Contents

Chapter 3

Differential Calculus

3.1 Preliminairies

3.1.1 More on real numbers

Calculus depends on properties of the real numbers. Algebraic properties of real numbers should be well known. Roughly speaking they assert that real numbers can be added, subtracted, multiplied and divided (except by zero) to produce other real numbers and the usual laws of arithmetic apply (see Chapter 1).

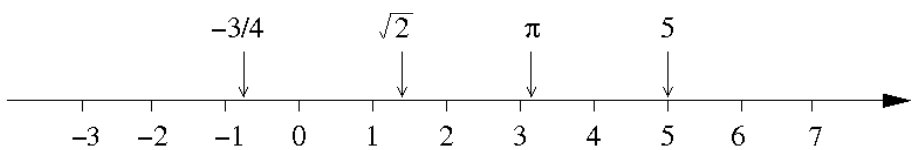

Figure 3.1: The Real Line

Recall that (real) numbers can represented by points on a line (see Figure 3.1). Indeed we often *identify* a number with the point that represents it, in which case the number is referred to as a 'point'. Formally we choose (Cartesian) coordinates on the line and identify a number a with the point whose coordinate is a, once the identification is made we refer to the *real line*. □

Real numbers have order properties. Recall that if a and b are real numbers then $a < b$ (or $b > a$) means a is *less* than b (or b is *greater* than a). If a is less than *or equal* to b then we write $a \leq b$. The basic order properties are summarised below.

For real numbers a, b and c:

- if $a < b$ and $b < c$ then $a < c$;
- if $a < b$ then $a + c < b + c$;
- if $a < b$ and $c > 0$ then $ac < bc$;
- if $a < b$ and $c < 0$ then $ac > bc$ (in particular $-a > -b$);

- if $0 < a$ then $0 < 1/a$;
- if $0 < a < b$ then $0 < 1/b < 1/a$.

3.1.1 Example. Show if $x \neq 1$ and $-3 < 1/(1-x) < -2$ then $4/3 < x < 3/2$.

Solution. Suppose that $-3 < 1/(1-x) < -2$ so that

$$-3 < 1/(1-x) \quad \text{and} \quad 1/(1-x) < -2.$$

If $1 - x > 0$ then from the second inequality above and the fifth and then first properties we see that $0 < -2$ which is not true. Thus $1 - x < 0$. Hence by the fourth property

$$-3(1-x) > 1 \quad \text{and} \quad 1 > -2(1-x),$$

thus

$$-3 + 3x > 1 \quad \text{and} \quad 1 > -2 + 2x.$$

Hence $x > 4/3$ and $x < 3/2$. $\square$

The *absolute value* of a real number a, denoted by $|a|$, is

$$|a| = \begin{cases} a & \text{if } a \text{ is positive,} \\ -a & \text{if } a \text{ is negative,} \end{cases}$$

($|0| = 0$). For example, $|5| = 5$ and $|-5| = 5$.

The absolute value of a real number is thought of as the magnitude (or size) of the number. If a and b are real numbers then $|b - a|$ is the length of the line segment between the points a and b on the real line. In particular $|a|$ is the length of the line segment joining points a and 0 on the real line. The following properties hold.

For real numbers a and b:

- $|a| \geq 0$, $a = 0$ if and only if $|a| = 0$;
- $|ab| = |a|\,|b|$;
- $|a \pm b| \leq |a| + |b|$

as you can easily check. $\square$

A subset of the real numbers, which has the property that it contains all the real numbers between any two of its members, is called an *interval*.

For instance if a and b are real numbers with $a < b$ then the subset consisting of all numbers x satisfying $a < x < b$ is an interval and is denoted by (a, b) (round brackets). An interval of this type is called an *open* interval.

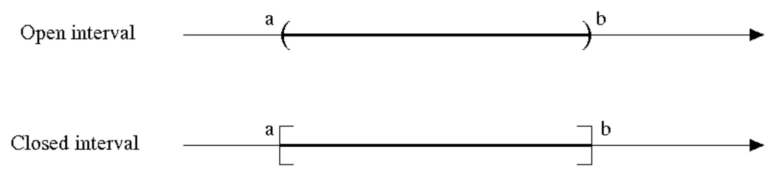

Figure 3.2: Open and Closed Intervals

The subset consisting of all numbers x satisfying $a \leq x \leq b$ is an interval and is denoted by $[a, b]$ (square brackets). An interval of this type is called a *closed* interval.

A closed interval contains the endpoints but an open interval does not contain them.

If a is a real number then the subset consisting of all numbers x satisfying $x > a$ is also an interval, as is the subset consisting of all numbers x satisfying $x < a$. These intervals are denoted by (a, ∞) and $(-\infty, a)$ respectively. [The symbol '∞' is read as 'infinity'. One can think of '∞' as the number which is larger than any real number (and is the least such number) so '$a < x < \infty$' just means x is a real number with $a < x$.] Both these intervals are called open intervals since they do not contain the endpoint points a and ∞.

The set $\mathbb{R}$ of all real numbers is also an open interval, it can be expressed as $\mathbb{R} = (-\infty, \infty)$.

3.1.2 Functions and variables

A *real function* is a rule which assigns to each number a in some subset of the real numbers (called the *domain* of the function) a *unique* real number called the *value* of the function at a.

A real function can be denoted by any symbol (or any combination of symbols) that takes your fancy, often f, g and h are used to denote functions. If f is a real function then the value of f at a is denoted by $f(a)$.

It is understood that in future 'function' means 'real function' and 'number' means 'real number'

(this is to save typing).

Often a function is defined using a formula. In these cases the domain of the function is taken to be the the largest subset of the real numbers for which the formula makes sense.

3.1.2 Example. Consider the function f defined by $f(x) = x^2 + 3x - 2$.

Replacing 'x' by a number gives the value of f at that number. For instance:

$$f(-3) = (-3)^2 + 3(-3) - 2 = -2, \quad f(1,000) = 1,002,998, \quad f(4.76) = 34.9376$$

In this case the formula for $f(x)$ makes sense for any (real) number x and so we can take the domain of the function f to be the set of all real numbers.

Remark: The symbol 'x' can be replaced by any convenient symbol. For instance the formula $f(y) = y^2 + 3y - 2$ defines exactly the same function. $\square$

3.1.3 Example. Consider the function g defined by $g(u) = \dfrac{u}{u-1}$. Some values are:

$$g(2) = \frac{2}{2-1} = 2, \quad g(-1.2) = \frac{-1.2}{-1.2-1} = \frac{1.2}{2.2} = \frac{6}{11} \ .$$

The formula does not make sense when $u = 1$ (remember it is not permissable to divide by zero) so g is not defined at 1. The domain of g is the set of all real numbers with 1 removed. In other words it is the union of the two open intervals $(-\infty, 1)$ and $(1, \infty)$. $\square$

Any function, p say, given by

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n$$

where a_0, a_1, a_2, a_3, ..., a_n are real numbers (called coefficients) with $a_n \neq 0$ is called a *polynomial* function of degree n. The domain of a polynomial function is $\mathbb{R}$ since the formula makes sense for any number x in $\mathbb{R}$. For example the function f given in Example 3.1.2 is an example of a polynomial function of degree 2.

