CH1204 Mathematics for Chemistry

Calculus for Chemists

James Pickering

General Information

Calculus Context

Calculus is the branch of maths that deals with understanding **how things change**. Specifically, calculus gives us a mathematical language for talking about **rates of change**. This is very useful for us as chemists, because we are generally dealing with things like:

- **Chemical reactions**: where concentrations of reactants and products *change* over time.
- **Quantum mechanics**: we will often want to talk about how wavefunctions *change* as a function of position or time.
- **Thermodynamics**: we'll often be interested in how heat or energy changes and flows through a system.

Clearly in the first example, if we could use calculus to be able to calculate how fast a reaction will go, and thus know when to stop it for optimum yield, that would be very helpful. However, it's much more widely applicable than just these chemical examples. For instance:

- **Mechanics**: calculating how fast something is moving or accelerating - how it's position or velocity *changes* with time.
- **Biochemistry**: understanding how enzymes affect the *rate* of chemical reactions in cells.
- **Biology**: understanding how populations of animals or microbes *change* over time in response to certain factors.
- **Medicine**: understanding how diseases grow and spread within populations over time (who remembers covid?).
- **Finance**: predicting how stock market prices will change over time.

All of the above examples are places where the behaviour of real-world systems are governed by the mathematics of how things change, which is called **calculus**. What you'll learn in this course will be applicable far beyond chemistry.

Course Structure

The course will be structure in three parts:

1. Differentiation
2. Integration
3. Differential Equations

If you've done A-level maths, parts of the first two topics will be familiar to you, but we will approach things a bit differently to A level, **and there will be a lot of new content**, so you should still come to the workshops if you want to be prepared for the exam. If you've done it before, you can also help your coursemates who haven't with it.

Each week there will be a workshop where I introduce you to some topics in the first 10-15 min and you have a go at some questions for the rest of the time. Any questions you don't finish are your homework, and then later in the week there'll be a lecture where we go through the solutions together. There will also be surgeries every other week where you can get additional help if needed.

This part of the course is assessed via a written exam in summer. It'll be 5 questions, 10 marks each.

How to Study

The lecture notes for this course will be a bit more minimal than in CH1200, as the best way to learn maths is to **do maths** rather than just reading about maths. Generally, the printed notes will contain key ideas, formulae, and some minimal examples. The problem sheets will then flesh these examples out and provide more insight - **the answers you make to the problem sheets are the most important resource you'll have in this course.**

Some textbooks you can look at for extra help are:

- *Maths for Chemistry* - Monk and Munro
- *Foundations of Science Mathematics* - Sivia and Rawlings

Parts of these books go beyond what we will do - the content of the lecture notes and problem sheets determines what is examinable.

The internet is also an excellent place to learn maths, and there are some really excellent people on YouTube who make beautifully explained mathematical content. Particularly highly regarded is **3blue1brown**: `https://www.youtube.com/@3blue1brown`. His video series **The Essence Of Calculus** is essential viewing for this course, and you should all go and watch it immediately.

Chapter 1

Differentiation

To start, let's make sure we know what the **gradient** of a line is. You've seen gradients before in this course, and probably learnt them in school. However have you ever thought about what the gradient is physically telling you, about the line you have plotted? What does it mean if the gradient is large, or small? Remember that in science, when we draw graphs, they **represent something physical** - we're not talking about abstract mathematics here, but real physical, tangible things. Maths just gives us a convenient language to talk about them in.

Imagine I am throwing a ball, and have a graph that represents how far the ball is from my hand at a given time, so it's a graph of position (y-axis) against time (x-axis). If the gradient of that graph is large, it means that for a *small* increase in time (small increase along x) I get a *large* increase in position (large increase along y). Physically, this means that the ball is moving quickly - it has a high **velocity**.

This means we can interpret the gradient as follows:

> **The gradient is telling us how big a change on the y-axis we get for a given change on the x-axis.**

Or alternatively:

> **The gradient is telling us the rate of change of the thing on the y-axis as a function of the thing on the x-axis.**

Always remember that the maths we do has a physical interpretation - we're not interested in maths for the sake of maths[1].

[1] Most of the time at least. I am unapologetically a maths fan-boy.

1.1 Basics: Gradients and Curves

We know how to calculate the gradient of a straight line from school:

- 'Change in y over change in x'
- 'Rise over run'

The main idea is illustrated in Figure 1.1. For a straight line, it doesn't matter

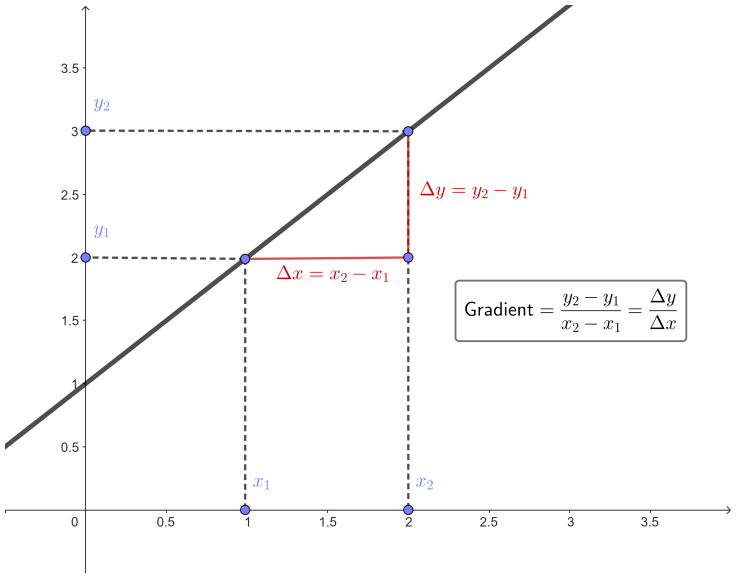

Figure 1.1: Gradient of a straight line.

which two points on the straight line we choose, we'll always calculate the same gradient. **The gradient of a straight line is constant.**

What happens if we wanted to calculate the gradient of a non-linear function, like a curve? One thing we can do is draw **tangent lines** to the function at different points and calculate the gradient of those lines, as shown in Figure 1.2.

Clearly in the example in Figure 1.2, the gradient is different depending on where we draw the tangent. This means that the **gradient is no longer constant**, it varies with x. Measuring the gradient at $x = 1$ gives us a shallower gradient than at $x = 3$. This means that the gradient of this curve is also **a function of** x. This function has a name, and we call it a **derivative**. If we have a function $y = f(x)$, then we can define the **derivative** of this function with respect to x as:

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

or as

$$f'(x)$$

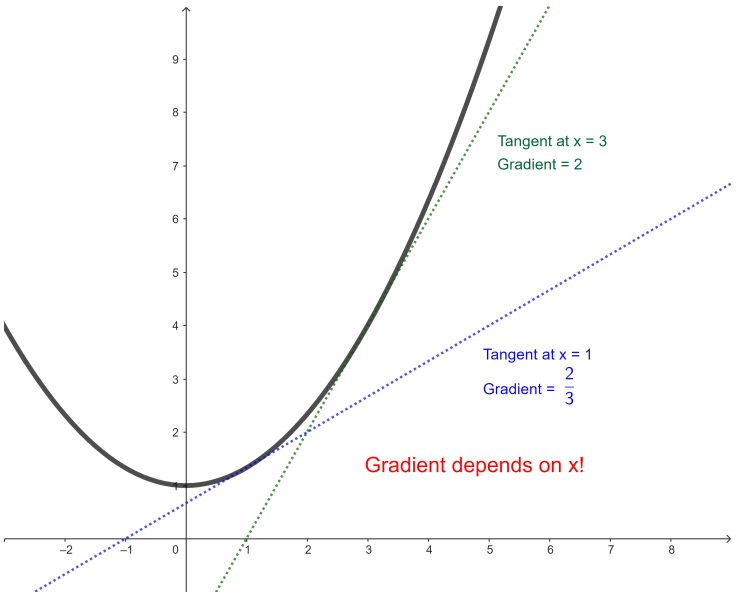

Figure 1.2: Gradient of a non-linear function - depends where we draw the tangent.

