3. DIFFERENTIATION

3.1 The Slope Graph

Suppose we have a smooth curve where there’s a tangent at every point. We could find the value of the slope at every point and plot this against the x-values. This new graph would be the slope graph corresponding to the original one, where heights on the slope graph correspond to slopes on the original graph.

Where the original graph becomes momentarily horizontal the slope graph would cross the x-axis. This is because, for that value of x, the slope would be zero. Whenever the original graph is increasing (going uphill from left to right) the slope would be positive and so the slope graph would be above the x-axis. When the original graph decreases the slope graph drops below the x-axis.

Here the original curve is the solid one and the dotted curve is the corresponding slope graph. Note that whenever the original graph peaks or troughs the slope graph crosses the x-axis. This is because the slope of the original graph is zero for those values of x.

Given a graph we can construct the slope graph as follows:
(1) Choose a number of points on the original graph.
(2) With a ruler, draw the tangent at each of these points.
(3) Work out the slope of each of these tangents and enter them in a table against the corresponding x-values.
(4) Plot these points on the same graph or on another graph.
(5) Join these points by a smooth curve.

This is very hard to do accurately because of step (2). It’s very difficult to draw an accurate tangent. We’ll ask you to carry out this procedure only to establish the concept of what a slope graph is. Later we’ll develop much simpler and more accurate ways of doing this.

3.2 Finding The Derivative of \( x^2 \)

If all we have is a graph and we have to find the slope graph then the above method, as cumbersome and inaccurate as it is, is all we can do. But fortunately behind many graphs there’s a formula which defines a function. In that case we can bypass the graph altogether and, with a few quick steps, find the formula for the slope graph. For example we’ll show that if our original graph is the graph of \( y = x^2 \) the slope graph is the graph of \( y = 2x \).
The function that corresponds to the slope graph is called the **derivative** and the process of finding the derivative is called **differentiation**. So if the original function is \( y = x^2 \) we’ll show that its derivative is \( y' = 2x \). Note that we couldn’t say that the derivative is \( y = 2x \). This would be too confusing because \( y \) refers to the original function. Rather than use another symbol altogether we put a little dash next to the \( y \) (just like a little slope). So \( y' \) (call it “\( y \) dash” if you’re reading it aloud) is the derivative (slope function) of the variable \( y \). Isaac Newton used a variation of this idea. Instead of the dash he put a dot over the \( y \) and wrote \( \dot{y} \).

Differential Calculus is the area of mathematics that studies derivatives, or slope functions together with their applications.

Let’s begin with a very easy case. What’s the derivative of \( y = 2x - 1 \)? This equation represents a straight line with slope 2. Like all straight lines the slope is the same at every point so the derivative (slope function) is a constant. The derivative is \( y' = 2 \). If we were to plot \( y' \) against \( x \), to get the slope graph, we’d get a horizontal line that’s 2 units above the \( x \)-axis. But the whole point of calculus is that we don’t need the graphs. We just go straight from formula to formula.

Now let’s attempt to differentiate \( y = x^2 \). Here the slope is constantly changing so the derivative won’t be constant. First, let’s work out the slope at \( x = 2 \). When \( x = 2 \) then \( y = 4 \) so the tangent passes through the point (2, 4). Can we find a second point on this tangent? If we take a point on the curve very near to (2, 4) it won’t lie on the tangent. But the chord, the line joining them, will have almost the same slope as the tangent. This approximation will get better and better the closer we make this second point to (2, 4). But of course we can never let the two points become one because we could no longer work out the slope.

![chord_approximates_tangent](image)

**A chord approximates a tangent. The closer the points the better the approximation.**

Suppose we take \( x = 2.1 \). Then \( y = (2.1)^2 = 4.41 \). So (2.1, 4.41) is a nearby point on the curve. The slope of the chord is \( \frac{4.41 - 4}{2.1 - 2} = \frac{0.41}{0.1} = 4.1 \). This isn’t the slope of the tangent at \( x = 2 \) but its a good approximation to it.

Suppose we take \( x = 2.01 \). Then \( y = (2.01)^2 = 4.0401 \). The slope of the chord is \( \frac{0.0401}{0.01} = 4.01 \). This gives us a better approximation.

Rather than repeating this process and getting better and better approximations, let’s take a general point near (2, 4). This will be the point \((2 + h, (2 + h)^2)\). If \( h > 0 \) this point will be to the right of (2, 4) and if \( h < 0 \) it will be to the left. And if \( h \) is small it will be close to (2, 4).

The smaller we make \( h \) the closer will be the two points and the better will be the approximation. What we have here is some sort of limiting process.
Now $(2 + h)^2 = 4 + 4h + h^2$. You can see this either by substituting into the formula $(a + b)^2 = a^2 + 2ab + b^2$ or by considering the areas in the following diagram.

![Diagram with labeled areas: h, 2h, h^2, 2, 4, 2h]

The slope of the chord joining these nearby points is \[
\frac{(4 + 4h + h^2) - 4}{(2 + h) - 2} = \frac{4h + h^2}{h} = 4 + h.
\]

By taking $h$ sufficiently small this can be made as close as we like to 4. In other words, in the limit, as $h$ approaches 0, this approaches 4. The derivative of $y = x^2$ at $x = 2$ then is exactly 4.

Now it might appear that we’ve taken $h = 0$. In practice we get the right answer, for the wrong reason, provided we put $h = 0$ after the division has been carried out. But strictly speaking we don’t put $h = 0$. To do so would give us a chord joining a point to itself which is nonsense. And the slope of the chord would come out as $\frac{0}{0}$ which is also nonsense. We are actually taking the limit as $h$ approaches 0, not putting $h = 0$.

So the derivative of $y = x^2$ at $x = 2$ is 4. What about at $x = 3$? When $x = 3$ then $y = 9$ so we consider a chord with $(3, 9)$ at one end and $(3 + h, (3 + h)^2)$ at the other. The slope of the chord is \[
\frac{(3 + h)^2 - 9}{h} = \frac{(9 + 6h + h^2) - 9}{h} = 6 + h.
\]

Clearly this approaches 6 as $h$ approaches 0.