The simplest type of polynomial function is a polynomial of degree 0, that is a constant function.

The identity function I is the polynomial function given by $I(x) = x$ for all x in $\mathbb{R}$.

Any function, r say, given by

$$r(x) = \frac{p(x)}{q(x)} \ ,$$

where p and q are polynomial functions with the degree of q greater than or equal to 1, is called a *rational* function, p is called the numerator and q the denominator of r. For example the function g in Example 3.1.3 is a rational function. The domain of a rational function is the set of all real numbers for which the denominator is *not* zero. $\square$

The trigonometric functions 'cos', 'sin' and 'tan' are particularly interesting. These were defined geometrically in Chapter 1. Recall that, in this course, for each real number a

$$\begin{aligned} \cos(a) &= \text{cosine of an angle of } a \textit{ radians} \\ \sin(a) &= \text{sine of an angle of } a \textit{ radians} \\ \tan(a) &= \text{tangent of an angle of } a \textit{ radians} \end{aligned}$$

(radians are emphasised). The domain of 'sin' and 'cos' is the whole of $\mathbb{R}$. Recall also that

$$\tan(x) = \frac{\sin(x)}{\cos(x)}.$$

The domain of 'tan' is all real numbers except the numbers whose cosine is zero. These numbers are

$$\ldots, \ -\tfrac{3}{2}\pi, \ -\tfrac{1}{2}\pi, \ \tfrac{1}{2}\pi, \ \tfrac{3}{2}\pi, \ \tfrac{5}{2}\pi, \ \ldots.$$

The domain of tan is the union of all the open intervals between these numbers.

Recall that if f is a function then $f(a)$ denotes the value of f at a (note the brackets round the number a). For the trigonometric functions it is common practice to drop the brackets writing, for instance, '$\cos a$' for '$\cos(a)$'. $\square$

In engineering one has to deal with quantities that may vary within the context of the problem at hand. A symbolic representation of such a quantity is called a *variable*. Any symbol (or combination of symbols) can be used to denote a variable, often letters from the end of the alphabet are used, that is ..., x, y or z (letters from the beginning of the alphabet are often used for quantities that remain constant).

In this course we restrict our attention to *real variables*, that is variables which represent quantities expressed as real numbers. Thus a real variable represents various real numbers, called the *values* of the variable. The possible values of a variable may be restricted as in Example 3.1.4.

Often a real variable will give a measure in terms of a *unit* of the quantity involved. To define such a real variable one has to specify the unit quantity. (The temperature may vary from point to point in some region in which case we could consider the real variable u which gives the temperature in degrees centigrade at each point of the region (the unit temperature is one degree centigrade)).

To save typing, in future it is understood that 'variable' means 'real variable'.

Let x and y be variables (that is real variables). We say that y *depends on* x if knowing the value of x determines the value of y uniquely; in this case we often write

$$y|_{x=a}$$

for the value of y when $x = a$.

More formally y depends on x means that y can be written as a function of x:

$$y = f(x) \quad \text{say}$$

so that $y = f(a)$ when $x = a$ (the function f is given by $f(a) = y|_{x=a}$ for all values a of x).

[For instance if u and v are variables which give the temperature at each point of a region in degrees centigrade and degrees Fahrenheit respectively then knowing the value of v determines the value of u, indeed you probably know that $u = \frac{5}{9}(v - 32)$. That is u is a function of v, $u = f(v)$ where in this case $f(v) = \frac{5}{9}(v - 32)$.]

Any quantity which remains fixed within the context of the problem at hand is called a *constant*.

In engineering problems it is important to establish relationships between the variables involved. In particular to determine if any variable depends on another variable. The next example demonstrates the ideas.

3.1.4 Example. A ladder of length l metres leans against a vertical wall, the foot of the ladder rests on a horizontal floor as in Figure 3.3. The ladder begins to slip down the wall. Suppose that we are interested in the motion of the ladder.

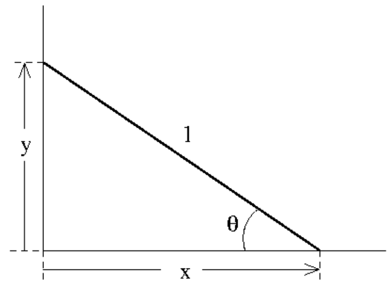

Figure 3.3: Moving Ladder

First we define the relevant variables. Let x be the distance in metres from the bottom of the wall to the foot of the ladder, let y be the distance in metres from the bottom of the wall to the top of the ladder and let θ be the angle in radians subtended by the ladder with the horizontal.

Since the ladder is moving x, y and θ are variables, real variables because distances and angles are expressed as real numbers. The length l of the ladder is not varying and is thus constant. The values of the variables are restricted. The values of x are between 0 and l, that is they lie in the interval $(0, l)$. Similarly for y. Also it is clear that $0 < \theta < \pi/2$, so θ is restricted to lie in the interval $(0, \pi/2)$.

There are some obvious relationships between the variables:

$$x^2 + y^2 = l^2, \quad \frac{y}{x} = \tan\theta, \quad \frac{y}{l} = \sin\theta, \quad \text{etc.}$$

Note these are *identities* because they hold for all possible values of the variables. (The first relationship follows from Pythagoras's theorem.) It is intuitively clear that y depends on x, indeed from the first relationship above (and the fact that y is non-negative) we see $y = \sqrt{l^2 - x^2}$ (the non-negative square root).

Similarly x depends on y: $x = \sqrt{l^2 - y^2}$.

From the third relationship we see y depends on θ: $y = l\sin\theta$. Also x depends on θ: $x = l\cos(\theta)$.

There is one very important variable that hasn't been mentioned as yet, that is time. Let t be the time (measured in seconds) from the moment the ladder begins to slip. The varaibles x, y and θ will depend on t (or at least we assume that they do), that is they are all functions of t:

$$x = f(t), \quad y = g(t), \quad \theta = h(t), \quad \text{say.}$$

The functions f, g and h are unknown. If we find $h(t)$ then

$$x = l\cos(\theta) = l\cos(h(t)) \quad \text{and} \quad y = l\sin(\theta) = l\sin(h(t))$$

which give x and y as functions of time. To determine the motion of the ladder it is sufficient to find a formula for $h(t)$. This can be done using Newton's laws of motion and calculus. □

Suppose that y depends on x, and thus $y = f(x)$ for some function f. In order to visualise the relationship between x and y we consider the 'graph $y = f(x)$', that is all points of the form $(a, f(a))$ where a is a value of x, plotted in the (x, y)-plane. [There is nothing special about x and y. (We could sketch the graph $v = f(u)$ in the (u, v)-plane. The only difference to the graph $y = f(x)$ would be the labels on the axes.) Essentially the graph only depends on f so we often refer to it as the 'graph of f'.]

