rightarrow \frac{dy}{dx} = -\sin(x)$$

So, in words, *'sine becomes cosine, and cosine becomes minus sine'*.

What is the derivative of $y = \sin(x) + \cos(x)$?

2.40

$$\frac{dy}{dx} = \cos(x) + (-\sin(x)) = \cos(x) - \sin(x)$$

Again just differentiating each term separately.

What about the derivative of $f = -\sin(t)$?

2.41

$$\frac{df}{dt} = -\cos(t)$$

Remembering that the exact letters we use don't matter, the process is the same. The negative sign is also really just a constant coefficient (-1), so stays at the front.

What about the derivative of $y = -\cos(x)$?

2.42

$$\frac{dy}{dx} = -(-\sin(x)) = \sin(x)$$

Because the two minus signs multiply to give a positive $(-1 \times -1 = 1)$.

For any of these functions we've just looked at, if we have a constant coefficient *inside* the function, like:

$$y = e^{3x}$$
$$y = \sin(\pi x)$$

Then when we differentiate the function that constant gets pulled out to the front of the derivative, due to something called the *chain rule*, which we'll learn about next time. So:

$$y = e^{3x} \rightarrow \frac{dy}{dx} = 3e^{3x}$$
$$y = \sin(\pi x) \rightarrow \frac{dy}{dx} = \pi \cos(\pi x)$$

In the second example, remember that we still differentiate the sine to a cosine.

What is the derivative of $y = \cos(4x)$?

2.43

$$\frac{dy}{dx} = -4\sin(4x)$$

How about the derivative of $y = e^{ax} + \sin(bx)$, where a and b are constants?

2.44

$$\frac{dy}{dx} = ae^{ax} + b\cos(bx)$$

Phew, lots of new stuff. We will just learn the derivative of one more function which will be useful, the derivative of the natural logarithm:

$$y = \ln(x) \rightarrow \frac{dy}{dx} = \frac{1}{x}$$

A bit weird, but that's what it is. We will look at why another time.

What's the derivative of $y = \ln(x) + e^{2x}$?

2.45

$$\frac{dy}{dx} = \frac{1}{x} + 2e^{2x}$$

To bring this back to the real world, imagine we have a chemical reaction where the amount of product, P, produced over time, t, is given by:

$$P = P_0 e^{kt}$$

Where P_0 is the initial amount of product, and k is a *rate constant*. What is an expression for the rate of change of P with time?

2.46

$$\frac{\mathrm{d}P}{\mathrm{d}t} = kP_0\mathrm{e}^{kt}$$

Remembering that P_0 is a constant coefficient and stays at the front. e^{kt} differentiates using the normal rule seen above, becoming $k\mathrm{e}^{kt}$.

Given that $P = P_0\mathrm{e}^{kt}$, how could you rewrite that derivative above so it's just in terms of k and P?

2.47

$$\frac{\mathrm{d}P}{\mathrm{d}t} = kP$$

This is actually an expression for the *rate of the reaction* producing P. The rate of a chemical reaction is just the rate of change of the amount of product over time. If k is a big number, is the reaction fast or slow?

2.48 Fast, because if k is big then $\frac{\mathrm{d}P}{\mathrm{d}t}$ is big, and the rate of reaction is big, so the reaction is fast. Remember that when we do derivatives in science, they *always* represent something physical and tangible.

To finish with, let's practice what we've learnt here. Make sure you can get the answers to the questions below correct before moving on – if you get stuck, have another look at the frames in this chapter.

Differentiate the following functions. You may assume anything other than y and x is a constant.

i $y = 4x^2 + 2x$

ii $y = \sin(x) + 2$

iii $y = \cos(3x) + \sin(x)$

iv $y = \mathrm{e}^{2x} + x^{-2}$

v $y = \ln(x) - x^3 - \sin(x)$

vi $y = \cos(Ax) + 4$

2.49

i $\frac{dy}{dx} = 8x + 2$

ii $\frac{dy}{dx} = \cos(x)$

iii $\frac{dy}{dx} = -3\sin(3x) + \cos(x)$

iv $\frac{dy}{dx} = 2e^{2x} - 2x^{-3}$

v $\frac{dy}{dx} = \frac{1}{x} - 3x^2 - \cos(x)$

vi $\frac{dy}{dx} = -A\sin(Ax)$

If you didn't get those all quite right, have another look through this chapter before continuing to the next.

In the next few chapters we are going to keep working to expand the range of functions we can differentiate, so that when we come to use them in anger we are armed and ready. We'll start by looking at a rule called the *product rule*.

Chapter 3

The Product Rule

In the last chapter we learnt how to differentiate the basic set of functions we need. Now we are going to learn a rule called the **product rule** that will help us differentiate a larger range of functions, and be helpful when we start to solve differential equations.