We could continue in this way, building up a table of values of slopes at various points. We could then plot these against $x$ to get a slope graph. But it would be much better, if we can, to get a formula giving the slope for a general value of $x$. Let’s see if we can do this for $y = x^2$.

Take the general point $(x, x^2)$ on the curve $y = x^2$. Now take a nearby point $(x + h, (x + h)^2)$. The slope of the chord is \[
\frac{(x + h)^2 - x^2}{h} = \frac{x^2 + 2xh + h^2 - x^2}{h} = 2x + h.
\]

As $h$ approaches 0 this approaches $2x$, so $2x$ must be the slope of the tangent (and hence the slope of the curve) at this point. In other words the derivative of $x^2$ is $2x$.

We use the symbol “$\rightarrow$” to mean “approaches” so we can write $x + h \rightarrow 2x$ as $h \rightarrow 0$. Another way of writing this is $\lim_{h \to 0} (2x + h) = 2x$. We read this as “the limit, as $h$ approaches 0, of $2x + h$ is $2x$”.

By doing the general case we can now substitute any value of $x$ we like and instantly work out the corresponding slope. For example, the slope of the tangent at $x = 10$ will be 20.

### 3.3 Limits

Limits are a very difficult concept, and in fact it took mathematicians a couple of hundred years to really come up with a precise, satisfactory definition. To say that $\lim_{h \to 0} (2 + h^2) = 2$ doesn’t quite mean that $2 + h^2$ gets closer and closer to 2 as $h$ gets closer to 0. After all $2 + h^2$ gets closer and closer to 1 as $h$ gets closer and closer to 0, but $\lim_{h \to 0} (2 + h^2) = 2$, not 1. Certainly the closer $h$ is to zero the closer $2 + h^2$ is to 1, but it never gets very close to 1. In fact it always stays above 2.

The reason why $\lim_{h \to 0} (2 + h^2) = 2$, not 1, is that we can make $2 + h^2$ as close as we like to 2 by making $h$ sufficiently close to zero. Those phrases capture the true meaning of limits. But even these phrases are a little imprecise, so a formal, technical definition of limits has been formulated and accepted as the basis of the theory of limits. But because it’s somewhat technical we’ll omit it here. We’ll be able to get by with a somewhat more intuitive, and less precise, concept of limits. After all, mathematicians managed to do so for a very long time!
To say that \( \lim_{h \to 0} (2 + h^2) = 2 \) looks as though we are simply substituting \( h = 0 \). In most cases, when an expression exists for \( h = 0 \), the limit as \( h \to 0 \) and the value when \( h = 0 \) will be the same. (This is not always the case, but it is true for functions that are said to be continuous.) But the limit can exist, as \( h \to 0 \), for functions that don’t exist at \( h = 0 \).

For example, we can’t put \( h = 0 \) in the expression \( \frac{h^2 + 2h}{h} \). But \( \lim_{h \to 0} \frac{h^2 + 2h}{h} \) exists. In fact it’s equal to 2. To see this we simply replace \( \frac{h^2 + 2h}{h} \) by something which, for \( h \neq 0 \), is equivalent to \( \frac{h^2 + 2h}{h} \). We simply divide by “h”. (If \( h = 0 \) this would be wrong. You must never divide by zero.) The limit as \( h \to 0 \) refers to the behaviour of points close to zero, not what happens, if anything at \( h = 0 \). For \( h \neq 0 \) the expression \( \frac{h^2 + 2h}{h} \) is the same as \( h + 2 \) and so \( \lim_{h \to 0} (h + 2) = 2 \).

If all this sounds very confusing simply follow the following protocol. An experienced mathematician would be able to devise an artificial situation where this wouldn’t work but for all situations that you’re likely to come across it does.

If you want the limit as \( h \to 0 \) of some expression try putting \( h = 0 \). If you get a value then this is the limit. If, however, you get the meaningless \( \frac{0}{0} \) you must try to replace the expression by one that’s equivalent to it, at least for \( h \neq 0 \). (For example you could divide numerator and denominator by \( h \).) Then if this new expression exists for \( h = 0 \), make this substitution.

There are many fundamental properties of limits that seem plausible, but can only be properly proved by getting embroiled in the precise, technical definition. Because this is a very elementary account of calculus we’ll avoid that. So you must accept, without proof, the following intuitively reasonable results as facts. (They all presuppose that the limits of the various functions exist.)

- The limit of a sum of functions is the sum of the limits.
- The limit of a difference between two functions is the difference between the limits.
- The limit of a product of functions is the product of the limits.
- The limit of a quotient of two functions is the quotient of the limits (provided that the limit of the denominator is non-zero).

### 3.4 The General Method For Finding Derivatives

What we did for \( y = x^2 \) we can do for any function, provided the algebra works out and we can find the limit. We’re going to develop this method using a new notation. We’ll use the symbol “\( \Delta \)” to denote “an increment of …”, that is “an extra little bit of …”. On its own it doesn’t make sense, just as \( \sqrt{} \) means “square root of …”. It must have something following it to make it complete. In this case “\( \Delta \)” is followed by a variable. So \( \Delta x \) means “an increment of x”. This is what we were calling “h” before. It’s just that \( \Delta x \) reminds us that it’s an extra little bit of the \( x \) variable.

Just as with subscripts \( \Delta x \) is a single quantity with a double-barrelled name. We mustn’t separate the \( \Delta \) from the \( x \). Now when \( x \) receives an increment it “increases” to \( x + \Delta x \). We say “increases”, but remember if \( \Delta x \) is negative it would actually represent a \textit{decrease}. Any variable that depends on \( x \) will receive a corresponding increment. If \( y \) is a function of \( x \) the corresponding increment of \( y \) is written as \( \Delta y \).
The slope of the chord joining \((x, y)\) to \((x + \Delta x, y + \Delta y)\) is simply \(\frac{-\Delta y}{\Delta x}\). So the derivative of \(y\), with respect to \(x\), is the limit of \(\frac{\Delta y}{\Delta x}\) as \(\Delta x \to 0\), which we write as \(\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}\). Notice the phrase “with respect to \(x\)”. We’ll have some more to say about that later.