3.1.5 Example. The graph $y = x^2 + 3x - 2$ is sketched in Figure 3.4.

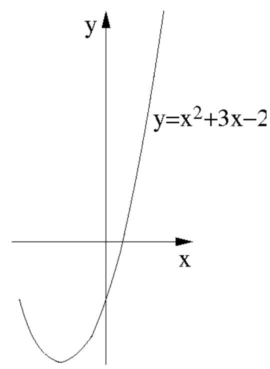

Figure 3.4: Graph of $y = x^2 + 3x - 2$

(In fact only part of the graph is sketched as the graph extends 'indefinitely'.) □

3.2 Limits and Continuity

3.2.1 Limits

The fundamental concept of calculus is the notion of a *limit.* Indeed calculus is basically concerned with the calculation of limits.

Given a function f and a number a, it is often of some interest to investigate the behaviour of $f(x)$ as a variable x tends to (approaches) the number a. The following examples demonstrate some possibilities.

3.2.1 Example. $f(x) = \dfrac{x^2+1}{x-1}$. In this case $f(x)$ is not defined when $x = 1$ (because we cannot divide by 0). It is of interest to investigate the behaviour of $f(x)$ as x tends to 1.

If $x > 1$ then it is fairly clear that x^2+1 tends to 2 and $x-1$ tends to 0 as x tends to 1 and also note that $x - 1 > 0$. Thus $f(x)$ will become large and *positive* as x tends to 1 from above.

If $x < 1$ then it is fairly clear that $x^2 + 1$ tends to 2 and $x - 1$ tends to 0 as x tends to 1 but in this case $x - 1 < 0$. Thus $f(x)$ will become large and *negative* as x tends to 1 from below.

Indeed $f(x)$ becomes arbitrarily large and positive if x is sufficiently close to 1 and $x > 1$ and $f(x)$ becomes arbitrarily large and negative if x is sufficiently close to 1 and $x < 1$. Figure 3.5 shows that as x tends to 1 from above $f(x)$ becomes arbitrarily

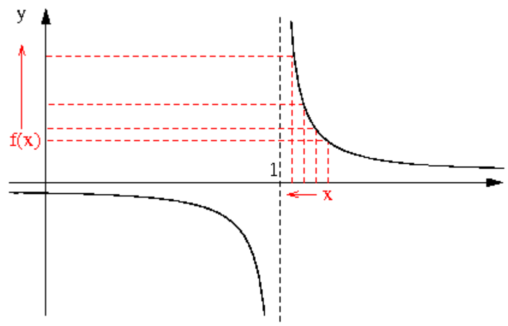

Figure 3.5: Graph $y = f(x)$

large and positive. (In the figure different scales are used for the x and y axes.) □

3.2.2 Example. $g(t) = \sin(1/t)$. It is of interest to investigate the behaviour of $g(t)$ as t tends to 0. We will just consider the case that t tends to 0 from above ($t > 0$). We sketch the graph $y = g(t)$ in Figure 3.6 (in the figure different scales are used for the t and y axes). We see that as t tends to 0 from above $g(t)$ will start to oscillate

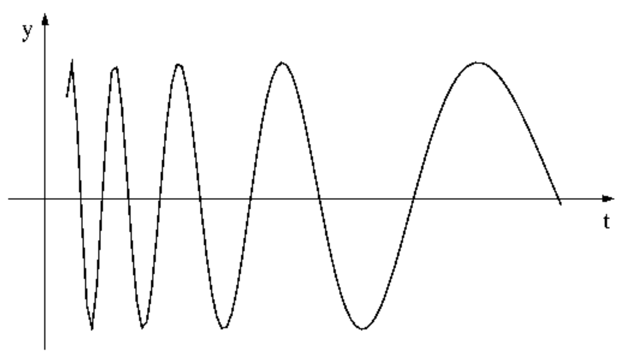

Figure 3.6: Graph $y = g(t)$

between -1 and 1 and $g(t)$ does not tend to a fixed number as t tends to 0. □

The next example demonstrates the 'simplest' type of behaviour.

3.2.3 Example. $h(x) = \dfrac{x^2-1}{x-1}$. As in Example 3.2.1. $h(x)$ is not defined when $x = 1$. We investigate the behaviour of $h(x)$ as x tends to 1.

If $x \neq 1$ ('x is *not* equal to 1') then we can simplify the expression for $h(x)$:

$$h(x) = \frac{x^2-1}{x-1} = \frac{(x-1)(x+1)}{x-1} = x+1 \ .$$

Thus the graph $y = h(x)$ is just the line $y = x + 1$ with the point $(1, 2)$ removed!!! From Figure 3.7 we can see that $h(x)$ tends to 2 as x tends to 1 from above (left

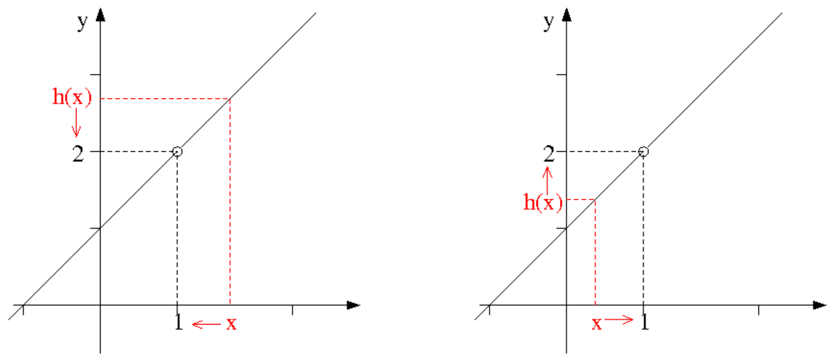

Figure 3.7: Graph $y = h(t)$

hand graph) and $h(x)$ tends to 2 as x tends to 1 from below (right hand graph). It is fairly clear (and indeed can be proved) that $f(x)$ is *arbirarily* close to 2 whenever x is *sufficiently* close to 1, that is we can make $h(x)$ as close to 2 as we please just by choosing x close enough to 1 (above or below). We say that $h(x)$ tends to the *limit* 2 as x tends to 1 and write

$$\lim_{x\to 1} h(x) = 2 \quad \text{or} \quad \lim_{x\to 1} \frac{x^2-1}{x-1} = 2\ .$$

$\square$

In general, if $f(x)$ tends to a (fixed) number, l say, as x tends to a (from above *and* from below) then we say the *limit* of $f(x)$ as x tends to a exists, the number l is called the *limit* of $f(x)$ as x tends to a and we write

$$\lim_{x\to a} f(x) = l \quad \text{or} \quad f(x) \to l \quad \text{as} \quad x \to a.$$

More precisely the terminology means that $f(x)$ is arbitrarily close to l whenever x is sufficiently close (above or below but not equal to) a.