I'll use the first of these throughout this course, but it's useful to know both - they mean the same thing. You can remember the notation $\frac{\mathrm{d}y}{\mathrm{d}x}$ by thinking about the definition of the gradient in Figure 1.1 - change in y over change in x.

So, if I can calculate the function $\frac{\mathrm{d}y}{\mathrm{d}x}$, then I have a function I can use to calculate the gradient of my original function $y = f(x)$ at any value of x. So how do I calculate it? Helpfully, clever people (notably Isaac Newton and Gottfried Leibniz) thought about this a few hundred years ago and figured it out. There's an appendix at the end of this document showing how to differentiate 'from first principles' for those of you who are interested, but for the purposes of this course, let's trust that those guys did it correctly.

1.1.1 Calculating Simple Derivatives

There are a set of rules for calculating derivatives of different functions, and we will see all the basic ones we'll need today. **I will expect you to know these**[2]. The process of calculating the derivative of a function is called **differentiating** the function - hence the name of this chapter.

[2]Normally I'm not into just rote learning, as you know, but this is important.

Constants

If our function y is constant C:

$$y = C \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

Think about if this makes sense, what is the gradient of the line $y = 2$?

Powers

If y is equal to x raised to some power n:

$$y = x^n \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = nx^{n-1}$$

'Multiply by the power and reduce the power by one'. This means that the gradient of a quadratic function is a straight line, and that the gradient of a straight line is a constant - can you see why?

Trigonometric Functions

If y is a sine or a cosine function:

$$y = \sin(ax) \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = a\cos(x)$$

$$y = \cos(ax) \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = -a\sin(x)$$

We will see next time why the constant a gets pulled out of the front of the derivative.

Remember that for trig functions, the derivatives go in a kind of cycle:

$$\sin(x) \xrightarrow{\text{Diff.}} \cos(x) \xrightarrow{\text{Diff.}} -\sin(x) \xrightarrow{\text{Diff.}} -\cos(x) \xrightarrow{\text{Diff.}} \sin(x)$$

We can talk about $\tan(x)$ another time.

Exponential Functions

For the exponential function, e^x:

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

Beautiful. This makes calculus with exponential functions incredibly simple, and is why we prefer natural logarithms to base-10 logarithms.

If we have a constant involved:

$$y = \mathrm{e}^{ax} \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = a\mathrm{e}^{ax}$$

Natural Logarithms

For the natural logarithm, $\ln(x)$:

$$y = \ln(x) \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{x}$$

We might come onto why this is later in the course.

1.1.2 Minor Complications

If there is a constant coefficient infront of the function, then just carry it through:

$$y = Ax^2 \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = A \times 2x$$

Where $2x$ is the derivative of x^2 following the rule above for powers.

If the thing to differentiate is a sum of different functions, just differentiate each term separately.

$$y = 2x^2 + 5x + 3 \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = 4x + 5$$

Note that the above does **not** apply if y is a product of different functions, such as $y = x\sin(x)$. We will discuss this next time.

1.1.3 Second Derivatives

Sometimes, we want to differentiate a function again - to find the gradient of a gradient function, or similar. We can just repeat the operation, and this is called taking the **second derivative**. The notation for this is:

$$y = x^2 \xrightarrow{\text{Differentiate Twice}} \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 2$$

You can extend this notation to third, fourth, and higher derivatives.

1.1.4 Context: Equations of Motion

At the start we talked about how if I threw a ball and it moved a large distance in a short time, then it was moving fast (high velocity). If I had a graph of the position of the ball against time, the gradient of this graph would tell me how fast the ball was moving. The gradient is just the derivative, so:

> The derivative of position *with respect to time* is **velocity**.

If we define position as s, velocity as v, and time as t, then:

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

Furthermore, the rate of change of *velocity* with respect to time is called **acceleration**. If acceleration is a:

$$a = \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\mathrm{d}^2s}{\mathrm{d}t^2}$$

So, this will come in handy if we need to work out how fast something is moving - which might be handy in the mechanics of molecular collisions.

You can now attempt the problems on **problem sheet 1**.

1.2 Advanced Differentiation

Now we are going to talk about two more advanced differentiation techniques, which will expand the range of functions you are able to differentiate.

1.2.1 Products of Functions

To differentiate a function which can be expressed as the **product** of two functions, such as $x\cos x$ or $\cos(x)\sin(x)$ we can use something called the **product rule**.

For a function $y = y(x)$ that is the product of two functions $u = u(x)$ and $v = v(x)$, such that $y = uv$:

$$y = uv \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = u\frac{\mathrm{d}v}{\mathrm{d}x} + v\frac{\mathrm{d}u}{\mathrm{d}x}$$

So, you just take the first function and multiply it by the derivative of the second, then take the second function, and multiply it by the derivative of the first, and then add the two results together. For example:

$$y = x\sin(x) \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = x\cos(x) + \sin(x)$$

Because the derivative of x with respect to x is 1 (as $x = x^1$, which becomes $1 \times x^0 = 1$ when differentiated). Another example would be:

$$y = x^2\mathrm{e}^x \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}x} = x^2\mathrm{e}^x + 2x\mathrm{e}^x$$

1.2.2 Functions of Functions

Sometimes we get functions nested inside other functions, such as:

$$y = \cos(3x + 2)$$

To differentiate this, we use something called the **chain rule**. To use the chain rule, we find a substitution that makes the function into something easily differentiable - you do this by intuition and will quickly get the hang of it. In the example above, let's use the substitution $u = 3x + 2$. Then we can write the function as:

$$y = \cos(u) \text{ where } u = 3x + 2$$

We don't have to call the new function u, it can be anything you want. Go wild.

Now we can use **the chain rule** which states that, for a function $y = y(u)$ where $u = u(x)$:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u} \times \frac{\mathrm{d}u}{\mathrm{d}x}$$

To remember which way around the derivatives go, you can look at the derivatives in the equation above and imagine the $\mathrm{d}u$ terms 'cancelling out' like in a fraction so that the RHS is equal to the LHS. However, I must stress that **this is not actually what happens** - the derivative notation merely **looks like a fraction**, but it's not **actually a fraction.** Ok?

All that remains is for us to compute the derivatives and plug them in to the chain rule formula. With our original example:

$$y = \cos(u) \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}y}{\mathrm{d}u} = -\sin(u)$$

$$u = 3x + 2 \xrightarrow{\text{Differentiate}} \frac{\mathrm{d}u}{\mathrm{d}x} = 3$$

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u} \times \frac{\mathrm{d}u}{\mathrm{d}x} = -\sin(u) \times 3 = -3\sin(3x + 2)$$

Where in the last step we undid our original substitution, and replaced u with $3x + 2$.

Let's do another example, with different letters to emphasise that it really doesn't matter what symbols we use for our derivatives. Imagine we have the function $F = F(t)$:

$$F(t) = \mathrm{e}^{4t^2+3t+6}$$

We make a substitution $\theta = 4t^2 + 3t + 6$, and get $F = \mathrm{e}^{\theta}$. We can then construct a chain rule, remembering that ultimately we want to find $\frac{\mathrm{d}F}{\mathrm{d}t}$:

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\mathrm{d}F}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}t}$$

I've dropped the $\times$ symbol for conciseness. Then we calculate:

$$\frac{\mathrm{d}F}{\mathrm{d}\theta} = \mathrm{e}^{\theta} \text{ and } \frac{\mathrm{d}\theta}{\mathrm{d}t} = 8t + 3$$

Putting it all together:

$$\frac{\mathrm{d}F}{\mathrm{d}t} = (8t + 3)\mathrm{e}^{\theta} = (8t + 3)\mathrm{e}^{4t^2+3t+6}$$

Not so bad, right?

1.2.3 Combinations of Chains and Products

Sometimes we will have to do the chain rule more than once, or use the chain rule on one or both of the parts of a product rule. It can get fiddly, but it's not difficult - you just have to be systematic and careful. Let's look at two examples.