Leibniz (1646-1716) introduced the notation \(\frac{dy}{dx}\) for the derivative. There are two things to note about this notation. Firstly notice the similarity between \(\frac{dy}{dx}\) and \(\frac{\Delta y}{\Delta x}\). The first is the slope of the tangent while the second is the slope of the chord. So we can write

\[
\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}
\]

Apart from the fact that it reminds us how to find the derivative, the Leibniz \(\frac{dy}{dx}\) notation incorporates both variables. Newton (1642-1727) developed calculus for the sole purpose of describing planetary motion, so for him every quantity, speed, position etc was a function of time. There was no need to be reminded of that when writing down derivatives. But there can be more than one choice of variable on which a given quantity depends.

Consider the volume of a cube. If the side of the cube is \(x\) the volume is \(V = x^3\) and the surface area is \(S = 6x^2\) (6 faces each an \(x\) by \(x\) square). We can combine these to express \(V\) in terms of \(S\). Since \(x = \sqrt{S/6}\) from the second equation we have \(V = \left(\sqrt{S/6}\right)^3\). This version enables us to give the volume if we happen to know the surface area. But it’s a different function to \(V = x^3\) which expresses volume in terms of the length of the side. If we plotted \(V = \left(\sqrt{S/6}\right)^3\) against \(S\) and \(V = x^3\) against \(x\) we’d get entirely different graphs with different derivatives. If we simply wrote the derivative as \(\dot{V}\) or \(V'\) it wouldn’t be clear which derivative we meant.

Another thing to note about the \(\frac{dy}{dx}\) notation is that we can’t cancel the d’s to get \(\frac{V}{x}\). In fact, unlike \(\frac{\Delta y}{\Delta x}\) which is made up of two pieces, \(\Delta y\) and \(\Delta x\), \(\frac{dy}{dx}\) is a single entity. It is not even a fraction, though it looks like one. It’s the limit of a fraction. So the \(dy\) and the \(dx\) have no independent existence.

Having said that \(\frac{dy}{dx}\) isn’t a fraction, in some circumstances it acts as if it is one. But more of this later.

The General Method For Finding \(\frac{dy}{dx}\).

(1) Write \(y\) as a function of \(x\).
(2) Substitute \(x + \Delta x\) for \(x\) and call the result \(y + \Delta y\).
(3) Subtract \(y\) to get \(\Delta y\).
(4) Divide by \(\Delta x\) to get \(\frac{\Delta y}{\Delta x}\).
(5) Take the limit as \(\Delta x \to 0\) to get \(\frac{dy}{dx}\).
**Example 1:** If \( y = x^2 \), find \( \frac{dy}{dx} \).

1. \( y = x^2 \).
2. \( y + \Delta y = (x + \Delta x)^2 \)
3. \( \Delta y = (x + \Delta x)^2 - x^2 = 2x \Delta x + (\Delta x)^2 - x^2 = 2x \Delta x + (\Delta x)^2 \).
4. \( \frac{\Delta y}{\Delta x} = \frac{2x \Delta x + (\Delta x)^2}{\Delta x} = 2x + \Delta x \).
5. \( \frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} (2x + \Delta x) = 2x \).

Note that we resisted the temptation to write \((\Delta x)^2\) as \(\Delta^2 x^2\). That would have meant splitting the \(\Delta\) from the \(x\) which we must never do.

**Example 2:** If \( y = \frac{1}{x} \), find \( \frac{dy}{dx} \).

1. \( y = \frac{1}{x} \).
2. \( y + \Delta y = \frac{1}{x + \Delta x} \)
3. \( \Delta y = \frac{1}{x + \Delta x} - \frac{1}{x} = \frac{x - (x + \Delta x)}{x(x + \Delta x)} = -\frac{\Delta x}{x(x + \Delta x)} \).
4. \( \frac{\Delta y}{\Delta x} = \frac{-1}{x(x + \Delta x)} \).
5. \( \frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{-1}{x(x + \Delta x)} = -\frac{1}{x^2} \).

**Example 3:** If \( y = \sqrt{x} \), find \( \frac{dy}{dx} \).

1. \( y = \sqrt{x} \).
2. \( y + \Delta y = \sqrt{x + \Delta x} \)
3. \( \Delta y = \sqrt{x + \Delta x} - \sqrt{x} \)
4. \( \frac{\Delta y}{\Delta x} = \frac{\sqrt{x + \Delta x} - \sqrt{x}}{\Delta x} \).
5. \( \frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{\sqrt{x + \Delta x} - \sqrt{x}}{\Delta x} \).

We have a bit of a problem here because there’s no \(\Delta x\) in the numerator that we can cancel. There’s a bit of a trick here. We *rationalise*, i.e. we multiply top and bottom by \(\sqrt{x + \Delta x} + \sqrt{x} \). Why? Well, just watch what happens.

\[
\frac{dy}{dx} = \lim_{\Delta x \to 0} \left( \frac{\sqrt{x + \Delta x} - \sqrt{x}}{\Delta x} \right) \left( \frac{\sqrt{x + \Delta x} + \sqrt{x}}{\sqrt{x + \Delta x} + \sqrt{x}} \right)
= \lim_{\Delta x \to 0} \frac{\sqrt{x + \Delta x} - \sqrt{x}}{\Delta x} \left( \frac{\sqrt{x + \Delta x} + \sqrt{x}}{\sqrt{x + \Delta x} + \sqrt{x}} \right)
= \lim_{\Delta x \to 0} \frac{(\sqrt{x + \Delta x})^2 - (\sqrt{x})^2}{\Delta x(\sqrt{x + \Delta x} + \sqrt{x})} \text{ (difference of two squares)}
= \lim_{\Delta x \to 0} \frac{(x + \Delta x) - x}{\Delta x(\sqrt{x + \Delta x} + \sqrt{x})}
\]
\[
\lim_{\Delta x \to 0} \frac{\Delta x}{\Delta x(\sqrt{x + \Delta x} + \sqrt{x})} \quad \text{(now we can cancel)}
\]
\[
= \lim_{\Delta x \to 0} \frac{1}{\sqrt{x + \Delta x} + \sqrt{x}}
\]
\[
= \frac{1}{2\sqrt{x}}.
\]

3.5 The Sum And Difference Rules

We’ve discovered that the derivative of \(y = x^2\) is 2x and the derivative of \(y = 5x + 2\) is 5 (the line \(y = 5x + 2\) has constant slope 5). What about \(y = x^2 + 5x + 2\)? Do we simply add the derivatives to get the derivative of a sum? The answer is “yes”.