In the notation, introduced above, the variable 'x' is called a *dummy* variable, it can be replaced by any symbol you like. The limit, should it exist, depends only on the function f and the number a, and not on the variable x. For example

$$\lim_{h\to a} f(h) = l, \quad \lim_{u\to a} f(u) = l \quad \text{or} \quad f(z) \to l \text{ as } z \to a$$

all mean the same thing.

In Example 3.2.1 we saw that $f(x) = (x^2+1)/(x-1)$ becomes arbitrarily large as x tends to 1 from above and so does not tend to a fixed number as x tends to 1, so the limit of $f(x)$ as x tends to 1 does *not* exist.

In Example 3.2.2 we saw that $g(t) = \sin(1/t)$ begins to oscillate between -1 and 1 as t tends to 0, more rapidly the smaller t and thus does not settle down to a fixed number (see Figure 3.6). The limit of $g(t)$ as t tends to 0 does *not* exist.

In Example 3.2.3 we saw that $h(x) = (x^2 - 1)/(x - 1)$ tends to the number 2 as x tends to 1 and so the limit of $h(x)$ as x tends to 1 does, indeed, exist and equals 2.

There are some useful properties of limits that we will use.

3.2.4 Theorem. *Let f and g be two functions and suppose that*

$$\lim_{x \to a} f(x) = l \quad \textit{and} \quad \lim_{x \to a} g(x) = m$$

then

$$\lim_{x \to a} f(x) + g(x) = l + m,$$
$$\lim_{x \to a} f(x)\, g(x) = l\, m$$

and if, in addition, $m \neq 0$ then

$$\lim_{x \to a} f(x)/g(x) = l/m.$$

We do not present a proof. The results are intuitively clear. □

In the notation of Theorem 3.2.4 if $m = 0$ and $l \neq 0$ then the limit of $f(x)/g(x)$ as x tends to a does not exist (because in this case $f(x)/g(x)$ will become arbitrarily large in size as x tends to a).

If $m = 0$ and $l = 0$ then it is not clear (in general) what happens (indeed the limit may or may not exist, see Examples 3.2.1 and 3.2.3). This is the 'interesting' case and one we meet when we discuss differentiation.

3.2.5 Example. Does $\displaystyle\lim_{x \to 1} \frac{x^2 - 1}{x^2 - 3x + 2}$ exist ?

[[If we take $f(x) = x^2 - 1$ and $g(x) = x^2 - 3x + 2$ then clearly $\lim_{x \to 1} f(x) = 0$ and $\lim_{x \to 1} g(x) = 0$ and so we have an interesting case. We must try to simplify the expressions so that we can see what's going on.]]

Note that

$$\frac{x^2 - 1}{x^2 - 3x + 2} = \frac{(x - 1)(x + 1)}{(x - 1)(x - 2)} = \frac{x + 1}{x - 2} \quad \text{if} \quad x \neq 1$$

(we remark that $f(x)/g(x)$ is not defined when $x = 1$ but this is of no concern since the limit, should it exist, depends only on the values of x close to 1 but not equal to 1). Hence

$$\frac{x^2 - 1}{x^2 - 3x + 2} = \frac{x + 1}{x - 2} \to \frac{1 + 1}{1 - 2} = -1 \quad \text{as} \quad x \to 1,$$

so the limit does exist and

$$\lim_{x \to 1} \frac{x^2 - 1}{x^2 - 3x + 2} = -1\ .$$

□

3.2.6 Example. Does $\displaystyle\lim_{x \to 1} \frac{(x^2 - x)}{(x - 1)^2}$ exist ?

Note that
$$\frac{(x^2-x)}{(x-1)^2} = \frac{x(x-1)}{(x-1)(x-1)} = \frac{x}{x-1} \quad \text{if} \quad x \neq 1$$
Thus we see that $|(x^2-x)/(x-1)^2|$ becomes arbitrarily large as x tends to 1 and so the limit does not exist.

3.2.7 Example. Does $\displaystyle\lim_{h\to 0} \frac{\sqrt{2+h}-\sqrt{2}}{h}$ exist ?

[[If we take $f(h) = \sqrt{2+h}-\sqrt{2}$ and $g(h) = h$ then $\displaystyle\lim_{h\to 0} f(h) = 0$ and $\displaystyle\lim_{h\to 0} g(h) = 0$ so we have another interesting case.]]

If $h \neq 0$ then
$$\begin{aligned}\frac{\sqrt{2+h}-\sqrt{2}}{h} &= \frac{(\sqrt{2+h}-\sqrt{2})(\sqrt{2+h}+\sqrt{2})}{h(\sqrt{2+h}+\sqrt{2})}\\ &= \frac{h}{h(\sqrt{2+h}+\sqrt{2})}\\ &= \frac{1}{\sqrt{2+h}+\sqrt{2}}\end{aligned}$$
(recall that $(a-b)(a+b) = a^2-b^2$) and so we see
$$\frac{\sqrt{2+h}-\sqrt{2}}{h} \to \frac{1}{\sqrt{2}+\sqrt{2}} = 2^{-\frac{3}{2}} \quad \text{as} \quad h \to 0.$$

□

3.2.2 Continuity

Most real functions that arise in engineering mathematics are *continuous* (at least on some interval of interest). Such functions are important because they have very nice, useful properties which can easily be established.

Let f be a function and let c be a number. We say that f is *continuous* at c if
$$\lim_{x\to c} f(x) = f(c).$$
We say that f is continuous *on an interval* of real numbers if it is continuous at each number (point) in the interval.

[Technical Remark: Our definition of 'continuity on an interval' is slightly stronger than the 'usual definition'. However our definion and the usual one agree for open intervals.]

Polynomial functions are continuous on $\mathbb{R}$. Rational functions are continuous at every point of $\mathbb{R}$ except the points where the denominator is zero. They are continuous on every interval between the points where the denominator is zero.

The functions *sin* and *cos* are continuous at every real number and thus continuous on $\mathbb{R}$.

If $\cos(c) \neq 0$ then the properties of limits show that

$$\lim_{x\to c} \tan(x) = \lim_{x\to c} \frac{\sin(x)}{\cos(x)} = \frac{\sin(c)}{\cos(c)} = \tan(c).$$

Thus *tan* is continuous at every number except those numbers where *cos* is zero. These numbers are

$$\ldots, \ -\tfrac{3}{2}\pi, \ -\tfrac{1}{2}\pi, \ \tfrac{1}{2}\pi, \ \tfrac{3}{2}\pi, \ \tfrac{5}{2}\pi, \ \ldots.$$

Thus *tan* is continuous on any interval between these numbers.

Intuitively, if you can sketch the graph of a function without taking your pencil off the piece of paper then the function is continuous.