3.50 As a motivating example, we're all familiar with atomic orbitals, and probably with something called a *radial distribution function* (RDF), F, which tells you the probability of finding the electron in the orbital at a distance r from the nucleus of the atom. For a 1s electron in a hydrogen atom, the equation that describes the RDF is:

$$F = r^2 e^{-2r}$$

It would be good to be able to differentiate this equation, as the derivative of F will eventually let us work out the place where the electron is most likely to be found, and start to explain the chemistry of the system.

The function F is a product of two individual functions: r^2 and e^{-2r}. Can you differentiate these functions on their own?

3.51 We can after last time, and find they differentiate to $2r$ and $-2e^{-2r}$.

So, the question is, *what is the derivative of F?* Unfortunately, it isn't as simple as just differentiating each function and multiplying the result. Instead, we have to use something called the **product rule**. The product rule says:

$$\text{For } F = g(r) \times h(r) \rightarrow \frac{\mathrm{d}F}{\mathrm{d}r} = g\frac{\mathrm{d}h}{\mathrm{d}r} + h\frac{\mathrm{d}g}{\mathrm{d}r}$$

In words, that is that the derivative of the product of two things is *'the first thing times the derivative of the second thing, plus the second thing times the derivative of the first thing'.* We've used F, g, and h here to represent the parts of our function, but the symbols don't matter.

We'll explore the reason this rule works shortly, but for now, can we differentiate our original RDF F using the product rule? Remember we had:

$$F = r^2\mathrm{e}^{-2r} = r^2 \times \mathrm{e}^{-2r} = g(r) \times h(r)$$

Can we use this, and the definition of the rule above to work out the derivative, $\frac{\mathrm{d}F}{\mathrm{d}r}$?

3.52

$$\frac{\mathrm{d}F}{\mathrm{d}r} = -2r^2\mathrm{e}^{-2r} + 2r\mathrm{e}^{-2r}$$

Let's go through this step by step. First we had to define what our functions g and h were, and we said:

$$g(r) = r^2 \text{ (the first function)}$$
$$h(r) = \mathrm{e}^{-2r} \text{ (the second function)}$$

Then we had to work out the derivative of these two functions individually. We did this just now, and found that:

$$\frac{\mathrm{d}g}{\mathrm{d}r} = 2r \text{ (derivative of the first function)}$$
$$\frac{\mathrm{d}h}{\mathrm{d}r} = -2\mathrm{e}^{-2r} \text{ (derivative of the second function)}$$

Then we just had to combine these all together as the product rule states:

$$\frac{\mathrm{d}F}{\mathrm{d}r} = g\frac{\mathrm{d}h}{\mathrm{d}r} + h\frac{\mathrm{d}g}{\mathrm{d}r} \rightarrow \frac{\mathrm{d}F}{\mathrm{d}r} = r^2(-2r\mathrm{e}^{-2r}) + \mathrm{e}^{-2r}(2r)$$

Then we just tidied up the expression a bit to find:

$$\frac{\mathrm{d}F}{\mathrm{d}r} = -2r^2\mathrm{e}^{-2r} + 2r\mathrm{e}^{-2r} = 2\mathrm{e}^{-2r}(r - r^2)$$

If you can't be arsed to tidy things up, that's absolutely fine. Although sometimes writing expressions in different forms helps you see them in different and interesting lights.

Anyway, can you now differentiate the function:

$$y = x^2 e^{-2x}$$

using the product rule?

3.53

$$\frac{dy}{dx} = -2x^2 e^{-2x} + 2x e^{-2x}$$

Hopefully you can see that that was the same example as above, just with different letters. I know I keep banging on about it, but *the letters we use don't matter!*

Let's do a new example. Let's use the product rule to differentiate the function:

$$y = x^3 \sin(x)$$

First we need to work out how to split it into two parts that you *can* differentiate individually, then apply the product rule. What are those two parts? Or, what are the two parts of our product?

3.54 x^3 and $\sin(x)$. We could write our original function as:

$$y = x^3 \times \sin(x) = f \times g$$

Where $f = x^3$ and $g = \sin(x)$.

Great. So, we have our two parts of our product, now we need to differentiate them. Can you differentiate each function f and g?

3.55 You should find that $\frac{df}{dx} = 3x^2$ and $\frac{dg}{dx} = \cos(x)$.

So we now have the four things we need: the two functions making up our original product, and their derivatives. Can you combine them using the product rule to get our final derivative, $\frac{dy}{dx}$?