Differentiate the function:

$$y = x^3 \sin(4x + \pi)$$

We can use the product rule to find that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x^3 \frac{\mathrm{d}}{\mathrm{d}x}[\sin(4x + \pi)] + 3x^2 \sin(4x + \pi)$$

Where I've used the standard notation $\frac{\mathrm{d}}{\mathrm{d}x}[f(x)]$ to denote that we want to differentiate the thing in square brackets ($f(x)$) with respect to x, but haven't actually worked out the derivative yet[3].

[3] Writing $\frac{\mathrm{d}}{\mathrm{d}x}$ like is is called writing it as a *differential operator*.

To work out the derivative, we use the chain rule, with the substitution $u = 4x + \pi$. Let's also call the bit of the function we want to use the chain rule on $Z = \sin(4x + \pi)$, to tidy up notation. With our substitution, we find:

$$Z = \sin(u) \text{ where } u = 4x + \pi$$

$$\frac{\mathrm{d}Z}{\mathrm{d}x} = \frac{\mathrm{d}Z}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = 4\cos(u) = 4\cos(4x + \pi)$$

Noting that:

$$\frac{\mathrm{d}Z}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}[\sin(4x + \pi)]$$

We can put the result above back into our original product rule formula to find:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 4x^3\cos(4x + \pi) + 3x^2\sin(4x + \pi)$$

Now let's look at a final example, differentiating the function:

$$y = \cos(\mathrm{e}^{6x+4})$$

Here we have a function in a function in a function! We can just use the chain rule twice, and be systematic. Firstly, let's define some substitutions:

$$y = \cos(g) \text{ where } g = \mathrm{e}^h \text{ where } h = 6x + 4$$

I've just broken it down into three functions that we can easily differentiate. Now let's think about chain rules, remembering we ultimately want $\frac{\mathrm{d}y}{\mathrm{d}x}$:

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

To find $\frac{\mathrm{d}y}{\mathrm{d}h}$, we can use another chain rule:

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

So that overall:

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

We can work out these three derivatives from our substitutions:

$$\frac{\mathrm{d}y}{\mathrm{d}g} = -\sin(g) \text{ and } \frac{\mathrm{d}g}{\mathrm{d}h} = \mathrm{e}^h \text{ and } \frac{\mathrm{d}h}{\mathrm{d}x} = 6$$

Combining it all and undoing our substitutions leads to:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -6\mathrm{e}^{6x+4}\sin(\mathrm{e}^{6x+4})$$

Yuk - but hopefully you can see that it's just logical. This is about as tricky as I'd ask in the exam.

1.2.4 Context: Differentiating Wavefunctions

The radial distribution function $R(r)$ for a 1s electron in a hydrogen atom is given by:

$$R(r) = Nr^2\mathrm{e}^{-2r}$$

This function gives a measure of how likely the electron is to be found at a distance r from the nucleus, and N is a constant. The radial distribution function is derived from the wavefunction, and atomic wavefunctions (Ψ) generally tend to have a form like:

$$\Psi(r) = \text{Some polynomial in r} \times \text{Some exponential function in r}$$

So to differentiate wavefunctions, like when we are solving the Schrödinger equation, we often need to use the product rule. If we differentiate the radial distribution function above with respect to r using the product rule, we find:

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

We'll use this result next time.

You can now attempt the problems on **problem sheet 2**.

1.3 Stationary Points

A very common use of differential calculus is to find **stationary points** of a function. To understand what this means, consider the graph in Figure 1.3, which shows a plot of yield of a chemical reaction as a function of reaction progress. Ideally, we want to stop the reaction at a point where the yield is highest - so how can we determine this point?

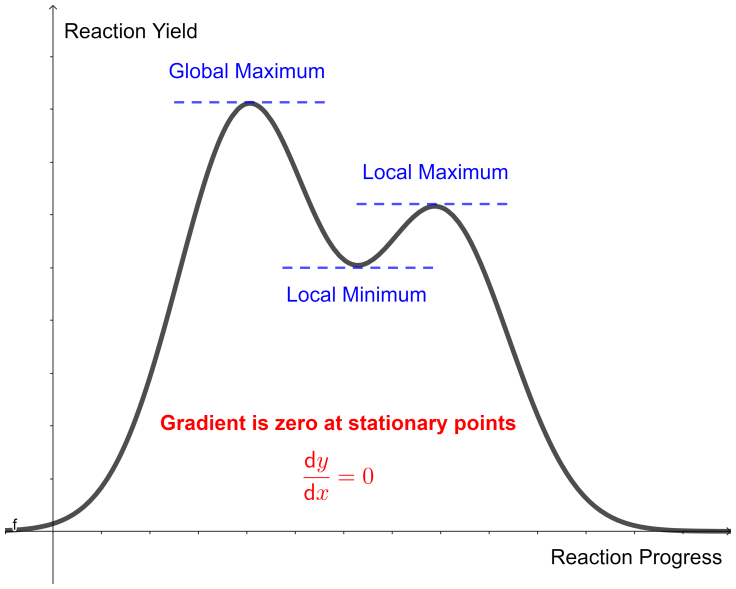

Figure 1.3: Stationary points: maxima, minima, saddle points (see later), occur when the gradient of a function passes through zero.

The key is to notice that at the maximum point, the gradient of the function is zero, because the gradient moves from being positive (on the left hand side of the maximum, where the function is rising) to negative (on the right hand side, where the function is falling). Thus, at the maximum, the gradient must pass through zero. This holds for any kind of stationary point, and we can have minima too - as shown on the diagram. The term **stationary point** simply means a point where the gradient (derivative) of the function is zero. We will see how to classify different types of stationary points soon.

1.3.1 Finding Stationary Points

Finding the stationary points of a function is as simple as setting the derivative of the function equal to zero and solving the resulting equation. For example, say we wanted to find the stationary points of $y = x^2$. First we differentiate:

$$y = x^2 \xrightarrow{\text{Differentiate}} \frac{dy}{dx} = 2x$$

then we set the derivative equal to zero, and solve it for x:

$$\frac{dy}{dx} = 2x = 0 \xrightarrow{\text{Solve}} x = 0$$

For $y = x^2$, there is one stationary point, and it occurs at $x = 0$. We could substitute this value back into our original equation to find the y coordinate, and we'd find the coordinates of the stationary point are (0,0). Does this make sense, thinking about the graph of $y = x^2$?

A more complex example could be to find the stationary points in the function $F(x)$:

$$F(x) = \frac{1}{3}x^3 - \frac{1}{2}x^2 - 6x + 3$$

First we differentiate:

$$\frac{\mathrm{d}F}{\mathrm{d}x} = x^2 - x - 6$$

Then set the equation equal to zero and solve. This time, the resulting equation is quadratic, but we can deal with that by factorisation:

$$\frac{\mathrm{d}F}{\mathrm{d}x} = x^2 - x - 6 = (x-3)(x+2) = 0 \xrightarrow{\text{Solve}} x = 3 \text{ or } x = -2$$

So now we have two stationary points, at $x = 3$ and $x = -2$. How can we know if these stationary points are maxima or minima?

1.3.2 Classifying Stationary Points

One simple way to classify stationary points is to simply look at a graph of the function and do it visually - looking at the example in Figure 1.3, it's clear that (from left to right), we have a maximum, minimum, and another maximum. The highest point of the function is called the **global maximum**, and any other maxima are called **local maxima**. Similarly, we have a **local minimum** - this is not the global minimum, because the function goes to zero, and so the lowest point of this function is lower than the local minimum.

However, it won't always be possible to easily sketch functions, so we can use algebra to classify stationary points too. The trick is to realise:

- At a **maximum**, the gradient changes from **positive** on the left side to **negative** on the right side.
- At a **minimum**, the gradient changes from **negative** on the left side to **positive** on the right side.

So we now need to think about the **rate of change of the gradient**, which requires we use the **second derivative**, as illustrated in Figure 1.4.

- At a **maximum**, the gradient is decreasing (+ve to -ve) so the second derivative is **negative** (or zero).
- At a **minimum**, the gradient is increasing (-ve to +ve) so the second derivative is **positive** (or zero).

