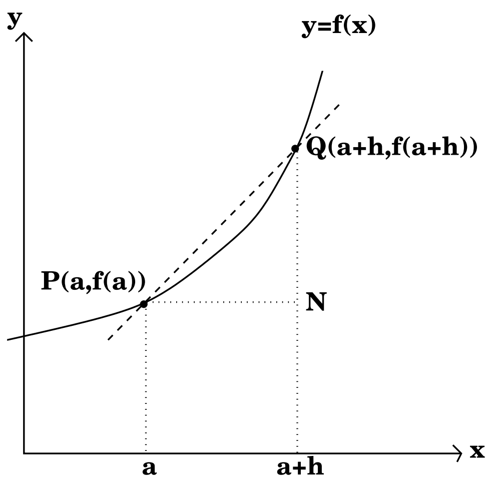
11 Derivatives

As you will remember from school or college mathematics, the derivative of a function tells us how rapidly the function is increasing or decreasing. We can think of it as measuring how steep the graph of the function is. We will now take this intuitive picture and make it mathematically precise.

11.1 Introduction and definition

Consider a function f, and take two points $\mathrm{P}:(a, f(a))$ and $\mathrm{Q}:(a+h, f(a+h))$ on the graph of f.



The slope of the secant line PQ is

$$\frac{f(a+h)-f(a)}{a+h-a}=\frac{f(a+h)-f(a)}{h}. \tag{11.1}$$

Intuitively, as Q approaches P, $h \to 0$ and (11.1) approaches the slope of the tangent line at P. This limiting process will work only if f is 'smooth' enough at a; it fails if f is not continuous or has 'corners'. (An example of a corner is the one $f(x) = |x|$ has at 0.) This notion of smoothness is so important that we make the following definition.

Definition 11.1 *The function f is **differentiable** at a if*

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

*If $f'(a)$ exists for all points a in an open interval I we call f **differentiable in** I.*

Definition 11.2 *Given a function f which is differentiable at a point a, the* ***derivative*** *of f at a is defined as*

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

This defines a new function f' whose value at a is the derivative of f at a.

The tangent line to the graph of f at a exists if and only if $f'(a)$ exists (i.e. if f is differentiable at a), and it is then the straight line through the point $(a, f(a))$ with slope $f'(a)$.

It is sometimes more convenient to use an alternative form of the limit in Definitions 11.1 and 11.2. If we define $b := a + h$ we have $h = b - a$ and then the slope in (11.1) can be written as

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

If f is differentiable at a, then

$$f'(a) = \lim_{b \to a} \frac{f(b) - f(a)}{b - a} \quad \left(= \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \right). \tag{11.2}$$

11.2 Elementary results for derivatives

In many cases, the derivative can be found by reverting to the definition and computing explicitly the limit $\lim_{h \to 0} \dfrac{f(a+h) - f(a)}{h}$.

Theorem 11.1 *Some elementary derivatives:*

1. *If $f(x) = c$ (where c is a constant) then $f'(a) = 0$.*
2. *If $f(x) = cx + d$ (where c and d are constants) then $f'(a) = c$.*
3. *If $f(x) = x^2$ (constant) then $f'(a) = 2a$.*
4. *If $f(x) = x^3$ (constant) then $f'(a) = 3a^2$.*

Proof: In each case the result follows from Definition 11.2 followed by some algebraic manipulation and Theorem 9.2.

1. $$f'(a) = \lim_{h \to 0} \frac{c - c}{h} = \lim_{h \to 0} \frac{0}{h} = 0.$$
2. $$f'(a) = \lim_{h \to 0} \frac{c(a+h) + d - (ca + d)}{h} = \lim_{h \to 0} \frac{ch}{h} = \lim_{h \to 0} c = c.$$

3.

$$f'(a) = \lim_{h\to 0} \frac{(a+h)^2 - a^2}{h} = \lim_{h\to 0} \frac{a^2 + 2ah + h^2 - a^2}{h} = \lim_{h\to 0}(2a+h) = 2a.$$

4.

$$f'(a) = \lim_{h\to 0} \frac{(a+h)^3 - a^3}{h} = \lim_{h\to 0} \frac{3a^2h + 3ah^2 + h^3}{h} = \lim_{h\to 0}(3a^2 + 3ah + h^2) = 3a^2.$$

□

We can also obtain some general results when we know the properties of a function but don't have a specific formula for it.

Example 11.1 Find the derivative of an arbitrary linear function f, that is, of a function for which $f(x+y) = f(x) + f(y)$ and $f(cx) = cf(x)$ for all x, y, and c.

By definition,

$$\begin{aligned}
f'(a) &= \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} && \text{(Definition 11.2)} \\
&= \lim_{h\to 0} \frac{f(a) + f(h) - f(a)}{h} && \text{(first property of a linear function)} \\
&= \lim_{h\to 0} \frac{f(h)}{h} && \text{(simplifying)} \\
&= \lim_{h\to 0} \frac{hf(1)}{h} && \text{(second property of a linear function)} \\
&= f(1).
\end{aligned}$$

(Note: this result could also be obtained as a special case of Theorem 11.1(2), since we already know that all linear functions on $\mathbb{R}$ are of the form $f(x) = \alpha x$ with some constant α, and that $f(1) = \alpha$.)

Example 11.2 Prove that the derivative of an even function f is an odd function.

An even function f satisfies $f(-x) = f(x)$ for all x. We have

$$\begin{aligned}
f'(-a) &= \lim_{h\to 0} \frac{f(-a+h) - f(-a)}{h} = \lim_{h\to 0} \frac{f(a-h) - f(a)}{h} = -\lim_{h\to 0} \frac{f(a-h) - f(a)}{-h} \\
&= -\lim_{k\to 0} \frac{f(a+k) - f(a)}{k} = -f'(a),
\end{aligned}$$

where we have set $k = -h$ and used the fact that

$$h \to 0 \iff -h \to 0.$$

So $f'(-a) = -f'(a)$ for all a, which shows that f' is odd.

