2. Derivatives and rates of change

Derivatives

Definition 1.1 (The derivative of f at a). *The derivative of a function f at a number a, denoted by $f'(a)$, is defined by*

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

if this limit exists and is finite. If $f'(a)$ exists, then we say that f is differentiable at a.

Here are some important points to note when you are answering a question like this.

- Always start by writing out the definition, e.g.

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \qquad \text{or} \qquad f'(3) = \lim_{h \to 0} \frac{f(3+h) - f(3)}{h}.$$

- Be careful when writing out $f(a+h)$. Take the definition of $f(x)$ and put $(a+h)$ in place of every x. Include the brackets! You will avoid mistakes that way.
- Make sure that you include the $\lim_{h \to 0}$ in every step, until you reach a point where you can actually compute the limit. (In the examples above, we had $\lim_{h \to 0}$ on every line until the very last one.)
- If the question asks you to work out $f'(2)$, then do that! Don't work out $f'(a)$ for a general a first. (There's probably a reason why the question is written that way. We will see examples where in some special values the calculation of the derivative is different than for other values).

Observe that $f'(a)$ is the gradient of the tangent line to $y = f(x)$ at $(a, f(a))$. Observe also that if we regard $f(t)$ as a position, then $f'(a)$ is the instantaneous velocity at time a.

2-2 The derivative as a function

Definition 1.2 (The derivative). *Let f be a function. The derivative of f, denoted f', is the function defined by*

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

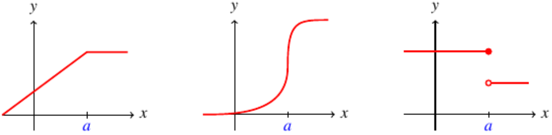
The domain of f' is

$$\mathrm{dom}(f') = \left\{ x \;\middle|\; \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \text{ exists and is finite} \right\}.$$

Theorem 1.3. *If f is differentiable at a, i.e. if $f'(a)$ exists, then f is continuous at a. Phrased differently, if f is not continuous at a then f cannot be differentiable at a.*

Three ways a function can fail to be differentiable

Here are three graphs depicting how a function can fail to be differentiable at a number a.



(a) corner (b) "infinite slope" (c) discontinuity

Definition 1.4 (Leibniz notation). *If we use the traditional notation* $y = f(x)$ *to indicate that the variable* y *depends on the variable* x *by means of the function* f, *then there are many different ways to denote the derivative, as follows.*

$$f'(x) = y' = \frac{dy}{dx} = \frac{df}{dx} = \frac{d}{dx} f(x)$$

The notation involving $\frac{d}{dx}$ *is called Leibniz notation. We will switch between notations frequently.*

Definition 1.5 (Higher derivatives). *Let* f *be a function. Its derivative* f' *is another function. That means that we can differentiate* f' *to produce another function,* $(f')'$, *which is called the second derivative and denoted* f''.
Differentiating once more gives $(f'')'$, *which is denoted* f''' *and called the third derivative. Repeating the process, we can define the* n*th derivative of* f, *which is denoted by* $f^{(n)}$.
In Leibniz notation, if $y = f(x)$, *then we would write*

$$f^{(n)}(x) = y^{(n)} = \frac{d^n y}{dx^n} = \frac{d^n f}{dx^n} = \frac{d^n}{dx^n} f(x).$$

Derivative of a constant function. Let c be any constant and let f be the function defined by $f(x) = c$. Then $f'(x) = 0$. Or in other words,

$$\frac{d}{dx} c = 0.$$

Derivative of a power function. If n is a real number, then

$$\frac{d}{dx} x^n = nx^{n-1}.$$

Next, we have some rules which tell us how to find the derivatives of new functions from the derivatives of old functions. We will write the proof to some of these rules and give examples.

Derivative of constant multiples. If c is a constant and f is differentiable, then so is cf, and

$$\frac{d}{dx}(cf(x)) = c\frac{d}{dx}(f(x)).$$

The sum rule.

If f and g are both differentiable, then

$$\frac{d}{dx}[f(x)+g(x)] = \frac{d}{dx}f(x) + \frac{d}{dx}g(x).$$

The difference rule.