3.2.8 Example. Define the function u by

$$u(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

(u is called the step function). Try to see if you can sketch the graph without taking your pencil from the paper!!! In this case $u(x)$ tends to 0 as x tends to 0 from below but tends to 1 as x tends to 0 from above. Thus the limit of $u(x)$ as x tends to zero does not exist (for a limit to exist both the limits from above and below must exist and be equal). Hence u is *not* continuous at 0. □

3.3 Differentiation

3.3.1 Introduction

One of the oldest problems in mathematics, going back at least as far as the ancient Egyptians, is that of determining area: tax gatherers needed to know how much land people had so that they could tax them on it. So when, in the 17th century, mathematicians came up with the idea of coordinate geometry, and with it the idea of a curve given as $y = f(x)$, one of the questions that obsessed them was "How do you find the area under the curve ?".

A less obviously useful but still interesting problem is "Given a curve $y = f(x)$ and a point (a, b) on the curve, what is the equation of the tangent at the point?".

What the inventors of calculus discovered was:

- a method for solving the tangent problem;
- that the tangent problem was key to the area problem;
- that their technique for solving the tangent problem enabled them to deal with problems concerning variability, movement, etc.

Previously mathematicians had handled static problems; now they could tackle dynamic ones. And of course it is this that has made calculus so important in applications of mathematics, not least in engineering. □

3.3.2 Differentiability

The tangent problem is the easiest place to start. For the equation of a line you need to know either two points or one point and the slope. Here we have one point, and so somehow we have to calculate the slope.

Consider a curve in the (x, y)-plane. Suppose that the curve can be expressed as the graph $y = f(x)$ for some function f. Let $P = (a, f(a))$ be a point on the curve (see Figure 3.8). To get the equation of the tangent at P we need to calculate the slope of the tangent. Let m be the slope of the tangent at P.

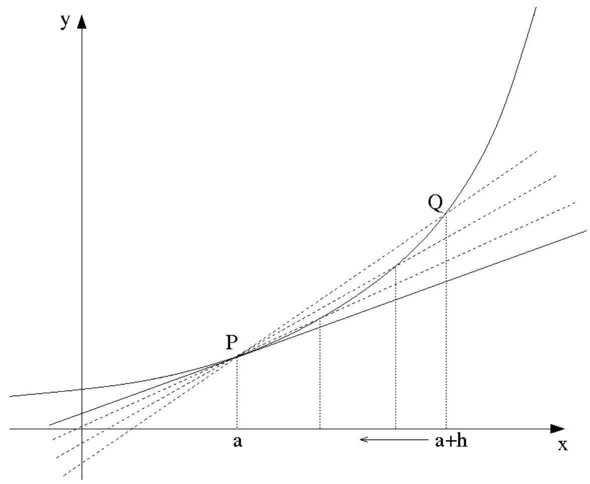

Figure 3.8: Graph $y = f(x)$

Let h be a small non-zero number and consider the point $Q = (a + h, f(a + h))$ so that Q is a point on the curve close to P. The slope of the chord PQ is

$$\frac{f(a+h) - f(a)}{(a+h) - a} = \frac{f(a+h) - f(a)}{h}$$

(this is called the Newton quotient of f at a.) In general, as h tends to zero from above the point Q approaches P along the curve and the slope of the chord PQ tends to the slope of the tangent to the curve at P (as indicated by the example in Figure 3.8). Likewise if h tends to 0 from below. Thus

$$\frac{f(a+h) - f(a)}{h} \to m \quad \text{as} \quad h \to 0 \ ,$$

In other words

Slope of the tangent to the curve $y = f(x)$ at the point $(a, f(a))$ is

$$\lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \ .$$

3.3.1 Example. Find the equation of the tangent to the curve given by $y = x^2$ at the point $(a,\, a^2)$ on the curve.

Solution. Let $f(x) = x^2$. Then if $h \neq 0$ we have

$$\begin{aligned}\frac{f(a+h)-f(a)}{h} &= \frac{(a+h)^2 - a^2}{h},\\ &= \frac{2ah + h^2}{h},\\ &= 2a + h.\end{aligned}$$

Hence

$$\begin{aligned}\text{Slope of tangent at the point } (a,\, a^2) &= \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}\\ &= \lim_{h\to 0}(2a+h)\\ &= 2a\end{aligned}$$

Consequently the equation of the tangent is $y - a^2 = 2a(x-a)$, which simplifies to $y = 2ax - a^2$. □

Let f be a (real) function. The *Newton quotient* of f at x is

$$\frac{f(x+h)-f(x)}{h}.$$

We say that f is *differentiable* at x if the limit of the Newton quotient of f at x as h tends to zero exists; in which case the limit is called the *derivative* of f at x and denoted by $f'(x)$. Thus

$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$

($f'(x)$ is read as 'f dashed at x'). f' is a (real) function, its domain is the set of all numbers x at which the limit above exists.

3.3.2 Remark. Remember h is a 'dummy' variable it can be replaced by any symbol or symbols we like. Leibniz used δx (read as 'delta x') to denote a small change (increment) in x and so h is often replaced by δx as discussed in Section 3.3.3. We could write

$$f'(x) = \lim_{\delta x\to 0} \frac{f(x+\delta x)-f(x)}{\delta x}\ .$$

□

In view of our earlier work we can now state the following result.

If f is differentiable at a then the slope of the tangent to the graph $y = f(x)$ at the point $(a,\, f(a))$ is $f'(a)$.

To find tangent to the graph of a function we usually calculate the derivative of the function from which we can get the slope and hence equation of the tangent. □

Differentiation is the process of calculating a derivative. In the next two very simple examples, we calculate the derivatives of two functions from *first principles* that is by evaluating the limit of Newton quotient directly. (Calculating derivatives from first principles can be quite difficult. Later we will develop techniques for calculating derivatives which avoids this.)

3.3.3 Example. $f(x) = c$ (a constant).

From first principles: if $h \neq 0$ then

$$\frac{f(x+h)-f(x)}{h} = \frac{c-c}{h} = \frac{0}{h} = 0, \quad \text{therefore} \quad \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = 0$$

(for any value of x). Thus f is differentiable at all points and $f'(x) = 0$. □

3.3.4 Example. $I(x) = x$ (identity map). From first principles: if $h \neq 0$ then

$$\frac{I(x+h)-I(x)}{h} = \frac{(x+h)-x}{h} = \frac{h}{h} = 1, \quad \text{therefore} \quad \lim_{h\to 0}\frac{I(x+h)-I(x)}{h} = 1$$

(for any value of x). Thus I is differentiable at all points and $I'(x) = 1$. □

Another simple example:

3.3.5 Example. The case $f(x) = x^2$ was dealt with in Example 3.3.1. There we calculated the derivative of f at a and showed that $f'(a) = 2a$. The same calculation holds for any number a. Thus we have $f'(x) = 2x$ for any value of x. □

3.3.3 Leibniz's notation

Let x and y be (real) variables and suppose that y depends on x,

$$y = f(x) \quad \text{say.}$$

Suppose that f is differentiable at all values of x so we can introduce the *variable*

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

dy/dx is called the derivative of y with respect to x. We sometimes write

$$\frac{d}{dx}(f(x)) \quad \text{for} \quad f'(x).$$

The notation is due to Leibniz.