However, for more complicated functions, finding the limit may be very hard work. We consider a few examples.

Example 11.3 Find the derivative of $f(x) = 1 + 2x + 3x^2$.

$$\begin{aligned} f'(a) &= \lim_{h\to 0} \frac{\{[1+2(a+h)+3(a+h)^2] - [1+2a+3a^2]\}}{h} \\ &= \lim_{h\to 0} \frac{(2h+6ah+3h^2)}{h} \\ &= \lim_{h\to 0}(2+6a+3h) \\ &= 2+6a. \end{aligned}$$

Example 11.4 Find the derivative of $f(x) = \sqrt{x}$.

An algebraic 'trick' is required:

$$\begin{aligned} f'(a) &= \lim_{h\to 0} \frac{\sqrt{a+h}-\sqrt{a}}{h} \\ &= \lim_{h\to 0} \frac{(\sqrt{a+h}-\sqrt{a})(\sqrt{a+h}+\sqrt{a})}{h(\sqrt{a+h}+\sqrt{a})} \\ &= \lim_{h\to 0} \frac{a+h-a}{h(\sqrt{a+h}+\sqrt{a})} \\ &= \lim_{h\to 0} \frac{1}{\sqrt{a+h}+\sqrt{a}} = \frac{1}{2\sqrt{a}} \end{aligned}$$

Example 11.5 Find the derivative of $f(x) = \dfrac{1}{\sqrt{x}}$.

A similar trick is required:

$$\begin{aligned}
f'(a) &= \lim_{h\to 0}\frac{1}{h}\left[\frac{1}{\sqrt{a+h}}-\frac{1}{\sqrt{a}}\right] \\
&= \lim_{h\to 0}\frac{1}{h}\left[\frac{\sqrt{a}-\sqrt{a+h}}{\sqrt{a}\sqrt{a+h}}\right] \\
&= \lim_{h\to 0}\frac{1}{h}\left[\frac{a-(a+h)}{\sqrt{a}\sqrt{a+h}(\sqrt{a}+\sqrt{a+h})}\right] \\
&= -\lim_{h\to 0}\frac{1}{\sqrt{a}\sqrt{a+h}(\sqrt{a}+\sqrt{a+h})} \\
&= -\frac{1}{\sqrt{a}\sqrt{a}\,(2\sqrt{a})} \\
&= -\frac{1}{2}a^{-3/2}.
\end{aligned}$$

Example 11.6 Find the derivative of $f(x)=\sin x$.

$$\begin{aligned}
(\sin)'(a) &= \lim_{h\to 0}\frac{\sin(a+h)-\sin a}{h} \\
&= \lim_{h\to 0}\frac{\sin a\cos h+\sin h\cos a-\sin a}{h} \\
&= \sin a\,\lim_{h\to 0}\frac{\cos h-1}{h}+\cos a\,\lim_{h\to 0}\frac{\sin h}{h} \\
&= \cos a.
\end{aligned}$$

Here we have used the two-angle formula for $\sin(a+h)$ and two results from our chapter on limits, $\lim_{h\to 0}\frac{\sin h}{h}=1$ and $\lim_{h\to 0}\frac{\cos h-1}{h}=0$. (See (9.2)) and ((9.3)).)

When we drew secants and tangents to the graph of f we assumed that f was continuous. The following theorem shows that continuity is indeed necessary (though by no means sufficient) for differentiability.

Theorem 11.2 *If f is differentiable at a, then f is continuous at a.*

Proof: We need to show that $\lim_{x\to a} f(x)=f(a)$. This is equivalent to showing that $\lim_{h\to 0} f(a+h)=f(a)$. (Here h can be both positive or negative, so $x=a+h$ can be either greater or less than a.)

$$\begin{aligned}
\lim_{h\to 0}(f(a+h)-f(a)) &= \lim_{h\to 0}\frac{f(a+h)-f(a)}{h}\cdot h \\
&= \lim_{h\to 0}\frac{f(a+h)-f(a)}{h}\cdot\lim_{h\to 0}h = f'(a)\cdot 0 = 0.
\end{aligned}$$

So $\lim_{h\to 0}(f(a+h)-f(a))=0$, which implies $\lim_{h\to 0} f(a+h)=\lim_{h\to 0} f(a)=f(a)$. $\square$

Note that Theorem 11.2 implies that a function *cannot* be differentiable at points where it is not continuous.

11.3 Alternative notation for derivatives

A frequently used alternative notation for the derivative f' of a function f is

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

or simply, when there is no question about the argument of f,

$$\frac{\mathrm{d}f}{\mathrm{d}x}.$$

This can be a concise way to express derivatives without having to assign names to functions. So, instead of writing *"the derivative of the function $f(x)=x^2$ is $f'(x)=2x$"* we can simply state that $\frac{\mathrm{d}x^2}{\mathrm{d}x}=2x$ without the need to introduce a name for the function $x\mapsto x^2$.

Another advantage of this notation is that we can specify the name of the variable with respect to which differentiation is to be performed. For example, if $f(s)=s^2t$ and $g(t)=s^2t$ (two very different functions!) then $f'(s)=2st$ but $g'(t)=s^2$. These two results can be simply stated as

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

and

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

without any doubt as to what is the variable and what merely a parameter. We would refer to these derivatives respectively as "the derivative of s^2t with respect to s" and "the derivative of s^2t with respect to t".

The danger of this notation is that we can mistake $\mathrm{d}f/\mathrm{d}x$ for a fraction. This can cause problems because neither $\mathrm{d}f$ nor $\mathrm{d}x$ on its own has been defined, and trying to manipulate them as though they were numbers or variables can lead to all sorts of errors…

12 Differentiation

As we have seen in the preceding chapter, **differentiation**, the process of finding derivatives[2], can be rather cumbersome when we start afresh every time from the definition of the derivative. While there are certainly cases that require us to do this, it is possible to find the derivative of most functions knowing only the derivatives of a few basic functions, plus some elementary rules.