**SUM RULE**

\[
\frac{d(u + v)}{dx} = \frac{du}{dx} + \frac{dv}{dx}
\]

Here \(u\) and \(v\) are functions of \(x\). In the above example \(u = x^2\) and \(v = 5x + 2\). The derivative of the sum is indeed the sum of the derivatives.

The reason for this is as follows:

1. Let \(y = u + v\).
2. If \(x\) receives the increment \(\Delta x\), and \(\Delta u\) and \(\Delta v\) are the corresponding increments of \(u, v\):
   \[
   y + \Delta y = (u + \Delta u) + (v + \Delta v).
   \]
3. Hence \(\Delta y = (u + \Delta u) + (v + \Delta v) - (u + v) = \Delta u + \Delta v\).
4. \[
   \frac{\Delta y}{\Delta x} = \frac{\Delta u + \Delta v}{\Delta x} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x}.
   \]
5. As \(\Delta x \to 0\) this becomes \(\frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx}\).

In the same way we have the Difference Rule:

**DIFFERENCE RULE**

\[
\frac{d(u - v)}{dx} = \frac{du}{dx} - \frac{dv}{dx}
\]

**Example 4:** Since the derivative of \(x^2\) is 2x and the derivative of \(3x + 5\) is 3, it follows from the sum rule that the derivative of \(x^2 + 3x\) is 2x + 3.

3.6 The Product Rule

If the derivative of a sum is the sum of the derivatives then surely the derivative of a product must be the product of the derivatives. WRONG! It’s somewhat more complicated than that. In fact if \(u\) and \(v\) are functions of \(x\) we’ll show that:

**PRODUCT RULE**

\[
\frac{d(uv)}{dx} = u \cdot \frac{dv}{dx} + v \cdot \frac{du}{dx}
\]
In other words, when you differentiate a product you get two terms. Each of the factors gets differentiated while the other remains as it is.

The reason for this is as follows:
(1) Let \( y = uv \).
(2) If \( x \) receives the increment \( \Delta x \), and \( \Delta u \) and \( \Delta v \) are the corresponding increments of \( u \), \( v \):
\[
y + \Delta y = (u + \Delta u)(v + \Delta v).
\]
(3) Hence \( \Delta y = (u + \Delta u)(v + \Delta v) - uv = u\Delta v + v\Delta u + \Delta u\Delta v \).
(4) \[
\frac{\Delta y}{\Delta x} = \frac{u\Delta v + v\Delta u + \Delta u\Delta v}{\Delta x} = \frac{\Delta v}{\Delta x}u + \frac{\Delta u}{\Delta x}v + \Delta u\frac{\Delta v}{\Delta x}.
\]
(5) As \( \Delta x \to 0 \) this becomes \[
\frac{dy}{dx} = \frac{du}{dx}v + \frac{dv}{dx}u + 0\cdot \frac{du}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}.
\]

We could have used the product rule to differentiate \( y = x^2 \) by writing it as \( y = x\cdot x \). The product rule gives:
\[
\frac{d}{dx}(x\cdot x) = \frac{dx}{dx}x + x\frac{dx}{dx} = 2x\frac{dx}{dx}. \quad \text{Now } \frac{dx}{dx} \text{ looks as though it should be 1 by cancellation. It is in fact 1 but for a different reason. (Remember we’re not allowed to treat } dx \text{ as a quantity on its own.) But } \frac{dx}{dx} \text{ is simply the derivative, or slope function, of } y = x. \text{ Since } y = x \text{ represents a straight line whose slope is 1, the derivative of } x \text{ with respect to } x \text{ is 1. This finally gives us the derivative of } x^2 \text{ as } 2x. \text{ Note that this is } \textit{not} \text{ the product of the derivatives is 1.}
\]

What about \( y = x^3 \)? We could treat that as \( y = (x^2)(x) \). The product rule then gives:
\[
\frac{dx^3}{dx} = x^2\frac{dx}{dx} + x\frac{dx^2}{dx} = x^2 + 2x^2 = 3x^2. \quad \text{In this way we can differentiate } x^n \text{ for any positive integer } n.
\]

Another use we can make of the product rule is the Constant Factor Rule. We know that the derivative of \( x^2 \) is 2x. Is the derivative of \( 10x^2 \) equal to 20x? Indeed it is, by the Constant Factor Rule.

If \( k \) is a constant then:

\[
\text{CONSTANT FACTOR RULE}
\]
\[
\frac{d(ku)}{dx} = k\frac{du}{dx}
\]

This is just a particular case of the product rule.
\[
\frac{d(ku)}{dx} = k\frac{du}{dx} + u\frac{dk}{dx} = k\frac{du}{dx} \text{ because the derivative of a constant is zero. Why? Well the line } y = k \text{ is horizontal and so has zero slope everywhere.}
\]

Example 5: Since the derivative of \( x^2 \) is 2x, the derivative of \( 5x^2 \) is 10x.

Example 6: Use the product rule to find the derivative of \( x^2\sqrt{x} \).

Solution: Let \( u = x^2 \) and \( v = \sqrt{x} \). Then \( \frac{du}{dx} = 2x \) and \( \frac{dv}{dx} = \frac{1}{2\sqrt{x}} \).