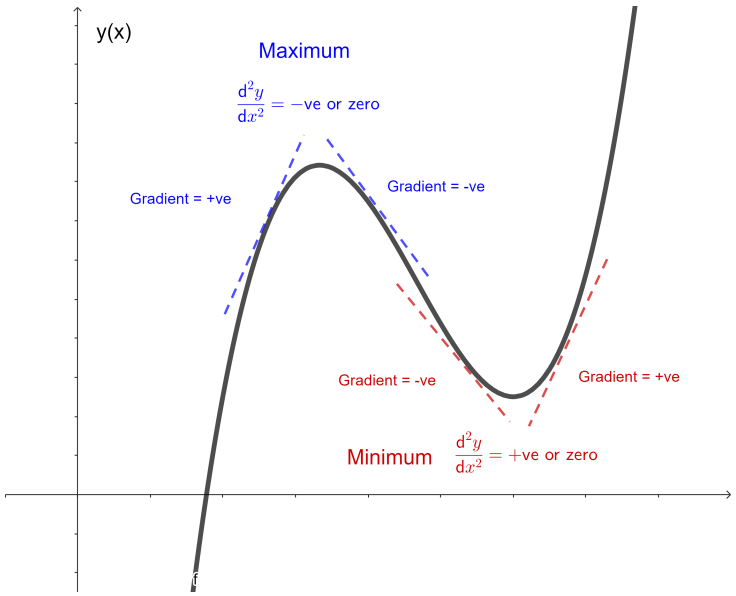

Figure 1.4: Classification of maxima and minima. If the second derivative is zero, further checks are needed.

To illustrate this, let's use the example we had earlier of the function $F(x)$, which had stationary points at $x = 3$ and $x = -2$. To classify these, we first calculate the second derivative:

$$\frac{\mathrm{d}^2F}{\mathrm{d}x^2} = 2x - 1$$

We then plug in each value of x and check the sign of the second derivative at each point:

$$x = 3 \rightarrow \frac{\mathrm{d}^2F}{\mathrm{d}x^2} = 5 \rightarrow \text{ Positive } \therefore \text{ Minimum}$$

$$x = -2 \rightarrow \frac{\mathrm{d}^2F}{\mathrm{d}x^2} = -5 \rightarrow \text{ Negative } \therefore \text{ Maximum}$$

1.3.3 Saddle Points

What happens if the second derivative of a function is zero? In this case, there are three options. The stationary point could be a higher order maximum or a minimum, or it could be a **saddle point**. A saddle point is a stationary point where the gradient has the same sign on both sides of the point. You may have heard this called a **point of inflection** in A level maths, but strictly you can have non-stationary points of inflection, so saddle point is a better term. Saddle points look something like the example in Figure 1.5. To classify a point with zero second derivative, you need to analyse the sign of the gradient on either side of the point:

- For a **maximum**, the gradient will be positive before and negative after the point.

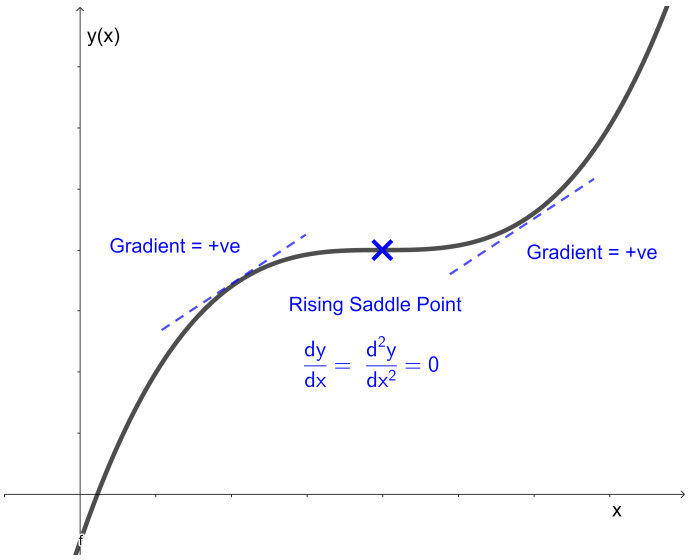

Figure 1.5: A rising saddle point - both the first and second derivatives are zero.

- For a **minimum**, the gradient will be negative before and positive after the point.
- For a **saddle point**, the gradient will have the same sign before and after the point. If the gradients are positive, it is a **rising saddle point**, and a **falling saddle point** if they are negative.

To check the sign of the gradient on each side, simply substitute the x-coordinate either side of the stationary point into your function for the derivative.

1.3.4 Context: Where is the electron?

Previously we had a radial distribution function $R(r)$ for a 1s electron in a hydrogen atom. This function tells us the probability of finding an electron at a distance r from the nucleus. Can we use our new knowledge of stationary points to find out the most likely distance that the electron will be found?

This amounts to finding the position of the maximum in $R(r)$, so we take the derivative and set it equal to zero:

$$\frac{\mathrm{d}R(r)}{\mathrm{d}r} = 2r\mathrm{e}^{-2r} - 2r^2\mathrm{e}^{-2r} = 0$$

Solving this equation:

$$2r - 2r^2 = 0 \rightarrow r = 1$$

So the electron is most likely to be found at a distance of $r = 1$, but one what? The function $R(r)$ is written in **atomic units**, where the standard unit of length is the **Bohr radius**, a_0. Thus, the 1s electron in a hydrogen atom is most likely to be found at $r = a_0$.

You can now attempt the problems on **problem sheet 3**.

1.4 Partial Differentiation

So far we have only considered functions of one variable, like $y = f(x)$, or $F = z(t)$. Lots of useful functions are functions of more than one variable, such as:

$$p = p(V, T) = \frac{nRT}{V}$$

the ideal gas equation that we are familiar with. How do we think about these functions? Unlike a function of one variable, which we can represent with a 2D graph, a function of two variables needs to be represented as a **3D surface**, as in Figure 1.6.

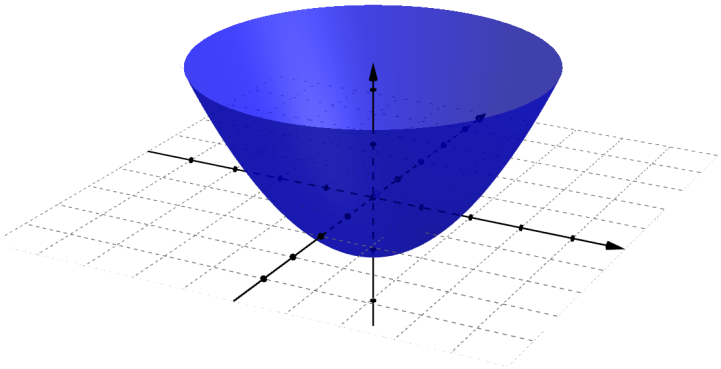

Figure 1.6: An example of a function of two variables, plotted in 3D. This function clearly has a minimum.

Clearly the function shown in Figure 1.6 has a minimum, but how do we find out where it is? To do this, we need to be able to differentiate this function somehow, and one way to do that is to take something called a **partial derivative**.

1.4.1 Partial Derivatives

Take the function $z = z(x, y)$ with two arguments:

$$z(x, y) = x^2 + y^2$$

To take **partial derivatives** of this function, we simply differentiate with respect to each variable, treating the other variable as a constant, and we use some new notation:

$$\left(\frac{\partial z}{\partial x}\right)_y = 2x$$

The notation $\left(\frac{\partial z}{\partial x}\right)_y$ means 'partial derivative of z with respect to x whilst holding y constant'. To calculate the derivative, we just treated every single y as a constant, and differentiated the x terms as normal. We could also calculate the other partial derivative:

$$\left(\frac{\partial z}{\partial y}\right)_x = 2y$$

Note that the partial derivative uses a funny d (∂) rather than the normal d, to differentiate it from the normal derivative. That's all there is to it though - just normal differentiation, but treating one variable as a constant. This amounts to taking a 'slice' through our 3D plot in Figure 1.6, and then calculating the derivative of that slice. We could work out the coordinates of the minimum of our 3D function by setting each of the partial derivatives equal to zero, but we don't do that in this course.