3.56

$$\frac{dy}{dx} = x^3 \cos(x) + 3x^2 \sin(x)$$

Let's go through this step-by-step again. Split our original function into two:

$$y = x^3 \sin(x) = g(x) \times h(x) \rightarrow g = x^3 \text{ and } h = \sin(x)$$

Work out both derivatives individually:

$$\frac{dg}{dx} = 3x^2 \text{ and } \frac{dh}{dx} = \cos(x)$$

Combine according to the rule:

$$\frac{dy}{dx} = x^3(\cos(x)) + \sin(x)(3x^2) = x^3 \cos(x) + 3x^2 \sin(x)$$

Easy! Let's do some more. How about the derivative of:

$$y = \sin(x)\cos(x)$$

3.57

$$\frac{dy}{dx} = -\sin^2(x) + \cos^2(x)$$

Hopefully you're getting a feel for how this works. When we see something that's a product of two functions, we can use the product rule to differentiate it. The steps to this are:

1. Identify how to split the function into a product of two functions that you can differentiate individually.
2. Differentiate them individually, and combine them according to the product rule.

Calculate the derivative of

$$y = x \ln(x)$$

3.58

$$\frac{dy}{dx} = 1 + \ln(x)$$

Because, following the rule, we'd find:

$$\frac{dy}{dx} = x \times \frac{1}{x} + \ln(x) \times 1$$

We could even use the product rule on things that don't *really* require it, just to check that it works. Think about the function $y = x^2$. We can write this as:

$$y = x \times x$$

What is the derivative of this, using the product rule?

3.59

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x \times 1 + x \times 1 = 2x$$

Just as we would predict without the rule.

Anyway, using the product rule is all fine and fun, but let's try and develop some intuition for where it comes from. In my experience, if you ask a group of students new to calculus how they think they should differentiate a product of two functions, many of them think you just differentiate both parts and multiply them together, like this:

$$y = f(x) \times g(x) \rightarrow \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}x}\frac{\mathrm{d}g}{\mathrm{d}x}$$

This feels logical, but it's easy to see why this doesn't work. Try to apply that rule to our previous example, the derivative of $y = x^2$ (writing x^2 as $x \times x$).

3.60 $\frac{\mathrm{d}y}{\mathrm{d}x} = 1 \times 1 = 1.$

So that would imply that the derivative of $y = x^2$ is 1, which we know is false from our previous rules. Of course, you might argue that those rules could be wrong, so let's look at it another way. Derivatives represent *rates of change*, so if we imagine that $\mathrm{d}y$ represents a change in area (units of metres squared) and $\mathrm{d}x$ represents a change in time (units of seconds), what are the units of the derivative $\frac{\mathrm{d}y}{\mathrm{d}x}$?

3.61 They'll be area over time, or *metres squared per second*.

Now imagine that we can express y as a product of two other functions f and g, where the units of f and g are distance (in metres), and test our incorrect version of the product rule:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}x}\frac{\mathrm{d}g}{\mathrm{d}x}$$

Then what are the units of both sides of the equation?

3.62 The LHS has units of *metres squared per second*, but the RHS has units of *metres squared per second squared*. The equation is **dimensionally inconsistent** because the units are different on each side, and so it cannot be true.

Alright, you probably now believe that the *'do each derivative and multiply them together'* idea is wrong. But that doesn't help us understand why the actual correct version of the product rule is correct. To do this, let's think about having a function, y, and making a small change to it. What symbol do we use to represent a small change in a function y?

3.63 $\mathrm{d}y$, remember that our derivatives, such as $\frac{\mathrm{d}y}{\mathrm{d}x}$, represent the ratio of two small changes.

Imagine we now have our original function, y, and we change it by the small amount $\mathrm{d}y$. What is the value of the changed function?

3.64 $y + \mathrm{d}y$. We just add that small change to our function.

Now imagine that $y = f \times g$, a product of two functions. If we change y by some amount $\mathrm{d}y$, these will also have changed by some amounts $\mathrm{d}f$ and $\mathrm{d}g$.

Given that $y = fg$, How could you write down an equation for $y + \mathrm{d}y$?

3.65

$$y + \mathrm{d}y = (f + \mathrm{d}f)(g + \mathrm{d}g)$$

We just swap each individual function in $y = fg$ for the corresponding function plus a small change.

Can you expand the bracket $(f + \mathrm{d}f)(g + \mathrm{d}g)$ above?

3.66 You should find:

$$fg + f\mathrm{d}g + g\mathrm{d}f + \mathrm{d}g\mathrm{d}f$$

Which means that:

$$y + \mathrm{d}y = fg + f\mathrm{d}g + g\mathrm{d}f + \mathrm{d}g\mathrm{d}f$$

Given that $y = fg$ from our original function, can you cancel something from the equation above to get an expression for dy?