We already know a few derivatives from Section 11.2.

- If c is a constant then
$$\frac{\mathrm{d}c}{\mathrm{d}x} = 0.$$
- If f is a linear polynomial, $f(x) = cx + d$, then
$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = c.$$
- If $f(x) = x^2$ then
$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{\mathrm{d}x^2}{\mathrm{d}x} = 2x.$$
- If $f = \sin$ then
$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{\mathrm{d}\sin x}{\mathrm{d}x} = \cos x.$$

These results are enough to get us started. We now derive general rules that allow us to compute the derivatives of a wide range of functions.

12.1 Rules of differentiation

12.1.1 The linearity rules

We start with two theorems that allow us to differentiate linear combinations of simpler functions.

Theorem 12.1 *If f and g are differentiable at a then $f + g$ is differentiable at a and $(f + g)'(a) = f'(a) + g'(a)$.*

[2]Terminology: we *differentiate* a function to find its *derivative*. To "derive" a function means something different, and a "differential" is not the same thing as a derivative. This is not logical, but you should blame the English language, not us.

Proof:

$$\begin{aligned}
(f+g)'(a) &= \lim_{h\to 0}\frac{(f+g)(a+h)-(f+g)(a)}{h} \\
&= \lim_{h\to 0}\left(\frac{f(a+h)-f(a)}{h}+\frac{g(a+h)-g(a)}{h}\right) \\
&= f'(a)+g'(a)
\end{aligned}$$

□

Theorem 12.2 *If f is differentiable at a and $g(x) = cf(x)$ with $c \in \mathbb{R}$ constant then g is differentiable at a and $g'(a) = cf'(a)$.*

Proof:

$$\begin{aligned}
g'(a) &= \lim_{h\to 0}\frac{g(a+h)-g(a)}{h} \\
&= \lim_{h\to 0}\frac{cf(a+h)-cf(a)}{h} \\
&= c\lim_{h\to 0}\frac{f(a+h)-f(a)}{h} \\
&= cf'(a)
\end{aligned}$$

□

Example 12.1 Reconsider Example 11.3: find f' for $f(x) = 1+2x+3x^2$.

We have

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}(1+2x+3x^2) = \frac{\mathrm{d}}{\mathrm{d}x}1 + 2\frac{\mathrm{d}}{\mathrm{d}x}x + 3\frac{\mathrm{d}}{\mathrm{d}x}x^2 = 0 + 2(1) + 3(2x) = 2+6x.$$

12.1.2 The product rule

The derivative of the product of two functions is *not* simply the product of their derivatives. Rather, it is found by adding the two terms obtained by (1) computing the derivative of the first function and multiplying it by the second function (as if the second function were constant), and and (2) computing the derivative of the second function and multiplying it by the first function (as if the first function were constant).

Theorem 12.3 ***Product Rule.*** *If f and g are differentiable at a then fg is differentiable at a and $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$.*

Proof:

$$\begin{aligned}
(fg)'(a) &= \lim_{h\to 0} \frac{f(a+h)g(a+h) - f(a)g(a)}{h} \\
&= \lim_{h\to 0} \left(\frac{[f(a+h)g(a+h) - f(a)g(a+h)}{h} + \frac{f(a)g(a+h) - f(a)g(a)}{h} \right) \\
&= \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \lim_{h\to 0} g(a+h) + \lim_{h\to 0} f(a) \lim_{h\to 0} \frac{g(a+h) - g(a)}{h} \\
&= f'(a)g(a) + f(a)g'(a).
\end{aligned}$$

For the last line we have used Theorem 11.2 to tell us that since g is differentiable at a it must also be continuous at a, so $\lim_{h\to 0} g(a+h) = g(a)$. $\square$

We can use the product rule to re-derive some of our earlier results:

- $(cf)' = c'f + cf' = 0 \cdot f + cf' = cf'$.
- $\dfrac{\mathrm{d}}{\mathrm{d}x}x^2 = \dfrac{\mathrm{d}}{\mathrm{d}x}(x \cdot x) = \left(\dfrac{\mathrm{d}}{\mathrm{d}x}x\right) x + x\dfrac{\mathrm{d}}{\mathrm{d}x}x = 1 \cdot x + x \cdot 1 = 2x.$
- $\dfrac{\mathrm{d}}{\mathrm{d}x}x^3 = \dfrac{\mathrm{d}}{\mathrm{d}x}(x \cdot x^2) = \left(\dfrac{\mathrm{d}}{\mathrm{d}x}x\right) x^2 + x\dfrac{\mathrm{d}}{\mathrm{d}x}x^2 = x^2 + x \cdot 2x = 3x^2.$
- $\dfrac{\mathrm{d}}{\mathrm{d}x}x^4 = \dfrac{\mathrm{d}}{\mathrm{d}x}(x \cdot x^3) = \left(\dfrac{\mathrm{d}}{\mathrm{d}x}x\right) x^3 + x\dfrac{\mathrm{d}}{\mathrm{d}x}x^3 = x^3 + x \cdot 3x^2 = 4x^3.$

Notice the pattern. The general result is given by the following theorem.

Theorem 12.4 *If $f(x) = x^n$ for $n \in \mathbb{N}$ then*

$$f'(a) = na^{n-1} \quad \forall a.$$

Proof: We prove this by induction on n.

First, for $n = 1$ we have $f(x) = x$ and so $f'(a) = 1$. This is indeed $f'(a) = na^{n-1} = 1 \cdot a^0$.