1.4.2 Chemistry Examples

A classic place where this gets used in chemistry is when describing gases, as the pressure of a gas depends on both volume and temperature - $p = p(V, T)$.

Take the ideal gas equation:

$$p = \frac{nRT}{V}$$

We can calculate the two partial derivatives:

$$\left(\frac{\partial p}{\partial T}\right)_V = \frac{nR}{V}\left(\frac{\partial}{\partial T}\right)_V (T) = \frac{nR}{V}$$

by treating V as a constant, and

$$\left(\frac{\partial p}{\partial V}\right)_T = nRT\left(\frac{\partial}{\partial V}\right)_t (V^{-1}) = -\frac{nRT}{V^2}$$

Now treating T as a constant - remember that n and R are constants by definition.

1.4.3 Context: Error Propagation

When analysing experimental errors, we often need to think about how the error in one variable affects our result. For example, we might be calculating the gradient of a straight line, and need to know how the error in y and x combine to give us an error in the gradient m. We do this by **error propagation**. Start with an expression for the thing we want to calculate in terms of the variables, for the gradient example:

$$m = \frac{y - c}{x}$$

Imagine we know the errors in y and x, called Δy and Δx respectively. It turns out[4] that the error in m, Δm is given by:

[4] I can prove this, and we'll see it next year - ask if you're impatient.

$$(\Delta m)^2 = \left(\frac{\partial m}{\partial x}\right)_y^2 (\Delta x)^2 + \left(\frac{\partial m}{\partial y}\right)_x^2 (\Delta y)^2$$

This expression holds for **any** function $m = m(x, y)$, not just the one above. We can understand it as follows:

- The first term is the product of how strongly m depends on x (the derivative) with the size of the error in x.
- The second term is the product of how strongly m depends on y with the size of the error in y.

- The total error in m is then given by the sum of these two - as both the size of the errors in the individual components x and y matter, and how strongly our desired result m depends on x and y. We might have a situation where there is a really large error in x, but actually x doesn't really affect m (small partial derivative), so that error doesn't affect m very much.

We'll see this more next year.

You can now attempt the problems on **problem sheet 4**.

Chapter 2

Integration

We've talked about differentiation, and know that if we have a function like $y = x^2$, we can define the derivative as:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 2x$$

However, what if we only knew the derivative, and not the original function? This might happen if we could only measure the change in something, and not it's actual value (like the speed of a car from a speed camera doesn't allow you to directly work out the position of the car). Is there a way we can 'undo' the derivative and get back our original function, y?

It turns out there is, and this process is called **integration**:

$$y \xrightarrow{\text{Differentiation}} \frac{\mathrm{d}y}{\mathrm{d}x}$$

$$\frac{\mathrm{d}y}{\mathrm{d}x} \xrightarrow{\text{Integration}} y$$

Integration is useful in chemistry, for example because we might know a chemical reaction rate (a derivative) and want to work out the concentration of product at a certain time:

$$\frac{\mathrm{d}[\text{Product}]}{\mathrm{d}t} \xrightarrow{\text{Integrate}} [\text{Product}]$$

2.1 Integration Basics

Initially we think of integration as the reverse of differentiation, and integration operation has a special symbol. The **integral** of a function $f(x)$ is denoted by:

$$\int f(x)\ \mathrm{d}x$$

In this notation, the thing inside the integral sign ($f(x)$) is called the **integrand**.

2.1.1 Indefinite Integration

To evaluate the derivative of a function, we just do the opposite of whatever we did to find the derivative. For example:

$$\int 2x \, \mathrm{d}x = \frac{2}{2}x^2 + C = x^2 + C$$

Let's unpack this:

- To differentiate an expression like x^n, we multiplied by the power and then reduced the power by one. To integrate, we do the opposite - raise the power by one, and then divide by the *new* power.
- We add a constant C to our result. This is called the **constant of integration**.

Why do we add a constant C? Remember any constants in our original expression are lost when we take a derivative, because the derivative of a constant is zero. This means that we need to add on a constant to fully reconstruct our original function when we integrate it. We'll see soon how to determine the value of this constant. Let's see how to integrate some common functions - remembering that it's just the inverse of differentiation that we had before.

Constants

If our function y is constant k:

$$y = k \xrightarrow{\text{Integrate}} \int y \, \mathrm{d}x = kx + C$$

Powers

If y is equal to x raised to some power n:

$$y = x^n \xrightarrow{\text{Integrate}} \int y \, \mathrm{d}x = \frac{x^{n+1}}{n+1} + C$$

The inverse function x^{-1} is special and:

$$y = x^{-1} \xrightarrow{\text{Integrate}} \int y \, \mathrm{d}x = \ln(x) + C$$

Try applying the normal rule to the inverse function and you'll see why it's weird.

Trigonometric Functions

If y is a sine or a cosine function:

$$y = \sin(ax) \xrightarrow{\text{Integrate}} \int y \, \mathrm{d}x = -\frac{1}{a}\cos(x) + C$$

$$y = \cos(ax) \xrightarrow{\text{Integrate}} \int y \, \mathrm{d}x = -\frac{1}{a}\sin(x) + C$$

Note that we divide by the constant, rather that multiplying now.

We just reverse the cycle of derivatives for integration:

$$\sin(x) \xleftarrow{\text{Int.}} \cos(x) \xleftarrow{\text{Int.}} -\sin(x) \xleftarrow{\text{Int.}} -\cos(x) \xleftarrow{\text{Int.}} \sin(x)$$

Exponential Functions

For the exponential function, e^x:

$$y = \mathrm{e}^x \xrightarrow{\text{Integrate}} \int y \, \mathrm{d}x = \mathrm{e}^x + C$$

If we have a constant involved:

$$y = \mathrm{e}^{ax} \xrightarrow{\text{Integrate}} \int y \, \mathrm{d}x = \frac{1}{a}\mathrm{e}^{ax} + C$$

Natural Logarithms

For the natural logarithm, $\ln(x)$:

$$y = \ln(x) \xrightarrow{\text{Integrate}} \int y \, \mathrm{d}x = x \ln x - x + C$$

Remember to always put in the constant of integration when doing an . The same rules regarding constants and expressions that are sums of terms that we saw for derivatives also apply:

$$\int Af(x) \, \mathrm{d}x = A \int f(x) \, \mathrm{d}x \text{ where A is a constant}$$

$$\int f(x) + g(x) + h(x) \, \mathrm{d}x = \int f(x) \, \mathrm{d}x + \int g(x) \, \mathrm{d}x + \int h(x) \, \mathrm{d}x$$

For example, I could simplify the integral below as:

$$\int 3x^2 + 6x + 7 \, \mathrm{d}x = 3 \int x^2 \, \mathrm{d}x + 6 \int x \, \mathrm{d}x + \int 7 \, \mathrm{d}x$$

To determine the constant of integration, we simply need to know a point on our integrated function. We'll see how this works in one of the problems. We call these kinds of integrals **indefinite integrals** because they are not evaluated between specific **limits**. To understand what this means, we have to think about integration graphically.

2.1.2 Definite Integration

We interpreted differentiation graphically as giving us the gradient of non-linear functions. We can think of integration graphically as giving us the **area between the function and the x-axis**. This is illustrated in Figure 2.1.