3.67

$$\mathrm{d}y = f\mathrm{d}g + g\mathrm{d}f + \mathrm{d}g\mathrm{d}f$$

Maybe you can see this is starting to look like the product rule a bit, but that dfdg at the end is messing it up. However, if df represents a very (infinitesimally) small change in f, do you think the product dfdg is a big or a small number?

3.68 It's a very small number, in fact a *very very* small number. So small that we can basically say it's roughly equal to zero. The smaller that our changes df and dg get, the product dfdg gets smaller much faster, so we can say that:

$$\mathrm{d}f\mathrm{d}g \simeq 0$$

This approximation becomes exact once we take a real derivative, and we'll see why later.

Anyway, if $\mathrm{d}f\mathrm{d}g \simeq 0$, how can we write our expression for dy from the previous step?

3.69

$$\mathrm{d}y = f\mathrm{d}g + g\mathrm{d}f$$

If we now divide everything by a small change in x, dx, what do we get?

3.70

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f\frac{\mathrm{d}g}{\mathrm{d}x} + g\frac{\mathrm{d}f}{\mathrm{d}x}$$

Which is our original product rule! So that shows us why the expression is what it is. Strictly we can't divide by these small dxs like this, but it works for reasons that we'll talk about later. We'll also see another way to arrive at this result once we start thinking about the *geometrical* interpretation of these derivatives.

We can make a neat rearrangement of the equation above to gain some more intuition for how the rule works. What do you get if you divide the equation above by y?

3.71

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{f}{y}\frac{\mathrm{d}g}{\mathrm{d}x} + \frac{g}{y}\frac{\mathrm{d}f}{\mathrm{d}x}$$

Which, if we remember that $y = fg$, becomes:

$$\frac{1}{y}\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{g}\frac{\mathrm{d}g}{\mathrm{d}x} + \frac{1}{f}\frac{\mathrm{d}f}{\mathrm{d}x}$$

This will be a bit less of a pain to write out if we introduce a contracted notation for the derivative, which is actually a pretty common thing to do. We'll let:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = f'$$

We call the derivative f' (*'f prime'*), and we don't make the variable we are differentiating by (here x) explicit. This shorthand often makes things neater to write down. Can you write down the result at the top of this frame using this notation?

3.72

$$\frac{y'}{y} = \frac{g'}{g} + \frac{f'}{f}$$

This looks quite neat, so let's think about what the fraction $\frac{y'}{y}$ physically means. It represents the *rate of change of y* divided by y, which basically tells you how big the rate of change of y is relative to y itself. If the rate of change of y (y') is much bigger than y, is this fraction large or small?

3.73 Large, because if $y' >> y$, then $\frac{y'}{y} >> 1$.

What about if the rate of change of y is much smaller than y?

3.74 It will be small, because if $y' << y$, then $\frac{y'}{y} << 1$.

Anyway, remember that:

$$\frac{y'}{y} = \frac{g'}{g} + \frac{f'}{f}$$

Is just the product rule but rewritten. So let's see if we can use these fractions to make sense of it. Remember that $y = fg$.

If $\frac{f'}{f} >> \frac{g'}{g}$, what we can say about $\frac{y'}{y}$?

3.75 We can say that:

$$\frac{y'}{y} \simeq \frac{f'}{f}$$

Because if $\frac{f'}{f} >> \frac{g'}{g}$ then $\frac{f'}{f} + \frac{g'}{g} \simeq \frac{f'}{f}$. A massive number plus a tiny number is basically still a massive number.

This makes some intuitive sense too, because what $\frac{f'}{f} >> \frac{g'}{g}$ really says is that the rate of change of f relative to f is much bigger than the rate of change of g relative to g, so it makes sense that the overall rate of change of y relative to y is largely determined by f:

$$\frac{y'}{y} \simeq \frac{f'}{f}$$

The value of g is still important though. Make the substitution $y = fg$ and rearrange the result to get an expression for y' alone.

3.76 You should find:

$$y' \simeq gf'$$

So in the case that $f' >> g'$, the rate of change of y is proportional to f', but is scaled by g. If $g = 0$, then we would find $y' = 0$, which would make sense because then $y = fg = f \times 0 = 0$, and if g was big, then y' would also be big, even if g' was small. This should give you a bit of a feel for why we have these combinations of *'first thing times the derivative of the second thing'* and vice versa in our product rules. .

What happens if $\frac{f'}{f} << \frac{g'}{g}$? Work through the argument again.