Now assume that the result holds for a given $n \in \mathbb{N}$, that is, $\frac{\mathrm{d}}{\mathrm{d}x}x^n = nx^{n-1}$. We have to show that $\frac{\mathrm{d}}{\mathrm{d}x}x^{n+1} = (n+1)x^{(n+1)-1} = (n+1)x^n$. The product rule does the trick:

$$\frac{\mathrm{d}}{\mathrm{d}x}x^{n+1} = \frac{\mathrm{d}}{\mathrm{d}x}(x^n \cdot x) = \frac{\mathrm{d}}{\mathrm{d}x}(x^n)x + x^n\frac{\mathrm{d}}{\mathrm{d}x}x = nx^{n-1} \cdot x + x^n \cdot 1 = (n+1)x^n.$$

Thus, by mathematical induction, the result holds for all $n \in \mathbb{N}$ as required. $\square$

With the results proved so far we can differentiate any polynomial.

Example 12.2 Differentiate the polynomial functions

$$f(x) = x^5 - 5x^3 + x^2 + 1, \quad g(x) = x^6 - 1, \quad h(x) = x^{27} - 103x^{11} - 1734x.$$

We have

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}(x^5 - 5x^3 + x^2 + 1) = 5x^4 - 5(3x^2) + 2x + 0 = 5x^4 - 15x^2 + 2x.$$

Similarly,

$$g'(x) = \frac{\mathrm{d}}{\mathrm{d}x}(x^6 - 1) = \frac{\mathrm{d}}{\mathrm{d}x}(x^6) - \frac{\mathrm{d}}{\mathrm{d}x}(1) = 6x^5.$$

Finally, we have

$$h'(x) = \frac{\mathrm{d}}{\mathrm{d}x}(x^{27} - 103x^{11} - 1734x) = 27x^{26} - 1133x^{10} - 1734.$$

12.1.3 The quotient rule

Our next task is to find the derivative of the quotient f/g of two functions f and g. We start by computing the derivative of $1/g$.

Theorem 12.5 ***Reciprocal Rule*** *If g is differentiable at a and $g(a) \neq 0$ then $1/g$ is differentiable at a and*

$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{[g(a)]^2}.$$

Proof: A bit of care is needed because g is in the denominator. Clearly, $g(a) \neq 0$ is necessary, but we also require $g(a+h) \neq 0$, at least for sufficiently small values of h. Because g is differentiable at a, by Theorem 11.2 it is also continuous at a. Therefore, we can find an interval around a in which g does not vanish.

We can now go back to Definition 11.2 and write

$$\begin{aligned}
\left(\frac{1}{g}\right)'(a) &= \lim_{h\to 0} \frac{\left(\frac{1}{g}\right)(a+h) - \left(\frac{1}{g}\right)(a)}{h} \\
&= \lim_{h\to 0} \frac{1}{h}\left[\frac{1}{g(a+h)} - \frac{1}{g(a)}\right] \\
&= \lim_{h\to 0} \frac{1}{g(a)g(a+h)} \cdot \frac{g(a)-g(a+h)}{h} \\
&= -\frac{g'(a)}{[g(a)]^2}.
\end{aligned}$$

$\square$

With this result, finding the derivative of the quotient of two functions is now straightforward.

Theorem 12.6 ***Quotient Rule*** *If f and g are differentiable at a and $g(a) \neq 0$ then f/g is differentiable at a and*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{[g(a)]^2}.$$

Proof:

$$\begin{aligned}
\left(\frac{f}{g}\right)'(a) &= \left(f \cdot \frac{1}{g}\right)'(a) \\
&= f'(a) \cdot \left(\frac{1}{g}\right)(a) + f(a) \cdot \left(\frac{1}{g}\right)'(a) \\
&= \frac{f'(a)}{g(a)} + \frac{f(a)(-g'(a))}{[g(a)]^2} \\
&= \frac{f'(a)g(a) - f(a)g'(a)}{[g(a)]^2}.
\end{aligned}$$

$\square$

Remembering where the sign goes in the numerator of the quotient rule is quite easy if you bear in mind that the quotient rule has to hold when the denominator is the constant function $g(x) = 1$ for all x: in that case, $(f/g)' = (f/1)' = f'$.

Note that the reciprocal rule can be regarded as a special case of the quotient rule with f being the constant function $f(x) = 1$ for all x:

$$\left(\frac{1}{g}\right)' = \left(\frac{f}{g}\right)' = \frac{0 \cdot g - 1 \cdot g'}{g^2} = -\frac{g'}{g^2}.$$

Therefore, there is no need to memorise the reciprocal rule separately.

Example 12.3 Differentiate the functions

$$f(x) = \frac{x}{1+x^2}, \qquad g(x) = \frac{3x+1}{x^2-x+1}.$$

We have

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x}{1+x^2}\right) = \frac{\left(\frac{\mathrm{d}}{\mathrm{d}x}x\right)(1+x^2) - x\frac{\mathrm{d}}{\mathrm{d}x}(1+x^2)}{(1+x^2)^2} = \frac{(1+x^2)-x(2x)}{(1+x^2)^2} = \frac{1-x^2}{(1+x^2)^2}$$

and

$$\begin{aligned} g'(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{3x+1}{x^2-x+1}\right) = \frac{3(x^2-x+1)-(3x+1)(2x-1)}{(x^2-x+1)^2} \\ &= \frac{3x^2-3x+3-6x^2+x+1}{(x^2-x+1)^2} = -\frac{3x^2+2x-4}{(x^2-x+1)^2}. \end{aligned}$$

Example 12.4 Show that the derivative of $f(x) = x^{-n}$ for $n \in \mathbb{N}$ is $f'(x) = -nx^{-n-1}$.

We have $f(x) = x^{-n} = \dfrac{1}{x^n}$ and so

$$f'(x) = -\frac{nx^{n-1}}{(x^n)^2} = -\frac{nx^{n-1}}{x^{2n}} = -nx^{n-1-2n} = -nx^{-n-1}.$$

This is formally the same as our earlier result for positive exponents. Thus the formula given in Theorem 12.4 is valid for all $n \in \mathbb{Z}$, given that also $\dfrac{\mathrm{d}}{\mathrm{d}x}x^0 = \dfrac{\mathrm{d}}{\mathrm{d}x}1 = 0 = 0 \cdot x^{0-1}$.