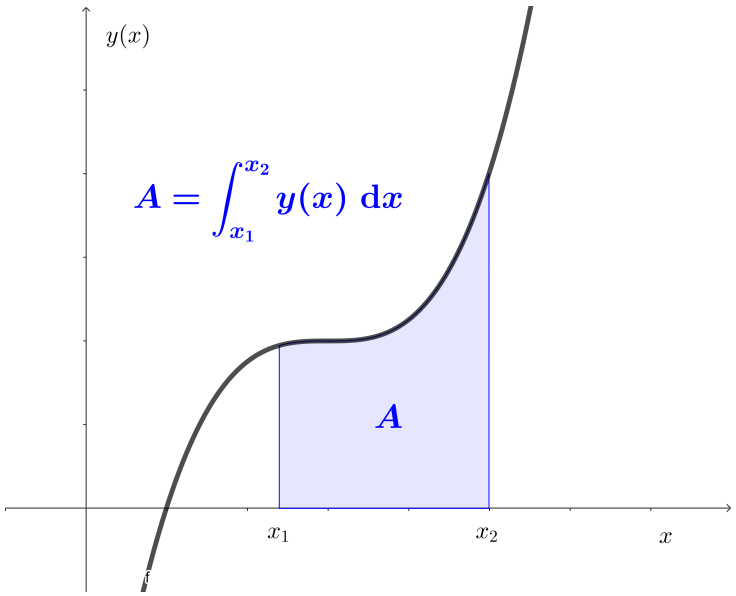

Figure 2.1: The definite integral of a function between x_1 and x_2 is defined as the area between the function and the x axis, evaluated between x_1 and x_2.

Evaluating an integral between two limits to calculate the area between a function an the x-axis is called doing a **definite integral**. To show this works, let's calculate the area between the curve $y = x^2 + 1$ and the x-axis between the limits 0 and 2.

$$\int_0^2 y(x)\ \mathrm{d}x = \int_0^2 x^2 + 1\ \mathrm{d}x$$

We now evaluate the integral as normal, but put the limits around the result. We also don't need a constant of integration when doing definite integrals[1].

[1] To understand why, do this example again leaving the constant of integration in.

$$\int_0^2 x^2 + 1\ \mathrm{d}x = \left[\frac{1}{3}x^3 + x\right]_0^2$$

To evaluate this, we simply evaluate the expression in the square bracket at the first limit, then at the second limit, and subtract them from each other:

$$\left[\frac{1}{3}x^3 + x\right]_0^2 = \left[\frac{1}{3}2^3 + 2\right] - \left[\frac{1}{3}0^3 + 0\right] = \frac{8}{3} + 2 = \frac{14}{3}$$

Leaving the answer answer as a fraction is fine, or you could convert it to decimals. To reiterate, this number is the area between the function $y = x^2 + 1$ and the x-axis, evaluated between $x = 0$ and $x = 2$. Remember that we **do not** need a constant of integration when evaluating a definite integral.

2.1.3 Context: Normalising Wavefunctions

In quantum mechanics, we have to ensure that our wavefunctions are **normalised**, which means that the integral of the square of the wavefunction $\Psi(x)$ over all

space is equal to one:

$$\int_{-\infty}^{\infty} \Psi(x)^2 \, \mathrm{d}x = 1$$

Evaluating this integral for given wavefunctions enables us to calculate **normalisation constants** for the wavefunctions. We'll see this again next year.

You can now attempt the problems on **problem sheet 5**.

2.2 Integration Hacks

Last time we saw how to integrate lots of things, and this'll serve us well for our needs as chemists. Calculating integrals quickly and efficiently is a large part of what computational and quantum chemistry is about, and to do this it's often possible to exploit symmetry, or other things, to make our lives easier (and calculations faster).

Note that in a previous version of this course, some other techniques for analysing integrals: 'integration by parts' and 'integration by substitution'. These are no longer part of the course, so don't panic if you see these on exam papers and cannot answer them.

2.2.1 Simple Shortcuts

The integral of anything multiplied by zero is still zero, even if that thing is very complicated:

$$\int_0^{\pi} \frac{\cos^2(x)\sin(\pi)}{x^3 + x^{1.5} - 5}\ \mathrm{d}x = 0$$

Because $\sin(\pi) = 0$.

The integral of anything between identical limits is zero by definition:

$$\int_{\pi}^{\pi} \frac{\Gamma(x)\mathrm{e}^{-x^2}}{1+\ln(x)}\mathrm{d}x = 0$$

Because in this case the area must be zero - imagine shrinking the limits in Figure 2.1 until there was no area between them.

We can also split up the limits of integration which can sometimes render the integral more straightforward to evaluate:

$$\int_{-2}^{3} f(x)\ \mathrm{d}x = \int_{-2}^{0} f(x)\ \mathrm{d}x + \int_{0}^{3} f(x)\ \mathrm{d}x$$

You can think of this as just doing the integral 'in stages' rather than all at once.

2.2.2 Exploiting Symmetry

The area produced by an integral is the area between the function and the x-axis - but what happens if the function is below the x-axis? By convention, this is defined as a **negative area**. So what does this mean when we calculate an area like the one shown in Figure 2.2?

Clearly, if the area below the x-axis is defined as negative area, the total area defined by this integral (the blue area in Figure 2.2) must be zero. This is because the function used here is an **odd function**, which means it satisfies the relationship:

$$f(-x) = -f(x)$$

By symmetry, the integral of an odd function between symmetrical limits (i.e. between $-a$ and a) is zero:

$$\int_{-a}^{a} f(x)\ \mathrm{d}x = 0 \text{ if } f(x) \text{ is an odd function}$$

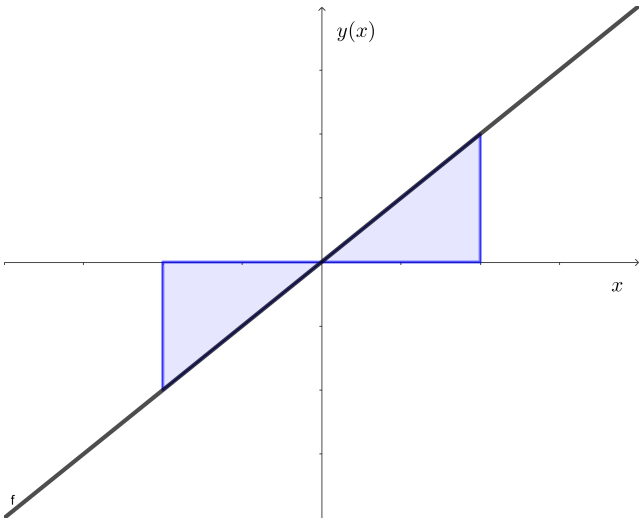

Figure 2.2: The integral of this function (an odd function) between symmetrical limits is zero, shown by the blue area.

Knowing this trick can make evaluating some horrific-looking integrals very easy.

Similarly, we can also have functions that are **even functions**, this means they satisfy the relationship:

$$f(-x) = f(x)$$

By symmetry, the integral of an even function between symmetrical limits is given by:

$$\int_{-a}^{a} f(x)\ \mathrm{d}x = 2\int_{0}^{a} f(x)\ \mathrm{d}x \text{ if } f(x) \text{ is an even function}$$

An example of this is shown in Figure 2.3. When evaluating difficult looking integrals between symmetrical limits, **always check if the function is even or odd first** - it can save you a lot of time!

2.2.3 Standard Integrals

Sometimes, and especially in quantum chemistry, there are integrals which cannot be evaluated in an indefinite way. A classic example of this is the integral of a **Gaussian function**:

$$\int \mathrm{e}^{-x^2}\ \mathrm{d}x = \text{ undefined!}$$

However, many of these integrals can be evaluated as definite integrals, such as:

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\ \mathrm{d}x = \sqrt{\pi}$$

That this happens to become the square root of π is pretty beautiful, but anyway.

The point is that quite a lot of the time we can look up these **standard integrals**, or just do them on a computer[2]. There's no shame in just using a result like this,

[2] I'll add some information at the end of this course about doing calculus on computers with Python, for interest.