3.77 You'd find:

$$\frac{y'}{y} \simeq \frac{g'}{g} \rightarrow y' = fg'$$

As in the previous frame. These ratios $\frac{y'}{y}$ are actually called **logarithmic derivatives** for reasons we might get into later, but I think they help give a bit of an intuition for why the product rule works, by rewriting:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f\frac{\mathrm{d}g}{\mathrm{d}x} + g\frac{\mathrm{d}f}{\mathrm{d}x} \rightarrow \frac{y'}{y} = \frac{f'}{f} + \frac{g'}{g}$$

Anyway, the main thing for now is that we can use the product rule, so to bring it back to actual maths, what is the derivative of:

$$F = te^{-t}$$

3.78 $\frac{dF}{dt} = -te^{-t} + e^{-t} = e^{-t}(1-t)$. Again, the tidying up is optional and very much dependent on the number of shits you give about using the result again later.

So that's the product rule. We will use this a lot in the future, and won't always say *'do this using the product rule'* – you have to be able to spot when to use it.

Before we continue, differentiate the following expressions with respect to x:

i $y = x^4 \sin(x)$

ii $F = 3xe^{3x}$

iii $y = Ax^2 \cos(Bx)$

iv $y = xe^x + x^2 \sin(x)$

3.79 You should get:

i $\frac{dy}{dx} = x^4 \cos(x) + 4x^3 \sin(x)$

ii $\frac{dF}{dx} = 9xe^{3x} + 3e^{3x}$

iii $\frac{dy}{dx} = -ABx^2 \sin(Bx) + 2Ax \cos(Bx)$

iv $\frac{dy}{dx} = xe^x + e^x + x^2 \cos(x) + 2x \sin(x)$

If you didn't get these - have a look back at the previous frames before continuing.

That was the product rule, which we use when we need to differentiate something we can write as the *product* of two functions. Now we are going to learn another rule, called the **chain rule**. With those these two rules, and our basic set of derivatives, we can differentiate anything that science will throw at us (pretty much).

Chapter 4

The Chain Rule

In the last chapter we learnt the **product rule** that enabled us to differentiate products of two functions. Now we are going to learn another rule that will let us differentiate expressions where we have functions inside other functions – or more practically, where have a function that depends on something, where that something depends on something else.

4.80 I think the intuition for the chain rule is simpler and easier than for the product rule, so let's start there. Imagine you are travelling somewhere and have the option to walk, cycle, or drive:

- Cycling is 5× faster than walking.
- Driving is 6× faster than cycling.

How much faster is driving than walking?

4.81 Driving would be $5 \times 6 = 30$ times faster than walking. Congratulations, you intuitively understand the chain rule!

To see how this relates to our derivatives and mathematics, can you remember how we defined speed, v, as a derivative involving distance L and time t?

4.82

$$v = \frac{dL}{dt}$$

Speed is distance divided by time, so in terms of derivatives is a change in distance dL over a change in time dt.

If I travel faster, will I cover a larger or smaller distance dL in a given time dt?

4.83 A larger distance, as my speed being higher means that $\frac{dL}{dt}$ will be higher, so for a given time interval dt I must travel a longer distance dL.

So, thinking about travelling by walking, cycling, or driving, I can define the speeds as:

$$v_{\text{walk}} = \frac{dL_{\text{walk}}}{dt} \text{ and } v_{\text{cycle}} = \frac{dL_{\text{cycle}}}{dt} \text{ and } v_{\text{drive}} = \frac{dL_{\text{drive}}}{dt}$$

Where I travel for the same amount of time dt in each case, but the distance I travel dL depends on whether I walk, drive, or cycle.

Cycling is 5 times faster than walking, can you express this in terms of these derivatives?

4.84 $v_{\text{cycle}} = 5v_{\text{walk}}$, or:

$$\frac{dL_{\text{cycle}}}{dt} = 5\frac{dL_{\text{walk}}}{dt}$$

Driving is 6 times faster than cycling, can you express this in terms of these derivatives?

4.85 $v_{\text{drive}} = 6v_{\text{cycle}}$, or:

$$\frac{dL_{\text{drive}}}{dt} = 6\frac{dL_{\text{cycle}}}{dt}$$

Can you combine these last two results to find an expression for $\frac{dL_{\text{drive}}}{dt}$ in terms of $\frac{dL_{\text{walk}}}{dt}$?

4.86

$$\frac{dL_{\text{drive}}}{dt} = 6\frac{dL_{\text{cycle}}}{dt} = 6 \times 5\frac{dL_{\text{walk}}}{dt} = 30\frac{dL_{\text{walk}}}{dt}$$

Which is what we said at the start: driving is 30 times faster than walking. Can you see that to do this we have basically *chained* the two relationships together? This is the essence of how the **chain rule** works.