12.1.4 The chain rule

We now learn how to compute the derivative of $f \circ g$ in terms of the derivatives of the functions f and g.

Theorem 12.7 ***Chain Rule*** *Consider two functions f and g with* $\mathrm{range}(g) \subseteq \mathrm{dom}(f)$. *If g is differentiable at a and f is differentiable at $y = g(a)$, then $f \circ g$ is differentiable at a and*

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a).$$

Proof: It is tempting to go straight ahead and write

$$\begin{aligned} \frac{f(g(a+h)) - f(g(a))}{h} &= \frac{f(g(a)+g(a+h)-g(a)) - f(g(a))}{g(a+h)-g(a)} \cdot \frac{g(a+h)-g(a)}{h} \\ &= \frac{f(g(a)+H) - f(g(a))}{H} \cdot \frac{g(a+h)-g(a)}{h}, \end{aligned}$$

with $H = g(a+h) - g(a) \to 0$ as $h \to 0$. Unfortunately, we cannot guarantee that $H \neq 0$ even for $h \neq 0$, so we have to be a bit more careful with the term $\dfrac{f(g(a)+H) - f(g(a))}{H}$. The case when $H = 0$, that is $g(a) = g(a+h)$, has to be treated separately.

To do so, we introduce a new function f^*, defined by

$$f^*(z) := \begin{cases} \dfrac{f(z) - f(y)}{z - y} & \text{if} \quad z \neq y \\ f'(y) & \text{if} \quad z = y, \end{cases}$$

with $y = g(a)$ as before.

Note that from the definition of $f^*(z)$ (and using (11.2)) we have that

$$\lim_{z \to y} \frac{f(z) - f(y)}{z - y} = f'(y).$$

This means that f^* is continuous at y:

$$\lim_{z \to y} f^*(z) = f^*(y) = f'(y).$$

Furthermore,

$$f(z) - f(y) = f^*(z)(z - y).$$

Note that this works both when $z \neq y$ and when $z = y$: in the latter case, both sides of the equation are zero.

With this preparation, the proof of the chain rule is now easy. Bearing in mind that $y = g(a)$ and using $z = g(a+h)$ in $f^*(z)$ we find

$$\begin{aligned} (f \circ g)'(a) &= \lim_{h \to 0} \frac{f(g(a+h)) - f(g(a))}{h} \\ &= \lim_{h \to 0} \frac{f^*(g(a+h))(g(a+h) - g(a))}{h} \\ &= \lim_{h \to 0} f^*(g(a+h)) \lim_{h \to 0} \frac{g(a+h) - g(a)}{h} = f'(g(a))g'(a). \end{aligned}$$

$\square$

Using the $\dfrac{\mathrm{d}}{\mathrm{d}x}$ notation, the chain rule is sometimes written as

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}g} \cdot \frac{\mathrm{d}g}{\mathrm{d}x}. \tag{12.1}$$

This means that the derivative of $f \circ g$ with respect to its argument is the product of the derivatives of f and g with respect to their arguments. In particular, the three derivatives in (12.1) are to be interpreted as

$$\begin{aligned}
\frac{\mathrm{d}f}{\mathrm{d}x} &= \frac{\mathrm{d}f(g(x))}{\mathrm{d}x} = (f \circ g)'(x) \\
\frac{\mathrm{d}f}{\mathrm{d}g} &= \left.\frac{\mathrm{d}f(g)}{\mathrm{d}g}\right|_{g=g(x)} = f'(g(x)) \\
\frac{\mathrm{d}g}{\mathrm{d}x} &= \frac{\mathrm{d}g(x)}{\mathrm{d}x} = g'(x).
\end{aligned}$$

Note: although the chain rule, as written in (12.1), looks like an exercise in cancelling fractions, it isn't! (Recall Section 11.3.)

Example 12.5 Differentiate

$$\left(x + \frac{1}{x}\right)^{100}$$

with respect to x.

Put $g(x) = x + \dfrac{1}{x}$ and $f(g) = g^{100}$ so that

$$f(g(x)) = \left(x + \frac{1}{x}\right)^{100}.$$

Then $f'(g) = 100g^{99}$ and $g'(x) = 1 - \frac{1}{x^2}$ and so

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(x + \frac{1}{x}\right)^{100} = f'(g(x)) \cdot g'(x) = 100(g(x))^{99}\left(1 - \frac{1}{x^2}\right) = 100\left(x + \frac{1}{x}\right)^{99}\left(1 - \frac{1}{x^2}\right).$$

Example 12.6 Compute $\dfrac{\mathrm{d}r}{\mathrm{d}s}$ for $r(s) = \dfrac{1}{(s^2 + s + 1)^3}$.

Write $r(s) = w(v(s))$ with $v(s) = s^2 + s + 1$ and $w(v) = v^{-3}$. Then, from the chain rule,

$$\frac{\mathrm{d}r}{\mathrm{d}s} = \frac{\mathrm{d}w}{\mathrm{d}v}\frac{\mathrm{d}v}{\mathrm{d}s} = (-3v^{-4})(2s + 1) = -\frac{3(2s + 1)}{(s^2 + s + 1)^4}.$$

Note that the somewhat enigmatic $\dfrac{\mathrm{d}w}{\mathrm{d}v}$ here means $w'(v(s))$.

In practice, we can apply the chain rule without introducing names for auxiliary functions and variables. The rule is: "differentiate the outer function and multiply by the derivative of the inner function".

Example 12.7 Find $\dfrac{\mathrm{d}}{\mathrm{d}x}\dfrac{1}{(x^4 + 1)^2}$.

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{(x^4+1)^2} = -\frac{2}{(x^4+1)^3}\,4x^3 = -\frac{8x^3}{(x^4+1)^3}.$$

Example 12.8 Find $\dfrac{\mathrm{d}}{\mathrm{d}t}\left(t^2-\dfrac{1}{t^2}\right)^5$.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(t^2-\frac{1}{t^2}\right)^5 = 5\left(t^2-\frac{1}{t^2}\right)^4\left(2t+\frac{2}{t^3}\right).$$

The chain rule allows one to differentiate the composition of more than two functions in one fell swoop by applying it multiple times.