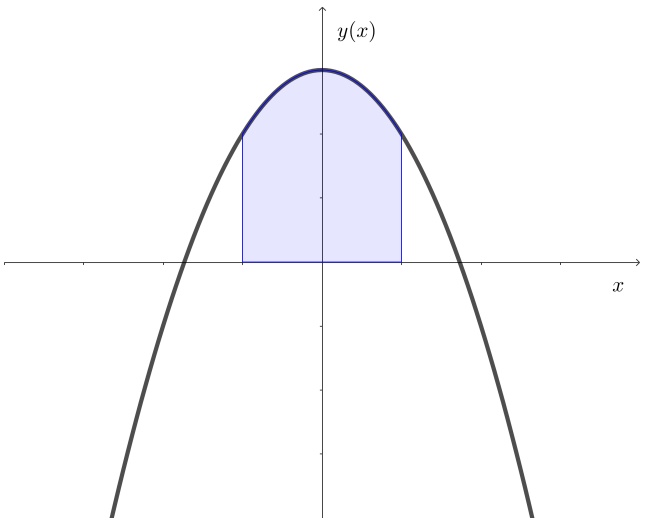

Figure 2.3: The integral of this function (an even function) between symmetrical limits is twice the integral of the function between 0 and the upper limit.

as we're chemists, not mathematicians. I Google integrals all the time when I can't remember how they work.

In an exam, you'll be given any standard integrals you need - but remember for your future life and career that if you need to find out what a certain integral is and can't work it out - Google, or standard maths textbooks, will contain lists of standard integrals that can help.

2.2.4 Context: Computational Chemistry

In computational chemistry, you use software (such as *Gaussian*) to calculate the structure and dynamics of molecules. This amounts to solving the Schrödinger equation for the molecules, and in practice the computer ends up evaluating tens of thousands of integrals (or more) to accomplish this.

When setting up a job in computational chemistry, you'll define something called a **basis set**, which is a set of mathematical functions that can be used to approximate the orbitals on the atoms in your molecule. Generally, these basis sets are made up of different kinds of Gaussian functions, because they have well known standard integrals. This means that it's blazingly fast for a computer to calculate the integrals - as you can just program the computer to know all the standard integrals. In fact, to calculate complex orbitals, one common strategy is to approximate the orbital as a series of different Gaussian functions, to increase the speed of the computation. You may see basis sets called things like *3-21G* or *6-31+G* - the numbers in these names refer to the number of different Gaussian functions that are used to approximate the orbitals. Larger numbers increase accuracy, but take longer to compute. Some basis sets go beyond Gaussian orbitals, to get increased accuracy, but at the expense of computational cost - integrating Gaussian functions like the one described above is really fast!

You can now attempt the problems on **problem sheet 6**.

Please note - there are many other tricks that people use to evaluate integrals than we have covered here. If you are interested in going deeper into physical chemistry/chemical physics/anything mathematical later on, it might be worth looking up things like **integration by parts** and **integration by substitution** in any maths textbook (or online). What we've learnt about integration here will suffice for most of what you see in the main chemistry course.

Chapter 3

Differential Equations

Finally, we are going to talk about **differential equations**. These are equations which involve some kind of derivative, such as:

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

Equations like this are incredibly commonplace in science, mathematics, and economics. This is because they allow us to relate *how things are going to change* (via derivatives) to *how things currently are* (via the other terms in the equation). There are entire university departments devoted just to studying differential equations, so we can only scratch the surface in two sessions - have a look online (3blue1brown has a nice series of videos) for more depth if you find this interesting. Differential equations are really beautiful.

These equations pop up in chemistry all the time in fields like **chemical kinetics**, where we are studying rates of reaction via **rate equations**:

$$\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = -k[\mathrm{A}]$$

These are equations that relate the rate of a chemical reaction (on the left - change in concentration $[\mathrm{A}]$ with time t) to concentration of reactants/products [A]. They also pop up in quantum mechanics - such as the time-independent Schrödinger equation:

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\Psi(x)}{\mathrm{d}x^2} = E\Psi$$

Other examples of things that ultimately come from differential equations are the Beer-Lambert law; the equations that govern radioactive decay; and the Clausius-Clapeyron and van t'Hoff equations from thermodynamics. We will start to see how these equations can be solved in the next two sessions.

3.1 First-Order Differential Equations

A **first-order** differential equation is one where the highest order derivative is $\frac{\mathrm{d}y}{\mathrm{d}x}$ - so there are no second or third derivatives. A simple example of an equation like this is:

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

You can solve an equation like this in a couple of ways:

- **By inspection**: we want a function $y = y(x)$ that will differentiate to just x. What about the function $y = \frac{1}{2}x^2 + C$, where C is a constant. This is often an quick way to solve simple equations.
- **By integration** - we can integrate both sides with respect to x:

$$\int \frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}x = \int x\,\mathrm{d}x$$

$$\int \mathrm{d}y = y = \frac{1}{2}x^2 + C$$

  On the LHS of the equation, you can imagine that the $\mathrm{d}x$ terms cancel under the integral sign - **again, this is not really what happens**. Note that we only need to add one constant of integration - if we added two, we could just as easily combine them into one, so we might as well just have one.

So we see for this simple example, the function that solves the equation is:

$$y = \frac{1}{2}x^2 + C$$

It's always a good idea to plug this function back into the equation and check that it solves it.

This function is known as the **general solution** to the differential equation, because we haven't yet determined C. To determine C we have to know something else about our system - for example, we might be told that $y = 1$ when $x = 2$. In this case, we can work out C:

$$1 = \frac{1}{2}2^2 + C \Rightarrow 1 = 2 + C \Rightarrow C = -1$$

Thus, in this situation, we can write down the **particular solution**:

$$y = \frac{1}{2}x^2 - 1$$

3.1.1 Separation of Variables

It's not always the case that we can just directly integrate our differential equations to solve them. Sometimes, we have to do some algebra before we can integrate them, and this is called **separating the variables**. We need to get it so that one side of the equation has all the y terms and the other has all the x terms (or whatever our variables are called). We'll illustrate this with an example from chemical kinetics - for the rate of a first order reaction of molecule A:

$$\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = -k[\mathrm{A}]$$

In this equation, $[\mathrm{A}]$ is the concentration of molecule A, t is time, and k is the rate coefficient. However, equations of this form govern all kinds of things - we'll see

one of them later. This equation tells us that the rate of change of concentration of A (the thing on the LHS) depends on the concentration of A (on the RHS). If there is a lot of A, then our rate will be a large - does it make sense?

To solve this equation, we first need to separate the variables:

$$\frac{1}{[\mathrm{A}]}\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = -k$$

Now we can integrate this equation with respect to t:

$$\int \frac{1}{[\mathrm{A}]}\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t}\ \mathrm{d}t = \int -k\ \mathrm{d}t$$

Which simplifies to:

$$\int \frac{1}{[\mathrm{A}]}\ \mathrm{d}[\mathrm{A}] = \int -k\ \mathrm{d}t$$

Doing the integrals results in:

$$\ln[\mathrm{A}] = -kt + C$$

Which we could recast as:

$$[\mathrm{A}] = \mathrm{e}^{-kt+c} = \mathrm{e}^{-kt}\mathrm{e}^{c}$$

This is fine as a general solution, but we would normally also told something else, like that the initial concentration of A $[\mathrm{A}] = [\mathrm{A}]_0$ at time $t = 0$. Then we can find a particular solution:

$$[\mathrm{A}]_0 = \mathrm{e}^{c} \Rightarrow [\mathrm{A}] = [\mathrm{A}]_0\mathrm{e}^{-kt}$$

Which might look like a familiar result from CH1203.

3.1.2 Context: Beer-Lambert Law

The Beer-Lambert law looks a lot like the solution we've just shown above for that differential equation - and that's no accident. We can derive[1] it as follow. The Beer-Lambert law says that the rate of change of intensity of light I as it passes through a volume of solution of depth L is proportional the concentration of the solution c, the molar absorptivity of the molecule in solution ϵ, and the intensity of light in the solution I:

$$\frac{\mathrm{d}I}{\mathrm{d}L} = -Ic\epsilon$$

Looks familiar! Separating the variables and integrating, we find:

$$\ln I = -\epsilon cL + k$$

Where I've called the integration constant k to avoid confusion. We know that the transmitted intensity when $L = 0$ is just the incident intensity, I_0, so we can say:

$$\ln I_0 = k$$

and therefore, if we rearrange it all:

$$\ln\left(\frac{I}{I_0}\right) = -\epsilon cL = -A$$

[1] I've skipped a couple of steps but the main idea is fine - look up *integrating factors* if you want a deeper derivation.

where A is the absorbance. Equally, we could write the above as:

$$I = I_0 \mathrm{e}^{-\epsilon c L}$$

Nice.