So \[
\frac{d(x^2\sqrt{x})}{dx} = u\frac{dv}{dx} + v\frac{du}{dx} = x^2\left(\frac{1}{2\sqrt{x}}\right) + \sqrt{x}(2x) = \frac{x\sqrt{x}}{2} + 2x\sqrt{x} = \frac{5x\sqrt{x}}{2}.
\]
Since \(x^2\sqrt{x} = x^{5/2}\) we can find its derivative more quickly using the Power Rule that we’re just about to develop.

### 3.7 The Power Rule

The derivative of \(x^2\) is \(2x\) and the derivative of \(x^3\) is \(3x^2\). There seems to be a pattern here. The power is copied out at the front as a factor and then reduces by 1 in its original position. If we’re right then the derivative of \(x^7\) should be \(7x^6\).

If we guess a correct formula in terms of a positive integer \(n\) and want to verify that it’s indeed correct for all \(n\) we can use a principle called the **Principle of Induction**. This involves two stages. We first check that our guess is true for \(n = 1\). Then assuming that it works for \(n\) we show that it works for \(n + 1\).

We’ve guessed that the derivative of \(x^n\) is \(nx^{n-1}\). Certainly we’re right for \(n = 1\) for this simply says that the derivative of \(x^1\) is \(1x^0\). Remember that \(x^1\) is just \(x\) and any non-zero number raised to the power zero is 1. So our general formula says, for \(n = 1\), that the derivative of \(x\) is 1. That’s certainly true.

Now suppose that we’re correct for \(n\), that the derivative of \(x^n\) is \(nx^{n-1}\). Must it work for \(n + 1\)? Well, we can write \(x^{n+1}\) as the product \(x^n x\) and so use the Product Rule.

\[
\frac{dx^{n+1}}{dx} = \frac{d(x^n x)}{dx} = x^n \frac{dx}{dx} + x \frac{dx^n}{dx} = x^n + x (nx^{n-1}) = x^n + nx^n = (n + 1)x^n.
\]

That’s what we predicted. The power is copied out the front and then drops by 1 in its original position.

So if it’s true for \(n\) we’ve shown that it works for \(n + 1\). But we did show that it works for \(n = 1\), so it must work for \(n = 2\). And since it works for \(n = 2\) it must work for \(n = 3\), and so on. It is like an infinite row of dominos, standing up one behind the other. If any one gets knocked over it will knock over the next and so by knocking over the first one they all fall over!

The Power Rule is:

<table>
<thead>
<tr>
<th>POWER RULE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{dx^n}{dx} = nx^{n-1} )</td>
</tr>
</tbody>
</table>

We’ve checked it for positive \(n\), and if \(n = 0\), \(x^n = x^0 = 1\). Being a constant, the derivative of 1 is 0, which is what the Power Rule is saying. What about negative values of \(n\)?

Suppose \(u = x^n\) and \(v = x^{-n} = \frac{1}{x^n}\). Then \(uv = 1\). We can use the Product Rule to find the derivative of \(uv\) with respect to \(x\). But this is 0, since \(uv = 1\), a constant. In the process we’ll get the derivative of \(v\).

\[
0 = \frac{d}{dx} \frac{du}{dx} = u \frac{dv}{dx} + v \frac{du}{dx} = x^n \frac{dv}{dx} + \frac{nx^{n-1}}{x^n} = x^n \frac{dv}{dx} + \frac{n}{x} \text{ so } \frac{dv}{dx} = -\frac{n}{x^{n+1}} = (-n)x^{-n-1}.
\]

This is what the Power Rule would say if we wanted to differentiate \(y = x^{-n}\). So it works if \(n\) is a negative integer as well.

The Power Rule even works for fractional powers. We won’t show this for a general fraction, but just in the case of the most important fractional power \(x^{1/2}\). Remember this is another way of writing \(\sqrt{x}\). Let \(u = v = \sqrt{x} = x^{1/2}\). Then \(uv = x\), and we know the derivative of \(x\). What does the Product Rule say?

\[
1 = \frac{dx}{dx} = \frac{d(uv)}{dx} = u \frac{dv}{dx} + v \frac{du}{dx} = 2u \frac{du}{dx} \text{ since } u = v. \text{ So } \frac{du}{dx} = \frac{1}{2u}.
\]

In other words \(\frac{d\sqrt{x}}{dx} = \frac{1}{2\sqrt{x}}\).

[Note that we derived this from first principles in example 3 on page 42.]
Expressing this using fractional powers, this says that \( \frac{dx^{1/2}}{dx} = \frac{1}{2} x^{-1/2} \). But this is just what the Power Rule predicts. A copy of the power goes out the front as a factor and then it reduces by 1 in its original position.

Although the Power Rule can be used to differentiate any power of \( x \) there are some special cases that occur so frequently that it’s much better to remember the answer rather than have to work it out each time. They are:

\[
\frac{d\sqrt{x}}{dx} = \frac{1}{2\sqrt{x}} \quad \text{and} \quad \frac{d(1/x)}{dx} = -\frac{1}{x^2}.
\]

### 3.8 Differentiating Polynomial-like Functions

Combining the Sum Rule, the Constant Factor Rule and the Power Rule we can differentiate any function which consists of terms of the form \( kx^n \). These are called “polynomial-like functions”. (A true polynomial function restricts itself to the case where the powers are all non-negative integers.)

**Example 7:** Differentiate \( 3x^4 - 5x^3 + 7x^2 - 2x + 17 \) (with respect to \( x \)).

**Solution:** The derivative is \( 12x^3 - 15x^2 + 14x - 2 \).

**Example 8:** If \( y = 3x^5 + 7(x + 1)^2 - 4x\sqrt{x} + \frac{5}{x^2} \), find \( \frac{dy}{dx} \).

**Solution:** The first thing we do is to write the terms in the \( kx^n \) form.

\[
y = 3x^5 + 7x^2 + 14x + 7 - 4x^{3/2} + 5x^{-2}.
\]

Then \( y' = 15x^4 + 14x + 14 - 6x^{1/2} - 10x^{-3} = 15x^4 + 14x + 14 - 6\sqrt{x} - \frac{10}{x^3} \).