Example 12.9 Find the derivative of $f\circ g\circ h$.

$$(f\circ g\circ h)'(a) = (f\circ(g\circ h))'(a) = f'((g\circ h)(a))\cdot(g\circ h)'(a) = f'((g\circ h)(a))\cdot g'(h(a))\cdot h'(a).$$

While this might look complicated at first sight, with some practice it is easy to apply. The recipe is to start with the outermost function, say f (that is, the last one to be applied, which is the leftmost in the composition). At this stage we completely ignore its argument. We compute the derivative f' and use as its argument the old argument of f. We then multiply the resulting expression by the derivative of that argument, which might in turn have to be computed by application of the chain rule.

Example 12.10 Find the derivative of $l(x) = \left\{x+\sin^3(x^5)\right\}^2$.

$$l'(x) = 2\left\{x+\sin^3(x^5)\right\}\cdot[1+3\sin^2(x^5)\cdot\cos(x^5)\cdot 5x^4].$$

Example 12.11 Find the derivative of $k(x) = \sqrt{\sin(x^3)}$.

$$k'(x) = \frac{1}{2\sqrt{\sin(x^3)}}\cdot\cos(x^3)\cdot 3x^2.$$

12.2 Derivatives of trigonometric functions

We have already established (in Example 11.6) that $\sin' = \cos$. This result is enough to find the derivatives of all other trigonometric functions. For cos and tan we find

$$\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) = \frac{\mathrm{d}}{\mathrm{d}x}\sin\left(\frac{\pi}{2} - x\right) = -\cos\left(\frac{\pi}{2} - x\right) = -\sin(x).$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan x = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\sin x}{\cos x} = \frac{\cos x\cos x - \sin x(-\sin x)}{(\cos x)^2} = \frac{1}{\cos^2 x} = \sec^2 x = 1 + \tan^2 x.$$

The following table lists the derivatives of the six basic trigonometric functions.

$f(x)$	$\sin x$	$\cos x$	$\tan x$	$\sec x$	$\csc x$	$\cot x$
$f'(x)$	$\cos x$	$-\sin x$	$\sec^2 x$	$\sec x\tan x$	$-\csc x\cot x$	$-\csc^2 x$

Example 12.12 Find $\dfrac{\mathrm{d}}{\mathrm{d}x}(\tan(4x))$.

$$\frac{\mathrm{d}}{\mathrm{d}x}(\tan(4x)) = \sec^2(4x)\frac{\mathrm{d}}{\mathrm{d}x}(4x) = 4\sec^2(4x).$$

Example 12.13 Find $\dfrac{\mathrm{d}}{\mathrm{d}x}(\cos^2 x + \sin^2 x)$.

$$\frac{\mathrm{d}}{\mathrm{d}x}(\cos^2 x + \sin^2 x) = 2\cos x(-\sin x) + 2\sin x\cos x = 0.$$

Example 12.14 Find $\dfrac{\mathrm{d}}{\mathrm{d}x}(\sqrt{1 + \sec x})$.

$$\frac{\mathrm{d}}{\mathrm{d}x}(\sqrt{1 + \sec x}) = \frac{1}{2\sqrt{1 + \sec x}}\frac{\mathrm{d}}{\mathrm{d}x}\sec x = \frac{\sec x\tan x}{2\sqrt{1 + \sec x}}.$$

Example 12.15 Find $\dfrac{\mathrm{d}}{\mathrm{d}t}(t^3\cot t)$.

$$\frac{\mathrm{d}}{\mathrm{d}t}(t^3\cot t) = 3t^2\cot t + t^3\{-\csc^2 t\} = 3t^2\cot t - t^3\csc^2 t.$$

Example 12.16 Find $\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{1-\tan\theta}{1+\tan\theta}\right)$.

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{1-\tan\theta}{1+\tan\theta}\right) = \frac{-\sec^2\theta(1+\tan\theta)-\sec^2\theta(1-\tan\theta)}{(1+\tan\theta)^2} = -\frac{2\sec^2\theta}{(1+\tan\theta)^2}.$$

12.3 Higher derivatives

Given a function f, its derivative f' is another function, which may again be differentiable. We denote the derivative of f' by f'', also written as $f^{(2)}$, and we write $f''(x) = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}$. Similarly, we denote the derivative of f'' by f''', also written as $f^{(3)}$, and we write $f'''(x) = \frac{\mathrm{d}^3 f}{\mathrm{d}x^3}$.

f' is called the **first derivative** of f, f'' the **second derivative**, etc. If $f(x)$ is differentiated n times in succession we obtain its n**-th derivative**, $f^{(n)}(x) = \frac{\mathrm{d}^n f}{\mathrm{d}x^n}$.

For example, if $f(x) = x^7$ then $f'(x) = 7x^6$; then $f''(x) = 7\cdot(6x^5) = 42x^5$; then $f^{(3)}(x) = 42\cdot(5x^4) = 210x^4$; and so on...

Example 12.17 Find $\tan''(x)$.

$$\frac{\mathrm{d}}{\mathrm{d}x}(\tan x) = \sec^2 x = 1+\tan^2 x.$$

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}(\tan(x)) = \frac{\mathrm{d}}{\mathrm{d}x}(1+\tan^2 x) = 2\tan x\sec^2 x = 2\tan x(1+\tan^2 x).$$

Example 12.18 Given $f(x) = \sin(x^2)$, evaluate $f''(x)$.

Differentiating once using the Chain Rule, we have $f'(x) = 2x\cos(x^2)$, so using the Product Rule we find $f''(x) = 2\cos(x^2) + 2x\cdot 2x\cdot(-\sin(x^2)) = 2\cos(x^2) - 4x^2\sin(x^2)$.