If we take the result above and manipulate it a bit, we can get it into the form that the chain rule is usually given in in calculus. We'll divide these derivatives and treat them like fractions, even though that's strictly an abuse of notation (don't tell any mathematicians you know):

$$30 = \frac{\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}t}}{\frac{\mathrm{d}L_{\text{walk}}}{\mathrm{d}t}} = \frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}L_{\text{walk}}} = \frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{walk}}}$$

Can you replace the number 30 in the expression at the top of this frame with this result here? Write down what you get.

4.87

$$\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}t} = \frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{walk}}}\frac{\mathrm{d}L_{\text{walk}}}{\mathrm{d}t}$$

This kind of expression is how **chain rules** are normally written. You can imagine that the $\mathrm{d}L_{\text{walk}}$s will cancel on the RHS of this expression so that the LHS equals the RHS (again, don't tell any mathematicians you know).

Of course, our original question was: *how much faster is driving than walking?*. That obviously relates to the derivative:

$$\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{walk}}}$$

If we didn't know this derivative, but *did* know a derivative that related both walking and driving to a third thing (say, cycling), we could construct a chain rule:

$$\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{walk}}} = \frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{cycle}}}\frac{\mathrm{d}L_{\text{cycle}}}{\mathrm{d}L_{\text{walk}}}$$

We use what we know about the relationship between driving and cycling $\left(\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{cycle}}}\right)$ and the relationship between cycling and walking $\left(\frac{\mathrm{d}L_{\text{cycle}}}{\mathrm{d}L_{\text{walk}}}\right)$ to work out a **chain rule** that will tell us the relationship between driving and walking, $\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{walk}}}$.

So, if cycling is 5 times faster than walking, and driving is 6 times faster than cycling, can we confirm the result from the start of this chapter using a chain rule?

4.88 From the information, we get: $\frac{\mathrm{d}L_{\text{cycle}}}{\mathrm{d}L_{\text{walk}}} = 5$ and $\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{cycle}}} = 6$. Using our chain rule:

$$\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{walk}}} = \frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{cycle}}}\frac{\mathrm{d}L_{\text{cycle}}}{\mathrm{d}L_{\text{walk}}} = 6 \times 5 = 30$$

Which we can then use to say that:

$$\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}t} = 30\frac{\mathrm{d}L_{\text{walk}}}{\mathrm{d}t}$$

Or, 'driving is 30 times faster than walking'.

Let's take stock of what we've done here. The chain rule tells us how to relate two different rates of change. In the example we've gone through above, it tells us how to relate the rate of change of distance and time while driving ($\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}t}$) to the rate of change of distance and time while walking ($\frac{\mathrm{d}L_{\text{walk}}}{\mathrm{d}t}$). We just have to know the ratio between the two of them, which is $\frac{\mathrm{d}L_{\text{drive}}}{\mathrm{d}L_{\text{walk}}}$, and we can learn that from another chain rule.

The trick was to introduce a *third* function that we can use to relate the two things we care about – we cared about the difference in speed between driving and walking. We didn't know what that was, but we *did* know how each of those things related to a third mode of transport: cycling. We used that relationship to construct the chain rule above.

It probably feels like that was a lot of time spent over-analysing something that you already knew, and I can understand that view. So let's see how we can use this chain rule to help us do derivatives.

If we are doing some hardcore spectroscopy, we might be interested in writing down an expression for the electric field, $E(t)$, of our light:

$$E(t) = \sin(2\pi\nu t + \phi)$$

Where ν, and ϕ are constants - we can worry about the physical interpretation later on.

Can you differentiate this expression for $E(t)$ using the rules we have learnt so far?

4.89 No, we can't. The problem is that we have this annoying $+\phi$ term in our sine function. So what we actually have here is like a *function inside* *a function*, like this:

$$E(t) = \sin(2\pi\nu t + \phi) = \sin(g(t))$$

where:

$$g(t) = 2\pi\nu t + \phi$$

Basically, we are introducing a *third function* $g = 2\pi\nu t + \phi$ so we can write:

$$E(g) = \sin(g)$$

Just like we had with our first case, we now have a function g that we can relate to both E and t (as we had with the cycling, that we could relate to both the walking and the driving).

Can you differentiate g (with respect to t) and E (with the substitution made, so with respect to g) on their own?

4.90 Yes we can:

$$\frac{\mathrm{d}E}{\mathrm{d}g} = \cos(g)$$
$$\frac{\mathrm{d}g}{\mathrm{d}t} = 2\pi\nu$$

Just using the normal rules we learnt before. So what is the overall derivative of our original function, $E(t)$?