### 3.9. Other Notations For The Derivative

The most useful notation for the derivative of a function is \( \frac{dy}{dx} \). It mentions not only the variable being differentiated but also the variable that it’s being differentiated with respect to.

We’ve also seen the \( y' \) notation and Newton’s \( \dot{y} \) notation.

Another notation, emphasising the fact that differentiation is a process, represents this process by a capital “\( D \)”. So we might write \( Dy \) for the derivative of \( y \).

Another variation is to use “functional notation”. This takes the emphasis off the variable and puts it more on the function itself. So we might write “\( y = f(x) \)” to represent a general function. The advantage of this notation is that if we want to substitute a specific value for \( x \) we can do so by putting that value in place of \( x \) in the \( f(x) \) notation.

So if \( f(x) = x^2 \) then \( f(2) = 4 \), \( f(-3) = 9 \), and so on. Sometimes we substitute another variable for \( x \). With \( f(x) = x^2 \) we would have \( f(t) = t^2 \). We can even substitute an expression for \( x \). So if \( f(x) = x^2 \) then \( f(x + 1) = (x + 1)^2 \) and \( f(\sqrt{x}) = (\sqrt{x})^2 = x \).

The derivative of \( f(x) \), with respect to \( x \), can therefore be written as \( \frac{df(x)}{dx} \) or \( f'(x) \) or \( Df(x) \).
3.10. Higher Derivatives

In some circumstances we need to differentiate the derivative, that is, differentiate the original function twice. Or perhaps we need to differentiate three or even more times. These higher order derivatives are called the second, third, fourth derivatives and so on. We extend the various notations for the first derivative to these higher order derivatives.

If we write \( y' \) or \( f'(x) \) for the first derivative we just put a second dash for the second derivative and write \( y'' \) or \( f''(x) \). In principle we can continue putting extra dashes to denote the higher derivatives, but after three or four dashes this is messy and instead put the number of dashes in brackets where the dash should be. So instead of \( y'''''' \) for the 7th derivative we write \( y^{(7)} \). The brackets are essential, as \( y^7 \) and \( y^{(7)} \) have totally different meanings and are almost never equal.

Something similar occurs with the dots over the variable. Two dots indicate the 2nd derivative. After this we revert to the \( y^{(n)} \) notation.

If we write the derivative of \( y \) as \( D_y \) then the second derivative could be written as \( D(D_y) \) but it’s more usual to write it as \( D^2y \). This makes it look as though \( D \) is a number, which it’s not. Remember that, like \( \Delta \), \( D \) cannot exist on its own. The third derivative would therefore be written as \( D^3y \), and so on. We might also write \( D^n f(x) \) to denote the \( n \)th derivative of \( f(x) \).

When it comes to the \( \frac{dy}{dx} \) notation there is a rather curious convention. Strictly speaking, the second derivative of \( y \), with respect to \( x \), should be written as \( \frac{d(dy)}{dx} \), but this is a bit of a mouthful. Instead it’s written as \( \frac{d^2y}{dx^2} \). Why?

Well, just pretend that the d’s in the \( \frac{dy}{dx} \) were numbers with their own separate existence. They’re not, of course, but I said “let’s pretend”. The rather cumbersome \( \frac{d(dy)}{dx} \) could then be simplified to \( \frac{d^2y}{(dx)^2} \). Then somebody must have said, “let’s leave out the brackets” and we got \( \frac{d^2y}{dx^2} \). In one sense it’s quite crazy and illogical. But once you get used to it, it seems natural enough. The third derivative is written \( \frac{d^3y}{dx^3} \) and the \( n \)th derivative \( \frac{d^ny}{dx^n} \).

We can summarise all these notations as follows:

<table>
<thead>
<tr>
<th>Derivatives of ( y = f(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Derivative</strong></td>
</tr>
<tr>
<td>( y' )</td>
</tr>
<tr>
<td>( y'' )</td>
</tr>
<tr>
<td>( D_y )</td>
</tr>
<tr>
<td>( \frac{dy}{dx} )</td>
</tr>
<tr>
<td>( f'(x) )</td>
</tr>
<tr>
<td>( Df(x) )</td>
</tr>
<tr>
<td>( \frac{df(x)}{dx} )</td>
</tr>
</tbody>
</table>
3.11. The Chain Rule

What’s the derivative of $y = (3x^2 + 2)^{10}$ with respect to $x$? We could attempt to expand it and differentiate but there’s got to be an easier way. There is! If we let $u = 3x^2 + 2$ we can write $y$ very easily in terms of $u$: $y = u^{10}$.

So this more complicated function can be broken down into a chain of two simpler ones:

$$y = u^{10}$$

$u = 3x^2 + 2$

This reflects the way we would evaluate $(3x^2 + 2)^{10}$ if we were given a specific value of $x$, except that we’d go in the reverse order. We’d first compute $3x^2 + 2$ and then raise this number to the 10'th power.

Now the derivatives of each of these simpler functions is easy to compute:

$$\frac{dy}{du} = 10u^9 \text{ and } \frac{du}{dx} = 6x.$$  Is there a simple way to combine $\frac{dy}{du}$ and $\frac{du}{dx}$ to get $\frac{dy}{dx}$? Yes. It’s simply that $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$.

It looks as though we’ve simply cancelled the $du$’s but remember that they are not separate entities. However it makes it easy to remember the result by pretending to cancel the $du$’s.

Now it is true that $\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x}$. At this level the $\Delta y$, $\Delta u$ and $\Delta x$ exist as individual quantities and we can cancel the $\Delta u$’s. Then we simply take limits of both sides. As $\Delta x \to 0$ so $\Delta u \to 0$. We also need to remember the fact that the limit of a product is the product of the limits. This gives us the Chain Rule.

**CHAIN RULE**

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

So to complete the above example, if $y = (3x^2 + 2)^{10}$ then $\frac{dy}{dx} = 10u^9 \cdot 6x = 60xu^9$, right?