Example 12.19 Show that if $u(t) = A\sin(\alpha t) + B\cos(\alpha t)$ then $u'' + \alpha^2 u = 0$.

$$u'(t) = \alpha A\cos(\alpha t) - \alpha B\sin(\alpha t) \implies u''(t) = -\alpha^2 A\sin(\alpha t) - \alpha^2 B\cos(\alpha t) = -\alpha^2 u.$$

Example 12.20 Prove by induction that

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^n) = n!\,.$$

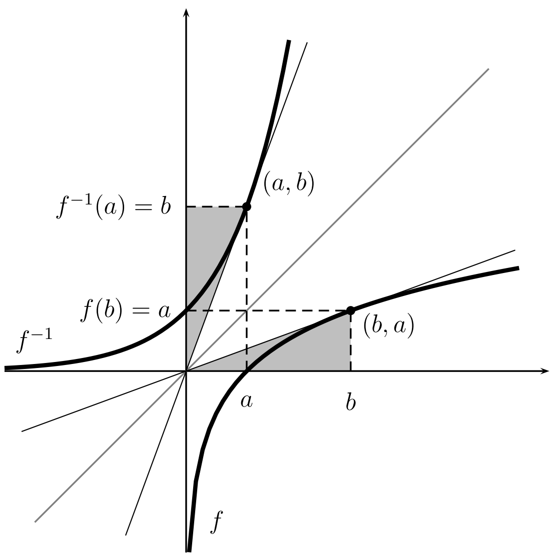
This is left to you!

12.4 Derivatives of inverse functions

Recall from Section 4.7 that when a function f is bijective then we can define a corresponding inverse function f^{-1}. Note that we cannot necessarily write this inverse function explicitly in terms of elementary functions: for example, we cannot write arcsin in this way. However, we can still obtain a useful result for the derivative of an inverse function.

Theorem 12.8 *Consider a function f on an interval I on which it is differentiable and invertible. Denote its inverse on I by f^{-1}. For all $b \in I$ where $f'(b) \neq 0$, the inverse function f^{-1} is then differentiable at $a = f(b)$ and*

$$(f^{-1})'(a) = \frac{1}{f'(b)} = \frac{1}{f'(f^{-1}(a))}. \tag{12.2}$$



Proof: The proof that the inverse function is indeed differentiable as long as $f'(b) \neq 0$ is a bit technical and we omit it here.[3] If we assume that f^{-1} can be differentiated, the formula (12.2) can be found easily. We start by observing that

$$f(f^{-1}(x)) = x$$

for all $x \in f(I)$. Computing the derivative with respect to x of both sides of this equation yields

$$f'(f^{-1}(x)) \cdot (f^{-1})'(x) = 1.$$

Setting $x = a$ gives

$$f'(f^{-1}(a)) \cdot (f^{-1})'(a) = 1,$$

and as $f'(f^{-1}(a)) = f'(b) \neq 0$ we can divide by it and hence obtain (12.2). $\square$

Example 12.21 Compute the derivative of $g(x) = \sqrt{x}$.

We have $g(x) = f^{-1}(x)$ where $f(x) = x^2$ for $x \geq 0$. If $a = f(b) = b^2$ then $b = g(a) = \sqrt{a}$. From (12.2) we find

$$g'(a) = \frac{1}{f'(b)} = \frac{1}{2b} = \frac{1}{2\sqrt{a}}$$

and so

$$\frac{\mathrm{d}}{\mathrm{d}x}\sqrt{x} = \frac{1}{2\sqrt{x}}.$$

12.5 Powers with rational exponents

From Example 12.21 we know that $\frac{\mathrm{d}}{\mathrm{d}x}x^{1/2} = \frac{1}{2}x^{-1/2}$. This is again formally the same as in Theorem 12.4, which so far we have proved only for integer exponents. We now show that the formula given in Theorem 12.4 is valid also for all rational exponents.

Theorem 12.9 *If $f(x) = x^n$ for $n \in \mathbb{Q}$ then*

$$f'(a) = na^{n-1},$$

for all a such that $f(a)$ and $f'(a)$ exist.

[3]If you're interested, you can find it in the books by Spivak (chapter 12) and Stewart (chapter 6).

Proof: We begin by looking at functions of the form $\sqrt[n]{x} = x^{\frac{1}{n}}$ for $n \in \mathbb{N}$.

For $n \in \mathbb{N}$ let

$$f_n(x) = x^n \begin{cases} \text{for all } x, & \text{when } n \text{ is odd,} \\ \text{for } x \geq 0, & \text{when } n \text{ is even.} \end{cases}$$

Then f_n is differentiable and invertible on its domain and

$$f_n^{-1}(x) = \sqrt[n]{x} = x^{1/n}.$$

Using Theorem 12.8 we find

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x} x^{\frac{1}{n}} &= \frac{1}{f_n'(f_n^{-1}(x))} = \frac{1}{n(f_n^{-1}(x))^{n-1}} = \frac{1}{n(x^{\frac{1}{n}})^{n-1}} \\ &= \frac{1}{nx^{1-\frac{1}{n}}} = \frac{1}{n} x^{\frac{1}{n}-1}. \end{aligned}$$

What we have established so far is that if $f(x) = x^a$, where a is an integer or one over a positive integer, then $f'(x) = ax^{a-1}$. We now extend this formula to all rational exponents, that is, exponents of the form $a = p/q$ with $p \in \mathbb{Z}$ and $q \in \mathbb{N}$.

$$\frac{\mathrm{d}}{\mathrm{d}x} x^{\frac{p}{q}} = \frac{\mathrm{d}}{\mathrm{d}x} (x^{\frac{1}{q}})^p = p(x^{\frac{1}{q}})^{p-1} \frac{1}{q} x^{\frac{1}{q}-1} = \frac{p}{q} x^{(\frac{p}{q}-\frac{1}{q}+\frac{1}{q}-1)} = \frac{p}{q} x^{\frac{p}{q}-1}$$

This means that for $f(x) = x^a$ with $a \in \mathbb{Q}$ we have $f'(x) = ax^{a-1}$, as required. $\square$

We will have to wait until Chapter 15 before we can show that this formula also holds for all real exponents that are not rational numbers. So far we cannot even interpret expressions such as $x^{\sqrt{2}}$ or x^{π}, let alone differentiate them!