3.3.6 Example. If $y = c$ where c is a constant then from example 3.3.3 (taking f(x)=c) we see that

$$\frac{dy}{dx} = \frac{d}{dx}(c) = 0.$$

(the derivative of a constant is zero). □

3.3.7 Example. If $y = x$ then from example 3.3.4 ($y = I(x)$) we see that

$$\frac{dy}{dx} = \frac{d}{dx}(x) = 1.$$

□

3.3.8 Example. If $y = x^2$ then from example 3.3.5 (taking $f(x) = x^2$) we see that

$$\frac{dy}{dx} = \frac{d}{dx}(x^2) = 2x \quad .$$

□

Next we explain where the notation comes from.

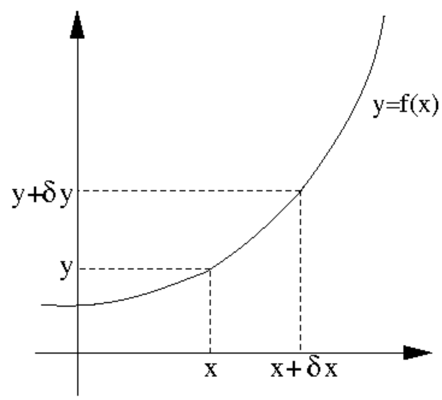

Figure 3.9: Graph of y against x

For each value of x consider the change δy in y caused by a small change δx in x, clearly

$$\delta y = f(x + \delta x) - f(x).$$

(As mentioned in Remark 3.3.2, Leibniz used δu to denote a 'small change' in a variable u, often referred to as an increment of u.) Thus

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

Assuming f is differnetiable at x

$$\lim_{\delta x \to 0} \frac{\delta y}{\delta x} = \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x} = f'(x)$$

and so from the definition of dy/dx we see

$$\frac{dy}{dx} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x}$$

which gives some explanation as to where the 'd/dx' notation comes from.

3.4 Standard Derivatives

3.4.1 Motivation

Calculating a derivative from *first principles*, that is by evaluating the limit of a Newton quotient as h tends to zero, can be time consuming and should be avoided if possible. In practice, based on a list of known derivatives, the derivatives of more complicated functions can be calculated using certain basic rules. Thus we want:

- to build up a list of the derivatives of commonly occurring functions;
- to develop a set of rules which allow the calculation of derivatives of more complicated functions by using known derivatives from the list.

In the next subsections we start to build our list of standard derivatives.

3.4.2 Derivative of x^n

First consider the case $n = 4$.

3.4.1 Example. Let $f(x) = x^4$. If $h \neq 0$ then

$$\begin{aligned}\frac{f(x+h) - f(x)}{h} &= \frac{(x+h)^4 - x^4}{h} \\ &= \frac{(x^4 + 4x^3h + 6x^2h^2 + 4xh^3 + h^4) - x^4}{h} \\ &= \frac{4x^3h + 6x^2h^2 + 4xh^3 + h^4}{h} \\ &= 4x^3 + h(6x^2 + 4xh + h^2).\end{aligned}$$

As $h \to 0$, $4x^3 + h(6x^2 + 4xh + h^2) \to 4x^3$ (for any value of x) and so $f'(x) = 4x^3$.

In Leibniz's notation: $\dfrac{d}{dx}(x^4) = 4x^3$. □

A similar argument works for any positive integer n. [Using the Binomial Theorem to expand $(x+h)^n$, the h from the bottom line of the corresponding Newton quotient will then cancel into the top, leaving you with an expression of the form $nx^{n-1} + h(\textit{something})$. The $h(\textit{something})$ then becomes negligible as h tends to 0, leaving nx^{n-1} as the derivative.]

The first derivative in our list is:

$$\frac{d}{dx}(x^n) = nx^{n-1} \quad \text{if } n \text{ is a positive integer}$$

(the formula holds for any value of x). Since $x^0 = 1$ the formula holds if $n = 0$ and, as we will see, the formula holds if n is negative *and* x is non-zero.

3.4.2 Example. Taking $n = -2$

$$\frac{d}{dx}\left(\frac{1}{x^2}\right) = \frac{d}{dx}\left(x^{-2}\right) = (-2)x^{-2-1} = -2x^{-3} = -\frac{2}{x^3} \quad (x \neq 0),$$

We remark that x^{-2} is not defined when $x = 0$. □

In fact the formula holds for *all* powers of x provided we restrict to positive values of x. The result is:

$$\frac{d}{dx}\left(x^a\right) = ax^{a-1}, \quad (x > 0).$$

3.4.3 Example. For $a = \frac{1}{2}$:

$$\frac{d}{dx}\left(\sqrt{x}\right) = \frac{d}{dx}\left(x^{\frac{1}{2}}\right) = \frac{1}{2}x^{\frac{1}{2}-1} = \frac{1}{2}x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}}, \quad (x > 0).$$

We remark that $\sqrt{x}$ is not defined when $x < 0$. □

3.4.3 Derivatives of sine and cosine

These are quite tricky to calculate from first principles. You need to use some of the trig formulae and then to do some delicate arguments with areas of segments of circles to justify the results. We quote the results without proof.

$$\frac{d}{dx}(\sin x) = \cos x, \qquad \frac{d}{dx}(\cos x) = -\sin x.$$

Recall: $\sin x$ is the sine of an angle x *radians*. These results are not true for the sine of an angle of x degrees.

More standard derivatives appear in the Examination Handbook (p. 16).

3.5 Rules for Differentiation

3.5.1 Preliminaries

The rules discussed here are the means by which you work out derivatives of functions which can be built up from functions in the 'standard list' by adding, multiplying, and so on. There are four rules. They are used so often that you must memorise them and practice using them. *It is pointed out that the rules do NOT appear in the Examination Handbook.*

Recall that given any two functions f and g we can form new functions $f+g$ and $f.g$, called the *sum* and *product* of f and g repsectively. Formally the definitions are

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

(that is the value of $f+g$ at x is the sum of the values of f and g at x) and

$$f.g(x) = f(x)g(x)$$

(that is the value of $f.g$ at x is the product of the values of f and g at x).

The domain of the sum or product consists of all numbers which lie in both the domain of f and the domain of g. The definitions extend to the sum or product of any number of functions.

Usually we write just fg for $f.g$ in future.

We can also form the quotient of two functions. For two functions f and g the quotient of f over g, denoted by $\dfrac{f}{g}$ (or f/g) is defined by

$$\frac{f}{g}(x) = \frac{f(x)}{g(x)},$$

the domain being all numbers in both the domain of f and the domain of g except those numbers at which g takes the value 0 (we cannot divide by zero).

3.5.2 Sum, product and quotient rules

Let f and g be functions. For any value of x at which both f and g are differentiable the following results hold.

Sum Rule.