You can now attempt the problems on **problem sheet 7**.

3.2 Second-Order Differential Equations

Finally, we will talk briefly about **second-order** differential equations - these are equations where the highest order term is a second derivative, such as:

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

These kinds of equations occur all the time in scientific problems - the Schrödinger equation is an example, as are many of the classical equations of motion. We're going to wrap up this course by doing a brief overview of solving these equations.

3.2.1 Educated Guessing

We don't solve equations like the ones above by integrating them twice (though you could do this in some circumstances). Instead, what we do is use our mathematical intuition to make some sort of educated guess as to what the solution is. This means we are looking for a function $y = y(x)$ which will differentiate to a multiple of itself, and also differentiate again to a multiple of itself. An obvious function that would fit this description is some kind of exponential function, as these differentiate to a multiple of themselves:

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

Plugging this into our equation at the top will result in:

$$a^2\mathrm{e}^{ax} + a\mathrm{e}^{ax} = 2\mathrm{e}^{ax}$$

We can tidy this up by factorising the exponential and collecting everything on one side:

$$(a^2 + a - 2)\mathrm{e}^{ax} = 0$$

How do we proceed from here? We can divide by the exponential term (as this can't ever be equal to zero), and are left with:

$$a^2 + a - 2 = 0$$

This is called the **auxiliary equation** for our differential equation.

3.2.2 The Auxiliary Equation

The auxiliary equation is what tells us the possible values of a that solve our equation, so we just need to solve it like any other equation. In this case, we can factorise it:

$$(a + 1)(a - 2) = 0 \Rightarrow a = -1 \text{ or } a = 2$$

So there are two possible values of a that satisfy our original equation, so we have two solutions:

$$y_1 = \mathrm{e}^{-x} \text{ and } y_2 = \mathrm{e}^{2x}$$

It turns out that actually any linear combination of these two solutions will also solve our equation, so we can write down the **general solution** as:

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

Where A and B are constants. We'll leave that here for now. Just remember that the main idea with solving differential equations is to **find functions that differentiate to themselves as guesses** - exponential functions always work. An old professor of mine used to say *'the solution's always some kind of exponential, just write that'* whenever he couldn't work one out.

3.2.3 Context: The Schrödinger Equation

We saw the Schrödinger equation at the start of this chapter, which is an example of a second-order differential equation. So can we solve it?

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\Psi(x)}{\mathrm{d}x^2} = E\Psi$$

We can, let's guess a wavefunction like:

$$\Psi = \mathrm{e}^{-kx}$$

Plugging this in, we'll find:

$$\frac{\hbar^2 k^2}{2m}\mathrm{e}^{-kx} = E\mathrm{e}^{-kx}$$

So we can clearly identify:

$$E = \frac{\hbar^2 k^2}{2m}$$

We just solved the Schrödinger Equation! We haven't determined k yet, because this requires that we impose some **boundary conditions**, which are ultimately what make our energy **quantised**. We'll talk about this over the next couple of years.

A final note, is that you've seen functions like:

$$\Psi = \sin(kx)$$

as solutions before too - these will also solve the equation above (try and see). The reason for this is that **sines and cosines are just exponential functions, but in a different form**. To go into this in more depth requires that we talk about **complex numbers**, which we don't have time to do now. But as a teaser, I can tell you that:

$$\sin(x) = \frac{\mathrm{e}^{ix} - \mathrm{e}^{-ix}}{2i} \text{ and } \cos(x) = \frac{\mathrm{e}^{ix} + \mathrm{e}^{-ix}}{2}$$

Where i is defined as $i^2 = -1$. We can talk about this another time, if you're interested :)

You can now attempt problems on **problem sheet 8**.

3.3 Appendices

Information below is **for interest only and will not be examined.**

3.3.1 Differentiation from First Principles

We can use the sketch in Figure 3.1 to understand how we can differentiate **from first principles**. We know that the gradient is defined as the change in y over

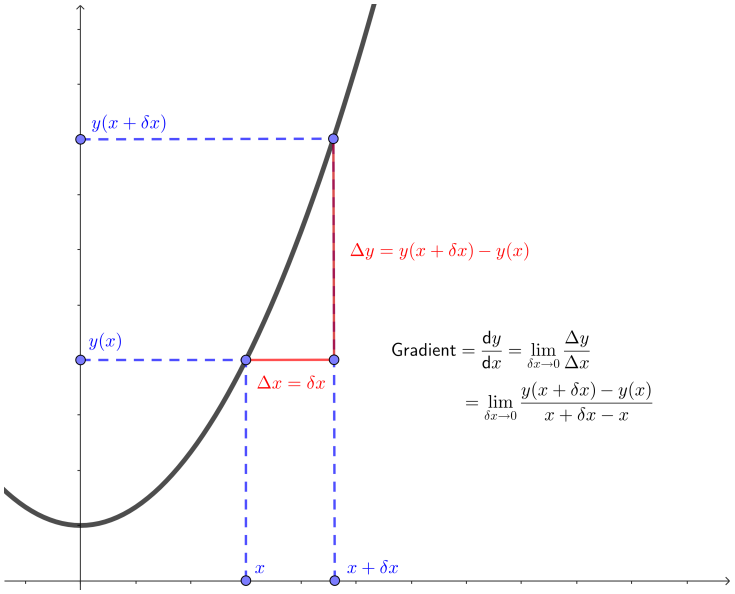

Figure 3.1: Differentiation from first principles.

change in x, for a generic function $y(x)$, we can define the change in x given a certain change in x, δx as:

$$\Delta y = y(x + \delta x) - y(x)$$

And similarly:

$$\Delta x = (x + \delta x) - x$$

Thus, we can define the gradient of our generic function $y(x)$ as:

$$\frac{\Delta y}{\Delta x} = \frac{y(x + \delta x) - y(x)}{(x + \delta x) - x}$$

This is an approximation, because our function is curved and so the gradient will change as we move along the function by the amount δx. The approximation becomes exact, however, if we take the limit that $\delta x \to 0$, and this is how we define $\frac{dy}{dx}$:

$$\frac{dy}{dx} = \lim_{\delta x \to 0} \left[\frac{y(x + \delta x) - y(x)}{(x + \delta x) - x} \right]$$

Which means we take the limit of the thing in square brackets δx tends towards zero. This is the true definition of the derivative, $\frac{dy}{dx}$. Let's show that it works by

differentiating $y = x^2$:

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x} &= \lim_{\delta x \to 0}\left[\frac{(x+\delta x)^2 - x^2}{(x+\delta x) - x}\right] \\ &= \lim_{\delta x \to 0}\left[\frac{(x^2 + 2x\delta x + \delta x^2) - x^2}{\delta x}\right] \\ &= \lim_{\delta x \to 0}[2x + \delta x] \\ &= 2x\end{aligned}$$

Just as the formulas from earlier showed. Nice.

3.3.2 Calculus with Computers

If you've measured some experimental data and want to integrate or differentiate it, how would you do that? We won't then have an analytical (algebraic) expression for the data, but could still want to know the gradient of a curve or the area between a graph and an axis. In this situation, we can use a computer to do **numerical differentiation** or **numerical integration**. This technique uses algorithms to find approximate (or 'numerical') values for the derivatives or integrals of our data - these will (strictly) be approximate values, but in reality they can be accurate to within a few hundredths of a percent or more.

In Python, the *numpy* package contains functions for doing this kind of numerical calculus. There are functions like `numpy.diff()` and `numpy.gradient()` which can perform numerical differentiation using things like *finite difference methods*. There are also functions like `numpy.trapz()` which does numerical integration using something called the *trapezoidal rule*.

There are also packages that can do **symbolic algebra** - where you can type in an equation and it will work out the analytical derivatives and integrals. One of these in Python is called *sympy*, but a more well known programming language for this is the *Wolfram Language* - you may have been to `www.wolframalpha.com`, this is a website where you can type in any mathematical expression and it can do algebra for you.