Wrong! The original function made no mention of the variable “$u$” so nor should the answer. To finish off we must substitute for “$u$” to get an answer entirely in terms of $x$.

$$\frac{dy}{dx} = 60x(3x^2 + 2)^9.$$  

**Example 9:** If $y = \sqrt{x^2 + 1}$, find $\frac{dy}{dx}$.

**Solution:** Let $u = x^2 + 1$. Then $y = u^{1/2}$.

Hence $\frac{du}{dx} = 2x$ and $\frac{dy}{du} = \frac{1}{2\sqrt{u}} = \frac{1}{2\sqrt{x^2 + 1}}$.

Hence $\frac{dy}{dx} = \frac{2x}{2\sqrt{x^2 + 1}} = \frac{x}{\sqrt{x^2 + 1}}$.

With a bit of practice it’s possible to use the use the Chain Rule to write down the final answer without having to explicitly make the substitution. In the above example, you could say to yourself, “$y$” is the square root of “something” so the derivative of “$y$” is $1$ over twice the square root of “something” (adapting the result that the derivative of $\sqrt{x}$ is $\frac{1}{2\sqrt{x}}$).
So at this stage you’d write down \( \frac{1}{2\sqrt{x^2+1}} \). But you have a feeling that this is incomplete, and so you should. Because you’ve only differentiated with respect to the “something” (the \( x^2 + 1 \)) you have to apply the “correction factor” as dictated by the Chain Rule. You now have to differentiate the “something” with respect to “x”. That gives the second factor, 2x. So with all this going on in your head you need only write down:

\[
\frac{dy}{dx} = \frac{1}{2\sqrt{x^2+1}} \cdot 2x = \frac{2x}{2\sqrt{x^2+1}} = \frac{x}{\sqrt{x^2+1}}.
\]

3.12. The Quotient Rule

Example 10: If \( y = \frac{x^2}{x^2+1} \), find \( \frac{dy}{dx} \).

Solution: We could write this as \( y = x^2 \left( \frac{1}{x^2+1} \right) \). This can be differentiated using the Product Rule, using \( u = x^2 \) and \( v = \frac{1}{x^2+1} \). But first we need to differentiate \( \frac{1}{x^2+1} \). We can do that by writing it as \((x^2+1)^{-1}\) and using the Chain Rule.

The derivative of \((x^2+1)^{-1}\) is \(- \left( \frac{1}{(x^2+1)^2} \right) \cdot 2x = - \frac{2x}{(x^2+1)^2} \).

Thus the derivative of \( \frac{x^2}{x^2+1} \) is \( x^2 \left( \frac{-2x}{(x^2+1)^2} \right) + \frac{1}{x^2+1} \cdot 2x = \frac{2x(x^2+1) - 2x^3}{(x^2+1)^2} = \frac{2x}{(x^2+1)^2} \).

We can handle any quotient in the same way. However it’s worth following through the method in the general case in order to develop the Quotient Rule.

\[
\text{QUOTIENT RULE} \quad \frac{d}{dx} \left( \frac{u}{v} \right) = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2}
\]

It’s a fairly difficult formula to learn. If you ever do forget it, or are not quite sure whether you’ve remembered it correctly, you can always resort to using the Product Rule in conjunction with the Chain Rule.

Rather than remember the formula as you would a piece of poetry, try remembering it as a process.

1) If the question is a quotient, so is the answer. Draw a long fraction bar: 

2) Square the denominator: \( \frac{1}{v^2} \)

3) Now put a “−” in the middle of the top: \( \frac{-}{v^2} \)

4) Each of numerator and denominator gets differentiated, but only one at a time.

Important: Differentiate the top first: 

\[
\frac{\frac{du}{dx} - \frac{dv}{dx}}{v^2}
\]
Example 11: Use the Quotient Rule to differentiate \[ \frac{x^2}{x^2 + 1} \] (this is the function we differentiated in Example 10, using the product rule).

Solution: \[ \frac{2x(x^2 + 1) - x^2(2x)}{(x^2 + 1)^2} = \frac{2x}{(x^2 + 1)^2} \]. (This was much easier.)

EXERCISES FOR CHAPTER 3

Exercise 1: Differentiate the following with respect to x:
(i) \( 5x^6 \);
(ii) \( 5x^6 - 7x^3 + 2x + 13 \);
(iii) \( 3x^4 + 7x + 1093 \);
(iv) \( x^2(x^{98} + 1) \);
(v) \( (x^{10} + x)^2 \).

Exercise 2: Find \( \frac{dy}{dx} \) for each of the following:
(i) \( y = x^{\frac{7}{2}} \);
(ii) \( y = \frac{2}{x^3} \);
(iii) \( y = \frac{x + 1}{x^2 + x} \) [Be careful here – there’s a trap which you must avoid!];
(iv) \( y = \frac{x + 1}{x^2 + x} \) [Hint: Don’t use the quotient rule It’s much easier than that!];
(v) \( y = \frac{2}{x^5} + x^2\sqrt{x} + 13 \).

Exercise 3: (a) Use the Product Rule to differentiate \((5x^2 + 2x)\sqrt{x}\).
(b) Now expand \((5x^2 + 2x)\sqrt{x}\) into two separate terms and differentiate. Compare your answers.

Exercise 4: Use the Chain Rule to find \( \frac{dy}{dx} \) when:
(i) \( y = (x^2 + 1)^{10} \);
(ii) \( y = \sqrt{x^2 + 1} \);
(iii) \( y = \frac{1}{x^2 - 1} \).

Exercise 5: Use the Product Rule to find the derivative of \( y = (x^2 - 1)\sqrt{x^2 + 1} \). [Note that you found the derivative of \( \sqrt{x^2 + 1} \) in exercise 4(ii).]

Exercise 6: Without using the Product Rule find the derivative of \( y = (x^2 + 1)\sqrt{x^2 + 1} \).

Exercise 7: If \( y = x^4 + 5x^2 + 7x + 2 \), find:
(i) \( Dy \);
(ii) \( y'' \);
(iii) \( \frac{d^3y}{dx^3} \);
(iv) \( y^{(5)} \).