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

equivalently, using Leibniz's notation,

$$\frac{d}{dx}\big(f(x) + g(x)\big) = \frac{d}{dx}\big(f(x)\big) + \frac{d}{dx}\big(g(x)\big)$$

or, in words, the derivative of a sum is the sum of the derivatives.

Product Rule.

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

equivalently, using Leibniz's notation,

$$\frac{d}{dx}\big(f(x)g(x)\big) = \frac{d}{dx}\big(f(x)\big)\,g(x) + f(x)\,\frac{d}{dx}\big(g(x)\big)$$

or, in words, the derivative of a product is the derivative of the first times the second plus the first times the derivative of the second.

Quotient Rule. If $g(x) \neq 0$ then

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2},$$

equivalently, using Leibniz's notation,

$$\frac{d}{dx}\Big(\frac{f(x)}{g(x)}\Big) = \frac{\frac{d}{dx}\big(f(x)\big)\,g(x) - f(x)\,\frac{d}{dx}\big(g(x)\big)}{g(x)^2}$$

or, in words, the derivative of a quotient is the derivative of the first times the second minus the first times the derivative of the second all over the second squared.

3.5.1 Example. Find the derivative of $x^3 + \sin x$ with respect to x.

Solution. Using the sum rule and our list of standard derivatives:

$$\begin{aligned}\frac{d}{dx}(x^3 + \sin x) &= \frac{d}{dx}(x^3) + \frac{d}{dx}(\sin x)\\ &= 3x^2 + \cos x.\end{aligned}$$

□

3.5.2 Example. Find the derivative of $t^3 \sin t$ with respect to t.

Solution. Using the product rule and our list of standard derivatives:

$$\begin{aligned}\frac{d}{dt}(t^3 \sin t) &= \frac{d}{dt}(t^3)\ \sin t + t^3\,\frac{d}{dt}(\sin t)\\ &= 3t^2 \sin t + t^3 \cos t.\end{aligned}$$

□

3.5.3 Example. Find the derivative of $\dfrac{y^3}{\sin y}$ with respect to y.

Solution. Using the quotient rule and our list of standard derivatives:

$$\begin{aligned}\frac{d}{dy}\left(\frac{y^3}{\sin y}\right) &= \frac{\frac{d}{dy}(y^3)\ \sin y - y^3\,\frac{d}{dy}(\sin y)}{\sin^2 y}\\ &= \frac{3y^2 \sin y - y^3 \cos y}{\sin^2 y} \qquad \text{(provided } \sin y \neq 0\text{)}.\end{aligned}$$

□

For variables u and v which depend on x, the rules can be expressed as

$$\frac{d}{dx}(u + v) = \frac{du}{dx} + \frac{dv}{dx},$$

$$\frac{d}{dx}(uv) = \frac{du}{dx}\,v + u\,\frac{dv}{dx},$$

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{\frac{du}{dx}\,v - u\,\frac{dv}{dx}}{v^2} \quad \text{(provided } v \neq 0\text{)},$$

which are a little easier to remember.

The derivative of cg where c is a constant is cg' because, by the product rule,

$$(cg)' = c'g + cg', \quad \text{and so} \quad (cg)' = cg'$$

because the derivative of a constant is zero. In the case of a variable u

$$\frac{d}{dx}(cu) = c\,\frac{du}{dx}.$$

3.5.4 Example. Find the derivative of $5x^3 \sin x \cos x$ with respect to x.

Solution. First note that since 5 is a constant

$$\frac{d}{dx}(5x^3 \sin x \cos x) = 5\frac{d}{dx}(x^3 \sin x \cos x).$$

Using the product rule and our list of standard derivatives:

$$\begin{aligned}
\frac{d}{dx}(5x^3 \sin x \cos x) &= 5\frac{d}{dx}(x^3 \sin x \cos x) \\
&= 5\Big(\frac{d}{dx}(x^3).\sin x \cos x + x^3.\frac{d}{dx}(\sin x \cos x)\Big) \\
&= 5\Big(\frac{d}{dx}(x^3).\sin x \cos x + x^3\Big(\frac{d}{dx}(\sin x).\cos x + \sin x.\frac{d}{dx}(\cos x)\Big)\Big) \\
&= 5\big(3x^2 \sin x \cos x + x^3 \cos x \cos x - x^3 \sin x \sin x\big) \\
&= 5x^2\big(3 \sin x \cos x + x(\cos^2 x - \sin^2 x)\big)
\end{aligned}$$

□

The sum rule extends in an obvious way to the sum of more than two functions, for instance the derivative of $f + g + h$ is $f' + g' + h'$.

The product rule extends in an obvious way to the product of more than two functions, for instance

$$(fgh)' = f'gh + f(gh)' = f'gh + fg'h + fgh'.$$

For variables u, v and w which depend on x:

$$\frac{d}{dx}(uvw) = \frac{du}{dx}\,vw + u\,\frac{dv}{dx}\,w + uv\,\frac{dw}{dx}.$$

In practice, to differentiate a product of three or more functions, it is advisable to do it in stages as in Example 3.5.4 rather than by memorising the formula above.

Using the quotient rule we can expand our list of standard derivatives by getting the derivatives of the other trig functions.

3.5.5 Example. Find the derivative of $\tan x$ with respect to x.

Solution. Recall that $\tan x = \dfrac{\sin x}{\cos x}$. Therefore

$$\begin{aligned}\frac{d}{dx}(\tan x) &= \frac{\frac{d}{dx}(\sin x)\ \cos x - \sin x\,\frac{d}{dx}(\cos x)}{(\cos x)^2} \\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} \\ &= \sec^2 x \quad \text{provided } \ \cos x \neq 0.\end{aligned}$$

□

The derivatives of $\cot x$, $\sec x$ and $\operatorname{cosec} x$ can be calculated in similar fashion. The result is the following table of derivatives (these can be found in the Engineering Mathematics Handbook):

$$\frac{d}{dx}(\tan x) = \sec^2 x \qquad \frac{d}{dx}(\cot x) = -\operatorname{cosec}^2 x$$

$$\frac{d}{dx}(\sec x) = \sec x \tan x \qquad \frac{d}{dx}(\operatorname{cosec} x) = -\operatorname{cosec} x \cot x$$

We claimed earlier that $\dfrac{d}{dx}(x^n) = nx^{n-1}$ for any integer n and justified the result in the case n is non-negative. We can now show that the result is true if n is negative.

Suppose that $n < 0$ so that $m = -n > 0$. Then

$$x^n = x^{-m} = \frac{1}{x^m} \quad (x \neq 0).$$

By the quotient rule

$$\begin{aligned}\frac{d}{dx}(x^n) = \frac{d}{dx}\left(\frac{1}{x^m}\right) &= \frac{\frac{d}{dx}(1)\,x^m - 1\,\frac{d}{dx}(x^m)}{(x^m)^2} \\ = \frac{-\frac{d}{dx}(x^m)}{(x^m)^2} = \frac{-mx^{m-1}}{x^{2m}} &= -mx^{-m-1} = nx^{n-1} \quad (x \neq 0).\end{aligned}$$

□

3.5.3 Chain rule

Let g and h be functions, we can consider the function f given by

$$f(x) = g(h(x)),$$

that is the value of f at x is the value of g at $h(x)$. f is called the *composition* of g and h and often denoted by $g \circ h$ (*do not confuse this with the product $g.h$ of g and h*).