If f and g are both differentiable, then

$$\frac{d}{dx}[f(x) - g(x)] = \frac{d}{dx}f(x) - \frac{d}{dx}g(x).$$

The sum rule, difference rule, and constant multiple rule can be combined to show that, if f and g are differentiable and a, b are constants, then

$$\frac{d}{dx}[af(x) + bg(x)] = a\frac{d}{dx}f(x) + b\frac{d}{dx}g(x).$$

And indeed, this works when we add together scalar multiples of any number of functions. So:
If $f, g, \ldots, h$ are differentiable and $a, b, \ldots c$ are constants, then

$$\frac{d}{dx}[af(x) + bg(x) + \cdots + ch(x)] = a\frac{d}{dx}f(x) + b\frac{d}{dx}g(x) + \cdots + c\frac{d}{dx}h(x).$$

We can use this together with the power law to differentiate any polynomial.

The product rule.

If f and g are both differentiable, then

$$\frac{d}{dx}[f(x)g(x)] = f(x)\frac{d}{dx}[g(x)] + \frac{d}{dx}[f(x)]g(x).$$

Or, in other notation,

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

The quotient rule.

If f and g are differentiable, then if $g(x) \neq 0$ the function $f(x)/g(x)$ is differentiable, and

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)\frac{d}{dx}[f(x)] - f(x)\frac{d}{dx}[g(x)]}{[g(x)]^2}.$$

Or, using briefer notation on the right hand side,

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

2-4 Derivatives of trigonometric functions

In this section we will study trigonometric functions and their derivatives.

Definition 1.6 (The trigonometric functions). *The trigonometric functions are as follows.*

$$\sin(\theta) \qquad \csc(\theta) = \frac{1}{\sin(\theta)}$$

$$\cos(\theta) \qquad \sec(\theta) = \frac{1}{\cos(\theta)}$$

$$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)} \qquad \cot(\theta) = \frac{\cos(\theta)}{\sin(\theta)}$$

Here, $\sin(\theta)$ *and* $\cos(\theta)$ *are defined as follows. Take a line segment of length* 1*, based at the origin, and making an anticlockwise angle of* θ *with the positive* x*-axis. Then* $\cos(\theta)$ *is defined to be the* x*-coordinate of the end of the line segment, and* $\sin(\theta)$ *is defined to be the* y*-coordinate of the end of the line segment:*

Two special limits.

$$\lim_{\theta \to 0} \frac{\sin(\theta)}{\theta} = 1 \qquad \lim_{\theta \to 0} \frac{\cos(\theta) - 1}{\theta} = 0$$

Now that we've computed our special limit, we are in a position to work out the following.

Derivatives of $\sin$ and $\cos$.

$$\frac{d}{d\theta}\sin(\theta) = \cos(\theta)$$

$$\frac{d}{d\theta}\cos(\theta) = -\sin(\theta)$$

Further derivatives of trigonometric functions

$$\frac{d}{d\theta}\csc(\theta) = -\csc(\theta)\cot(\theta)$$

$$\frac{d}{d\theta}\sec(\theta) = \sec(\theta)\tan(\theta)$$

$$\frac{d}{d\theta}\tan(\theta) = \sec^2(\theta)$$

$$\frac{d}{d\theta}\cot(\theta) = -\csc^2(\theta)$$

We would like to find the derivative of the composition of two functions $f \circ g$. We will prove the following:

The chain rule. If g is differentiable at x and f is differentiable at $g(x)$, then $f \circ g$ is differentiable at x and

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

2-6 Implicit differentiation

Theorem 1.7. *Let $F(x, y)$ be a nice function of two variables (we will not say precisely what "nice" means here. All the functions which we will consider here will be nice). Assume that $F(a, b) = 0$ and that $F_y(a, b) \neq 0$. Then there is a continuous differentiable function $f(x)$ such that:*

1. $f(a) = b$.
2. *For x close enough to a and y close enough to b it holds that $F(x, f(x)) = 0$ and moreover $F(x, y) = 0$ exactly when $f(x) = y$.*
3. *The derivative of f is given by $f'(x) = \frac{-F_x(x,y)}{F_y(x,y)}$*

Remark 1.8.