Example 12.22 Differentiate $h(y) = (2y^3 + 4y^{1/2} - 5)^{1/3}$.

$$h'(y) = \frac{1}{3}(2y^3 + 4y^{1/2} - 5)^{-2/3} \left(6y^2 + 2y^{-1/2}\right).$$

12.6 Inverse trigonometric functions

As we have seen earlier, derivatives of trigonometric functions are again trigonometric functions or combinations of trigonometric functions. Derivatives of inverse trigonometric functions — surprisingly — are more closely related to rational functions.

Theorem 12.10

1. *Let* $f(x) = \arcsin x$*. Then* $f'(x) = \dfrac{1}{\sqrt{1-x^2}}$.

2. *Let* $f(x) = \arccos x$*. Then* $f'(x) = -\dfrac{1}{\sqrt{1-x^2}}$.

3. *Let* $f(x) = \arctan x$*. Then* $f'(x) = \dfrac{1}{1+x^2}$.

Proof:

1. Take $f(x) = \sin x$ for $x \in [-\pi/2, \pi/2]$. Then $f^{-1}(x) = \arcsin x$. We find the derivative of arcsin by using (12.2).

$$\arcsin'(x) = \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-\sin^2(\arcsin x)}} = \frac{1}{\sqrt{1-x^2}}.$$

We have used the fact that $\sin^2 x + \cos^2 x = 1$ and that $\cos x$ is positive over the range of the arcsin function, $[-\pi/2, \pi/2]$, that is, $\cos x = +\sqrt{1-\sin^2 x}$.

2. The derivative of arccos is computed in exactly the same way. The following paragraph is a verbatim copy of the preceding one with sin and cos interchanged, $-$ signs added where necessary, and an adjustment for the domains.

Take $f(x) = \cos x$ for $x \in [0, \pi]$. Then $f^{-1}(x) = \arccos x$. We find the derivative of arccos by using (12.2).

$$\arccos'(x) = \frac{1}{\cos'(\arccos x)} = -\frac{1}{\sin(\arccos x)} = -\frac{1}{\sqrt{1-\cos^2(\arccos x)}} = -\frac{1}{\sqrt{1-x^2}}.$$

We have used the fact that $\sin^2 x + \cos^2 x = 1$ and that $\sin x$ is positive over the range of the arccos function, $[0, \pi]$, that is, $\sin x = +\sqrt{1-\cos^2 x}$.

3. As a last example, we consider $f(x) = \tan x$, $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ with $f^{-1}(x) = \arctan x$. Using $f'(x) = \sec^2 x = 1 + \tan^2 x$ we find

$$\arctan'(x) = \frac{1}{(\tan)'(\arctan x)} = \frac{1}{1+\tan^2(\arctan x)} = \frac{1}{1+x^2}.$$

$\square$

Similar results can be derived for the inverse functions of cot, sec, and csc. However, the most important derivatives are those for arcsin and arctan. You should memorise these because of their applications in integration (see Chapter 16 and MM102).

Example 12.23 Find $\dfrac{\mathrm{d}}{\mathrm{d}x}\arcsin 3x$.

$$\frac{\mathrm{d}}{\mathrm{d}x}\arcsin 3x = \frac{1}{\sqrt{1-(3x)^2}}\frac{\mathrm{d}}{\mathrm{d}x}(3x) = \frac{3}{\sqrt{1-9x^2}}.$$

Example 12.24 Find $\dfrac{\mathrm{d}}{\mathrm{d}x}\arctan\dfrac{x}{7}$.

$$\frac{\mathrm{d}}{\mathrm{d}x}\arctan\frac{x}{7} = \frac{1}{1+(x/7)^2}\frac{1}{7} = \frac{7}{7^2[1+(x/7)^2]} = \frac{7}{49+x^2}.$$

Example 12.25 Find $\dfrac{\mathrm{d}}{\mathrm{d}x}(\arcsin x + \arccos x)$.

$$\frac{\mathrm{d}}{\mathrm{d}x}(\arcsin x + \arccos x) = \frac{1}{\sqrt{1-x^2}} - \frac{1}{\sqrt{1-x^2}} = 0.$$

(Note that from basic trig, if $\arcsin x = \theta$ then $\arccos x = \frac{\pi}{2} - \theta$, so the function we're differentiating here is a constant in disguise.)

Example 12.26 Find $\dfrac{\mathrm{d}}{\mathrm{d}t}\arctan\sqrt{t}$.

$$\frac{\mathrm{d}}{\mathrm{d}t}\arctan\sqrt{t} = \frac{1}{1+(\sqrt{t})^2}\frac{\mathrm{d}}{\mathrm{d}t}(\sqrt{t}) = \frac{1}{1+t}\frac{1}{2\sqrt{t}} = \frac{1}{2\sqrt{t}(1+t)}.$$

12.7 Summary

Given two differentiable functions f and g (g different from zero where necessary) and a constant $c \in \mathbb{R}$, we have the following basic rules for computing derivatives.

$$\begin{aligned}
(c\cdot f)' &= c\cdot f' \\
(f+g)' &= f'+g' \\
(f\cdot g)' &= f'\cdot g + f\cdot g' \\
\left(\frac{f}{g}\right)' &= \frac{f'\cdot g - f\cdot g'}{g^2} \\
(f\circ g)' &= (f'\circ g)\cdot g'
\end{aligned}$$

We also already know the derivatives of a few elementary functions. The most important results so far are given in the following table.

$f(x)$	$f'(x)$
x^a	$a \cdot x^{a-1},\ a \in \mathbb{Q}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\arcsin x$	$\dfrac{1}{\sqrt{1-x^2}}$
$\arctan x$	$\dfrac{1}{1+x^2}$