3.5.6 Example. $f(x) = \sin(x^3)$. Taking $g(u) = \sin u$ and $h(x) = x^3$ we see that $f(x) = g(h(x))$ so that f is the composition of g and h. $\square$

3.5.7 Example. $f(x) = \sqrt{1 + \sin x}$ is the composition of $g(u) = \sqrt{u}$ and $h(x) = 1 + \sin x$. $\square$

Assume that g is differentiable at $h(x)$ and h is differentiable at x.

Chain Rule (Newton's notation).

$$f'(x) = g'(h(x))\, h'(x),$$

or

$$(g \circ h)' = (g' \circ h)\, h'.$$

for short.

This is the Newtonian way of writing the chain rule. An alternative formulation, using Leibniz's notation is much easier to use and remember. Let

$$y = f(x) = g(h(x)).$$

The idea is to introduce another variable $u = h(x)$ then

$$y = g(u) \quad \text{where} \quad u = h(x),$$

$$\frac{dy}{du} = g'(u) \quad \text{and} \quad \frac{du}{dx} = h'(x).$$

By the chain rule

$$\frac{dy}{dx} = f'(x) = g'(h(x))\, h'(x) = g'(u)h'(x) = \frac{dy}{du}.\frac{du}{dx}.$$

Chain Rule (Leibniz's notation). If $y = g(u)$ where $u = h(x)$ then

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

[To remember the rule you just 'imagine' dy, du and dx as being 'infinitesimally' small quantities in their own right and then visualise the 'cancellation' of the du.]

3.5.8 Example. Differentiate $\dfrac{1}{\sqrt{\sin x}}$ with respect to x.

Solution. Let $y = \dfrac{1}{\sqrt{\sin x}}$ so that $y = \dfrac{1}{\sqrt{u}} = u^{-\frac{1}{2}}$ where $u = \sin x$. Then

$$\begin{aligned} \frac{dy}{dx} &= \frac{dy}{du}.\frac{du}{dx} \\ &= -\frac{1}{2}u^{-\frac{3}{2}}.\cos x \\ &= -\frac{1}{2}(\sin x)^{-\frac{3}{2}}.\cos x \\ &= -\frac{\cos x}{2(\sin x)^{\frac{3}{2}}} \end{aligned}$$

The calculations are valid as long as $\sin x > 0$. $\square$

3.5.9 Example. Differentiate $(t^2 + \tan t)^8$ with respect to t.

Solution. Put $x = (t^2 + \tan t)^8$ then

$$x = z^8 \quad \text{where} \quad z = t^2 + \tan t.$$

By the chain rule

$$\begin{aligned} \frac{dx}{dt} &= \frac{dx}{dz}.\frac{dz}{dt} \\ &= 8z^7.(2t + \sec^2 t) \\ &= 8(t^2 + \tan t)^7(2t + \sec^2 t) \end{aligned}$$

After a little practice there is no need to introduce the variables x and z:

$$\frac{d}{dt}(t^2 + \tan t)^8 = 8(t^2 + \tan t)^7.\frac{d}{dt}\big(t^2 + \tan t\big) = 8(t^2 + \tan t)^7(2t + \sec^2 t)$$

□

The Leibniz version also makes it easy to deal with longer chains.

Suppose that $y = g(h(k(x)))$. Introduce a new variable $u = h(k(x))$ so that $y = g(u)$. Introduce another variable $v = k(x)$ so that $u = h(v)$. Thus

$$y = g(u) \quad \text{where} \quad u = h(v) \quad \text{and} \quad v = k(x)\ .$$

Hence

$$\frac{dy}{dx} = \frac{dy}{du}.\frac{du}{dx} = \frac{dy}{du}.\frac{du}{dv}.\frac{dv}{dx}$$

so that

$$\frac{dy}{dx} = \frac{dy}{du}.\frac{du}{dv}.\frac{dv}{dx}$$

(one can imagine both the 'du' and the 'dv' cancelling).

3.5.10 Example. Differentiate $\sin(\sin(1 + x^2))$ with respect to x.

Solution. Put $y = \sin(\sin(1 + x^2))$. Break y down as

$$y = \sin u \quad \text{where} \quad u = \sin(1 + x^2),$$

and break down further as

$$y = \sin u \quad \text{where} \quad u = \sin v \quad \text{and} \quad v = 1 + x^2.$$

Then

$$\begin{aligned} \frac{dy}{dx} &= \frac{dy}{du}.\frac{du}{dv}.\frac{dv}{dx} \\ &= \cos u.\cos v.2x \\ &= \cos(\sin v).\cos(1 + x^2).2x \\ &= \cos(\sin(1 + x^2)).\cos(1 + x^2).2x \end{aligned}$$

□

With complicated functions you often have to use several of the four rules in combination.

3.5.11 Exercise. Differentiate $\dfrac{x^2 \cos\left(x^{\frac{3}{2}}\right)}{\tan x}$ with respect to x. $\square$

We have not justified the rules of differentiation. Next we will justify the product rule (the justification of the other rules are just as 'simple'). You should be able to get the justifications of the other rules from the internet if you are interested.

[[Let f anf g be (real) functions. For any value of x at which both f and g are differentiable we have:

$$\begin{aligned}\frac{(f.g)(x+h)-(f.g)(x)}{h} &= \frac{f(x+h)g(x+h)-f(x)g(x)}{h} \\ &= \frac{f(x+h)g(x+h)-f(x)g(x+h)+f(x)g(x+h)-f(x)g(x)}{h} \\ &= \frac{(f(x+h)-f(x))}{h}\,.\,g(x+h)+f(x)\,.\,\frac{(g(x+h)-g(x))}{h}\end{aligned}$$

but we know

$$\frac{f(x+h)-f(x)}{h} \to f'(x) \quad \text{and} \quad \frac{g(x+h)-g(x)}{h} \to g'(x) \quad \text{as} \quad h \to 0$$

(by the differentiability of f and g). Also note that

$$g(x+h) = \frac{(g(x+h)-g(x))}{h}\,.\,h + g(x) \to g'(x)\,.\,0 + g(x) \quad \text{as} \quad h \to 0\ .$$

By the properties of limits we see

$$\frac{(f.g)(x+h)-(f.g)(x)}{h} \to f'(x)g(x)+f(x)g'(x) \quad \text{as} \quad h \to 0$$

and so $f\,.\,g$ is differentiable at x and

$$(f\,.\,g)'(x) = f'(x)g(x)+f(x)g'(x) \qquad //$$

which justifies the product rule.]] $\blacksquare$