Well, we are going to use a **chain rule** just like we did before. In the end we want $\frac{\mathrm{d}E}{\mathrm{d}t}$, but we can't differentiate that using the rules we have so far, so we use our introduced third function g to define a **chain rule**:

$$\text{For } E = E(g(t)) \rightarrow \frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}E}{\mathrm{d}g}\frac{\mathrm{d}g}{\mathrm{d}t}$$

Remember that you can get a feel for why this works by imagining that the dg terms cancel each other out in the above, but this isn't strictly an 'allowed' thing to do. The main point is to see that introducing that simplifying third function allowed us to turn *one derivative we couldn't do into two that we can do*. The chain rule tells us how to put them together.

Anyway, for our example:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \cos(g) \times (2\pi\nu) = 2\pi\nu\cos(2\pi\nu t + \phi)$$

Where we got the last bit by undoing our original substitution, so we get the answer just in terms of t.

Can you differentiate:

$$E(x) = \cos(x^2)$$

using the chain rule? Our third function will now be that $g = x^2$, so $E = \cos(g)$. Remember we choose this one because it gives expressions $g(x)$ and $E(g)$ that are easy to differentiate.

4.91 You should find that:

$$\frac{dg}{dx} = 2x \text{ and } \frac{dE}{dg} = -\sin(g)$$

So by the chain rule:

$$\frac{dE}{dx} = \frac{dE}{dg}\frac{dg}{dx} = -2x\sin(g) = -2x\sin(x^2)$$

Let's summarise the steps required to use the chain rule before we practice it a bit. Imagine you've got a complicated looking function that can be expressed as a function inside a function. Let's call that function $F(t)$, so we want the derivative $\frac{dF}{dt}$.

1. Identify a substitution, $g(t)$, that will turn the complicated 'function-inside-a-function' F, into two simpler functions: $F(g)$ and $g(t)$ We need to be able to differentiate both of these functions on their own.
2. Differentiate both of these simpler functions, to find $\frac{dF}{dg}$ and $\frac{dg}{dt}$
3. Multiply them together to find $\frac{dF}{dt}$, as per the chain rule.

Use these steps to differentiate:

$$F(t) = e^{t^3+t}$$

You need to identify the appropriate substitution first!

4.92 You should have found the substitution $g(t) = t^3 + t$. Then, our function becomes $F(g) = e^g$. Our two derivatives are then:

$$\frac{dF}{dg} = e^g \text{ and } \frac{dg}{dt} = 3t^2 + 1$$

Using the chain rule, we multiply these to find:

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}g} \times \frac{\mathrm{d}g}{\mathrm{d}t} = (3t^2 + 1)\mathrm{e}^{t^3+t}$$

To make the link back to our initial intuition with the walking, cycling, and driving clear, there we knew how two things (walking and driving) related to a third thing (cycling). We didn't explicitly know how walking and driving were related, but could use their relationship to cycling to figure it out.

It's the same idea here, if we call driving y and walking x, except that now we *create* our third thing (cycling – call it g) to relate to y and x in a way that makes it easy to find $\frac{\mathrm{d}y}{\mathrm{d}g}$ and $\frac{\mathrm{d}g}{\mathrm{d}x}$. The chain rule then tells us how to find $\frac{\mathrm{d}y}{\mathrm{d}x}$:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}g}\frac{\mathrm{d}g}{\mathrm{d}x}$$

Let's see another example. Peaks in infrared spectra often have a shape, P, that can be modelled using something called a *Gaussian* function:

$$P(\omega) = A\mathrm{e}^{-\omega^2}$$

Where ω is the infrared wavenumber (a variable), and A is a constant.

Can you calculate the derivative $\frac{\mathrm{d}P}{\mathrm{d}\omega}$?

4.93 Make the substitution $g(\omega) = -\omega^2$. Then, our function becomes $P(g) = A\mathrm{e}^g$, like before. Our two derivatives are then:

$$\frac{\mathrm{d}P}{\mathrm{d}g} = A\mathrm{e}^g \text{ and } \frac{\mathrm{d}g}{\mathrm{d}\omega} = -2\omega$$

We now construct and use our chain rule, to find:

$$\frac{\mathrm{d}P}{\mathrm{d}\omega} = \frac{\mathrm{d}P}{\mathrm{d}g} \times \frac{\mathrm{d}g}{\mathrm{d}\omega} = -2\omega\mathrm{e}^{-\omega^2}$$

We could use this derivative to work out where the center of our peak was, and other things. But that's a story for later. For now, let's finish this part on the chain rule with a few examples. Differentiate the following with respect to x:

i $y = \sin(x^3)$

ii $F = e^{3x - x^2}$

iii $y = \cos(Bx + Ax^4)$

iv $y = e^{\cos(x) + x}$

4.94 You should get:

i $\frac{dy}{dx} = 3x^2 \cos(x^3)$

ii $\frac{dF}{dx} = (3 - 2x)e^{3x - x^2}$

iii $\frac{dy}{dx} = -(B + 4Ax^3)\sin(Bx + Ax^4)$

iv $\frac{dy}{dx} = (1 - \sin(x))e^{\cos(x) + x}$

The previous three chapters, on basic principles, the product rule, and the chain rule, will allow us to differentiate most things that science throws our way. So now we're going to start thinking about these derivatives in different ways, and using them to solve some interesting problems.