Exercise 8: Find the fifth derivative of \( x^5 + x^4 + x^3 + x^2 + x + 1 \).
Exercise 9: Find $\frac{d^{100}y}{dx^{100}}$ when $y = x^5 + x^4 + x^3 + x^2 + x + 1$. [You have to differentiate 100 times but it’s not as bad as it seems!]

Exercise 10: Find the slope of the tangent to the following curves at $x = 1$:
(i) $y = 3x^4 - 4x^2 + 13$;
(ii) $y = \frac{1}{x} + \sqrt{x}$;
(iii) $y = \sqrt{x^2 - 1}$.
[Something very peculiar happens in this case. Can you work out what’s going on?]

SOLUTIONS FOR CHAPTER 3

Exercise 1:
(i) $30x^5$; (ii) $30x^5 - 21x^2 + 2$; (iii) $12x^3 + 7$; (iv) $x^2(x^{98} + 1)$ can be expanded to $x^{100} + x^2$ so the derivative is $100x^{99} + 2x$ [unless requested to, never use the Product Rule if the function can be easily expanded]; (v) $(x^{10} + x)^2 = x^{20} + 2x^{11} + x^2$ so the derivative is $20x^{19} + 22x^{10} + 2x$.

Exercise 2:
(i) $\frac{dy}{dx} = \frac{7}{2}x^{5/2}$; (ii) $\frac{dy}{dx} = -10x^{-6} = -\frac{10}{x^6}$; (iii) 0 [don’t forget that $\pi^2$ is a constant]; (iv) We can write $y = \frac{x + 1}{x(x + 1)} = \frac{1}{x}$ so $\frac{dy}{dx} = -\frac{1}{x^2}$; (v) $y = 2x^{-6} + x^{5/2} + 13$ so $\frac{dy}{dx} = -12x^{-7} + \frac{5}{2}x^{3/2}$.

Exercise 3:
(a) Let $u = 5x^2 + 2x$ and $v = \sqrt{x}$. Then $\frac{du}{dx} = 10x + 2$ and $\frac{dv}{dx} = \frac{1}{2\sqrt{x}}$. So $\frac{dy}{dx} = \frac{5x^2 + 2x + 2(10x + 2)}{2\sqrt{x}} = \frac{25x^2 + 6x}{2\sqrt{x}} = \frac{\sqrt{x}(25x + 6)}{2}$.

Exercise 4:
(i) Let $u = x^2 + 1$. Then $y = u^{10}$. So $\frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx} = 10u^9 \cdot 2x = 20x(x^2 + 1)^9$;
(ii) Let $u = x^2 + 1$. Then $y = u^{1/2}$. So $\frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx} = \frac{1}{2\sqrt{u}} \cdot 2x = \frac{x}{\sqrt{x^2 + 1}}$;
(iii) Let $u = x^2 - 1$. Then $y = \frac{1}{u}$. So $\frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx} = -\frac{1}{u^2} \cdot 2x = -\frac{2x}{(x^2 - 1)^2}$.

Exercise 5:
Let $u = x^2 - 1$ and $v = \sqrt{x^2 + 1}$. Then $\frac{du}{dx} = 2x$ and $\frac{dv}{dx} = \frac{x}{\sqrt{x^2 + 1}}$. So $\frac{dy}{dx} = u \frac{dv}{dx} + v \frac{du}{dx} = (x^2 - 1) \frac{x}{\sqrt{x^2 + 1}} + \sqrt{x^2 + 1} \cdot 2x = \frac{x^3 - x + 2x}{\sqrt{x^2 + 1}} = \frac{x^3 + x}{\sqrt{x^2 + 1}}$.
**Exercise 6:** We can write \( y = (x^2 + 1)^{3/2} \). Let \( u = x^2 + 1 \). So \( y = u^{3/2} \). Since \( \frac{dy}{du} = \frac{3}{2} u^{1/2} \) and \( \frac{du}{dx} = 2x \) we have \( \frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx} = 3xu^{1/2} = 3x\sqrt{x^2 + 1} \).

**Exercise 7:**
(i) \( D_y = 4x^3 + 10x^2 + 7 \); (ii) \( y'' = 12x^2 + 20x \); (iii) \( \frac{d^3y}{dx^3} = 24x + 20 \); (iv) \( y^{(5)} = 0 \).

**Exercise 8:** Let \( y = x^5 + x^4 + x^3 + x^2 + x + 1 \). Then \( \frac{dy}{dx} = 5x^4 + 4x^3 + 3x^2 + 2x + 1 \), \( \frac{d^2y}{dx^2} = 20x^3 + 12x^2 + 6x + 2 \), \( \frac{d^3y}{dx^3} = 60x^2 + 24x + 6 \), \( \frac{d^4y}{dx^4} = 120x + 24 \), \( \frac{d^5y}{dx^5} = 120 \).

**Exercise 9:** From exercise 8 we see that \( \frac{d^6y}{dx^6} = 0 \). From this point on, all higher derivatives are zero so \( \frac{d^{100}y}{dx^{100}} = 0 \).

**Exercise 10:** To find the slope of the tangents, differentiate and put \( x = 1 \).
(i) \( \frac{dy}{dx} = 12x^3 - 12x^2 \) so the tangent at \( x = 1 \) has slope zero [the tangent is horizontal];
(ii) \( \frac{dy}{dx} = -\frac{1}{x^2} + \frac{1}{2\sqrt{x}} \) so the tangent at \( x = 1 \) has slope \(-1 + \frac{1}{2} = -\frac{1}{2} \); 
(iii) \( \frac{dy}{dx} = \frac{x}{\sqrt{x^2 - 1}} \). Putting \( x = 1 \) gives \( \frac{1}{0} \) which doesn’t exist! If \( x \) is just a little bigger than 1 then \( \frac{dy}{dx} \) is very large (if \( x < 1 \) it doesn’t exist). So, what is happening at \( x = 1 \) is that the tangent is vertical.