- *This theorem sounds quite complicated at first, but using it is easier. It means that if we look close enough to the point (a, b), the collection of points (x, y) which satisfy the equation $F(x, y) = 0$ look like the graph of a function. Usually we will not be able to write a precise formula for y as a function of x, but that's completely fine.*
- *To find the derivative, we derive $F(x, y)$ with respect to the variable x, where we think of y as a function of x. This will be the same as calculating the quotient $\frac{-F_x(x,y)}{F_y(x,y)}$.*

Definition 1.9 (Absolute and local maxima and minima). *Let f be a function with domain D and let $c \in D$. Then $f(c)$ is:*

- *The absolute maximum of f if $f(c) \geqslant f(x)$ for all $x \in D$.*
- *The absolute minimum of f if $f(c) \leqslant f(x)$ for all $x \in D$.*
- *A local maximum of f if $f(c) \geqslant f(x)$ for all x near c.*
- *A local minimum of f if $f(c) \leqslant f(x)$ for all x near c.*

We say that f has an absolute maximum at c, or attained at c, and so on. As usual, 'near' means 'in some open interval containing c'. (This means in particular that we do not allow local maxima or minima at the endpoints of a closed interval $[a, b]$.)

The following theorem guarantees that there are absolute maximum and minimum values under certain conditions.

Theorem 1.10 (The extreme value theorem). *If f is continuous on a closed interval $[a, b]$, then f attains an absolute maximum $f(c)$ and an absolute minimum $f(d)$ for some $c, d \in [a, b]$.*

Theorem 1.11. *(Fermat) If f has a local minimum or maximum at c, then c is a critical number of f.*

Definition 1.12 (Critical number). *Here a critical number of a function f is a number c in the domain of f such that $f'(c) = 0$ or $f'(c)$ is not defined.*

The closed interval method.

Let f be a continuous function defined on a closed interval $[a, b]$. To find the absolute maximum and minimum values of f on $[a, b]$, we follow these steps.

1. Find the critical numbers c of f in (a, b), and for each one compute $f(c)$.
2. Compute $f(a)$ and $f(b)$.
3. The largest of the numbers from steps 1 and 2 is the absolute maximum, and the smallest of the numbers from steps 1 and 2 is the absolute minimum.

Theorem 1.13 (The Mean Value Theorem). *Let f be a function satisfying the following hypotheses:*

- f *is continuous on the closed interval* $[a, b]$
- f *is differentiable on the open interval* (a, b).

Then there is a number c *in* (a, b) *such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

To prove this theorem, we will begin with the case where $f(a) = f(b)$. This is called Rolle's Theorem:

Theorem 1.14 (Rolle's Theorem). *Let f be a function that satisfies the following three hypotheses:*

- *f is continuous on the closed interval $[a, b]$*
- *f is differentiable on the open interval (a, b).*
- *$f(a) = f(b)$.*

Then there is a number c in (a, b) such that $f'(c) = 0$.

The mean value theorem has the following corollary:

Theorem 1.15. *Assume that* $f'(x) = 0$ *for all* $x \in (a, b)$. *Then* f *is constant on* (a, b).

Proof.

Let c, d be two points in (a, b). We want to show that $f(c) = f(d)$. By the mean value theorem, there is a point e in (c, d) such that $f'(e) = \frac{f(d)-f(c)}{d-c}$. But this means that $f(d) - f(c) = 0$. □

Corollary 1.16. *Assume that* $f'(x) = g'(x)$ *for all* x *in an interval* (a, b)*. Then* $f - g$ *is constant on* (a, b)*.*

Proof.

Consider the function $F(x) = f(x) - g(x)$ and use the above theorem. Fill in the details! □

In this section we will see several more specific ways in which the derivative of a function affects its graph.

The increasing / decreasing test.

- If $f'(x) > 0$ for all x in an open interval I, then f is increasing on I.
- If $f'(x) < 0$ for all x in an open interval I, then f is decreasing on I.

Now we will see several tests designed to tell whether a critical point is a local maximum or local minimum or neither.

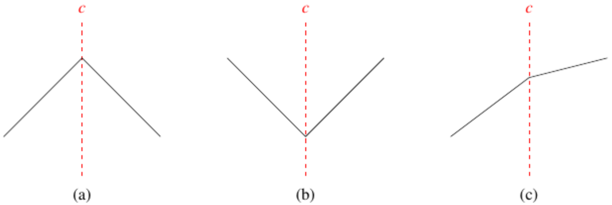
The first derivative test. Suppose that c is a critical number of a continuous function f.

(a) If f' changes from positive to negative at c, then f has a local maximum at c.

(b) If f' changes from negative to positive at c, then f has a local minimum at c.