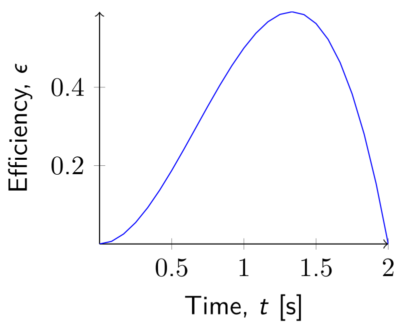
Chapter 5

Visualising Derivatives

We understand now that derivatives are **rates of change**, and are armed to the teeth with methods to calculate them, so it's time to use them for some interesting stuff. To facilitate this, it's helpful to have geometric intuition for what they mean – relating the derivative to the *gradient* of a curve at some point.

Later we'll extend this into a few dimensions and show how it has all kinds of exciting applications. For now, let's stick with functions of one variable, that we can plot in 2D.

5.95 Here's a graph showing how the efficiency, ϵ, of some process varies with time.



The process could be the yield of a chemical reaction, or the output from a factory, the efficiency of an engine, the rate at which a drug is metabolised, or something else. Choose the example that most tickles your fancy.

Roughly (to the nearest 0.5s), when is the process most efficient?

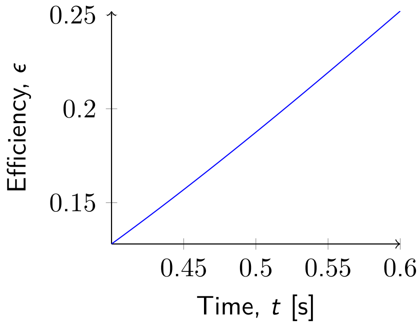
5.96 At around 1.5s, give or take a bit. It'd be good if we had a way to exactly calculate that though, rather than having to plot a graph and reading it off by eye like a humanities student. We'll work our way to figuring out how to do just that by the end of this chapter.

To begin with, if efficiency is ϵ and time is t, what derivative tells us the rate of change of efficiency with time?

5.97

$$\frac{d\epsilon}{dt}$$

Easy. Now, let's look at the graph from earlier, and zoom in on a section of it:



It looks roughly like a straight line over this period. Can you remember (from school) how to find the gradient of this line? What is the method, or equation, for finding the gradient?

5.98 *'Change in y over change in x'*, *'rise over run'*, or something daft like '*down the stairs and along the corridor*' (I was taught that last one in school, and it doesn't make sense however you look at it). The main idea is that you take the change in the thing on the y axis and divide by the change in the thing on the x axis:

$$\text{Gradient} = \frac{\text{Change in } y}{\text{Change in } x} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x}$$

So what is an expression for the gradient using the symbols from our graph above?

5.99

$$\text{Gradient} = \frac{\text{Change in } \epsilon}{\text{Change in } t} = \frac{\Delta \epsilon}{\Delta t}$$

So the gradient of this straight line is $\frac{\Delta \epsilon}{\Delta t}$. If I could actually be arsed to go to the trouble of finding two points, getting $\Delta \epsilon$ and Δt, and dividing them[1], I'd find that the gradient is about 0.6 or so.

If the gradient is 0.6, and I increase t by 0.1, what is the corresponding increase in ϵ?

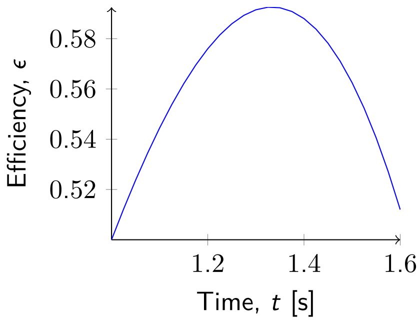
5.100 0.06, because:

$$\frac{\Delta \epsilon}{\Delta t} = 0.6 \rightarrow \Delta \epsilon = 0.6 \Delta t = 0.6 \times 0.1 = 0.06$$

Now, you might notice that we're talking about these $\Delta \epsilon$ things like we talked about our infinitely small changes dy and df from chapter 1. You might, if you parents paid for you to have a pointless and anachronistic education, also know that the Greek letter Δ is the equivalent of an upper case letter d in English. This is not a coincidence.

The gradient we just talked about was of a section of the function that looked like a straight line. What if I zoom on a different part of the function:

[1] And in general, I can't.



What is the gradient of the function here? Is it obvious to define?