(c) If f' does not change sign at c, then f has neither a local maximum nor a local minimum at c.

The kind of behaviour we are discussing can be depicted as follows.



The rule applies when $f'(x)$ exists for all x close to c but not necessarily equal to c, in particular can apply even when $f'(c)$ does not exist.

The second derivative test. Suppose that c is a critical number of a continuous function f. Suppose that f'' is defined and is continuous near c. Then:

(i) If $f'(c) = 0$ and $f''(c) > 0$, then f has a local minimum at c.

(ii) If $f'(c) = 0$ and $f''(c) < 0$, then f has a local maximum at c.

The test gives no conclusion if $f''(c) = 0$.

Alternative definition of e: The number e is the limit

$$\lim_{x \to \infty} (1 + \frac{1}{x})^x.$$

Combining the computation of $f'(x)$ with the definition of e, we get the following.

The derivative of a^x.

$$\frac{d}{dx} a^x = a^x \ln(a).$$

The derivative of e^x.

$$\frac{d}{dx} e^x = e^x.$$

And combining this with the chain rule gives us the following.

The derivative of $e^{f(x)}$.

$$\frac{d}{dx} e^{f(x)} = e^{f(x)} \cdot f'(x).$$

Definition 1.17 (The natural logarithm). *The function* f *defined by* $f(x) = e^x$ *is increasing, and so is one-to-one. Its domain is* $\mathbb{R} = (-\infty, \infty)$*, and its range is* $(0, \infty)$*.*
The natural logarithm, denoted $\ln$*, is the inverse function (we will say what we mean by inverse function in the next section). Its domain is* $(0, \infty)$*, its range is* $(-\infty, \infty)$*, and it is characterised by the fact that*

$$\ln(y) = x \iff y = e^x.$$

Example 1.18.

- $e^0 = 1$*, and so* $\ln(1) = 0$*.*
- $e^1 = e$*, and so* $\ln(e) = 1$*.*

The following properties of $\ln$ are all consequences of the definition of $\ln$ together with properties of the exponential function.

Properties of $\ln$.

$$\begin{aligned}
\ln(ab) &= \ln(a) + \ln(b) \\
\ln(a/b) &= \ln(a) - \ln(b) \\
\ln(a^r) &= r\ln(a) \\
e^{\ln(x)} &= x \\
\ln(e^x) &= x
\end{aligned}$$

2-12 Derivatives of logarithmic functions

The derivative of $\log_a(x)$.

$$\frac{d}{dx}\log_a(x) = \frac{1}{x\ln(a)}.$$

The derivative of $\ln$.

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

The derivative of $\ln(g(x))$.

$$\frac{d}{dx}\left[\ln g(x)\right] = \frac{g'(x)}{g(x)}.$$

2-13 Inverse functions

Definition 1.19 (One-to-one functions). *A function f is called a one-to-one function if it never takes the same value twice. In other words, f is one-to-one if*

$$f(x_1) \neq f(x_2) \text{ whenever } x_1 \neq x_2.$$

In terms of graphs, a function is one-to-one if and only if there is no horizontal line that intersects its graph more than once.

Definition 1.20 (Inverse functions). *Let f be a one-to-one function with domain A and range B. (Recall that the range of f is the set of all numbers of the form $f(x)$ for some $x \in A$.) Then the inverse function, denoted f^{-1}, has domain B and range A and is defined by the rule*

$$f^{-1}(y) = x \iff y = f(x).$$

Here are three important properties of the inverse function. The first is in fact the definition again.

$$f^{-1}(y) = x \iff y = f(x)$$

The next ones are called the *cancellation equations*.

$$f^{-1}(f(x)) = x \qquad \text{for every } x \in A.$$

$$f(f^{-1}(x)) = x \qquad \text{for every } x \in B.$$

How to find the inverse. Let f be a one-to-one function. We find the inverse of f as follows.

1. Write $y = f(x)$.
2. Solve the equation to give x in terms of y, i.e. $x = \cdots$, where the right hand side involves y but not x.
3. Swap x and y, so that you now have an equation of the form $y = \cdots$.
4. The resulting equation is $y = f^{-1}(x)$.

Derivatives of inverse functions. If f is a one-to-one differentiable function and $f'(f^{-1}(a)) \neq 0$, then f^{-1} is differentiable at a and

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