Advanced Mathematics I-1

University of Aberdeen

2022-2023

Contents

Chapter 1

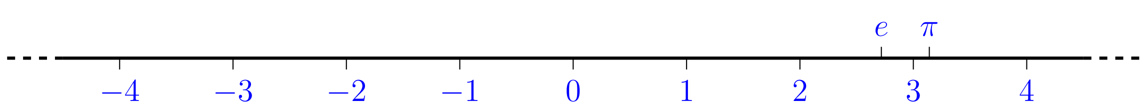
Functions

1.1 Sets

Definition 1.1 (Sets and elements). A *set* is a collection of objects (maybe numbers, maybe kittens, maybe something else). The objects in the set are called the *elements* of the set. If S is a set, and a is an element of S, then we write $a \in S$. If a is not an element of S, we write $a \notin S$.

Example 1.2.

- $\mathbb{N}$ denotes the set of *natural numbers*, i.e. whole positive numbers.
- $\mathbb{Z}$ denotes the set of *integers*, i.e. whole numbers, both positive and negative.
- $\mathbb{Q}$ denotes the set of *rational numbers*. These are numbers that we can write as the quotient of two integers, i.e. $2 = \frac{2}{1}, 1.5 = \frac{3}{2}, -2\frac{1}{3} = \frac{-7}{3}$.
- $\mathbb{R}$ denotes the set of *real numbers*, i.e. all the numbers on the number line.



- $\emptyset$ is the empty set, which is the set with no elements at all.

Definition 1.3 (Describing sets). Some sets can be described by listing their elements between *braces*, i.e. the symbols { and }. For example:

- $\{0\}$ is a set with one element, 0.
- $\{-1, 2, 5\}$ is a set with three elements.
- $\{\}$ is a set with no elements, or in other words it is the empty set $\emptyset$.
- $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$ is another way of writing $\mathbb{Z}$.

Sets can also be described using *set-builder notation*, for example

$$A = \{x \mid x \text{ is an even integer}\} = \{\ldots, -4, -2, 0, 2, 4, 6, \ldots\}.$$

Definition 1.4 (Intersections and unions). If S and T are sets, then their *intersection*, denoted $S \cap T$, is the set of all elements of S that are also elements of T. So in set-builder notation

$$S \cap T = \{x \mid x \in S \text{ and } x \in T\}.$$

Their *union*, denoted $S \cup T$, is the set consisting of all elements of S and all elements of T. So in set-builder notation:

$$S \cup T = \{x \mid x \in S \text{ or } x \in T \text{ or both}\}.$$

Their *set difference*, denoted $S \backslash T$ is the set containing all the elements of S which are not elements of T:

$$S \backslash T = \{x | x \in S \text{ and } x \notin T\}.$$

Example 1.5.

$$\begin{aligned}
\{-2, -1, 0\} \cup \{0, 1\frac{1}{2}, \pi\} &= \{-2, -1, 0, 1\frac{1}{2}, \pi\} \\
\{-2, -1, 0\} \cap \{0, 1\frac{1}{2}, \pi\} &= \{0\} \\
\{-2, -1, 0\} \cap \{1, 2\} &= \emptyset \\
\{-2, -1, 0\} \cup \{1, 2\} &= \{-2, -1, 0, 1, 2\} \\
\{-2, -1, 0\} \backslash \{1, 2\} &= \{-2, -1, 0\} \\
\{-2, -1, 0\} \backslash \{0, -1, \frac{1}{2}, \pi\} &= \{-2\}
\end{aligned}$$

Warning. The elements of a set are not ordered, so that for example $\{1, 2\} = \{2, 1\}$. And we can list the same element twice without changing the set, so that for example $\{1, 2, 2, 3, 3, 3, 3, 3\} = \{1, 2, 3\}$.

Definition 1.6 (Intervals). Let $a < b$ be real numbers.

- The *open interval* (a, b) denotes the set of all numbers between a and b, *not* including a or b. So in set-builder notation:

$$(a, b) = \{x \mid a < x < b\}$$

- The *closed interval* $[a, b]$ denotes the set of all numbers between a and b, *including* a and b themselves. So in set-builder notation:

$$[a, b] = \{x \mid a \leqslant x \leqslant b\}$$

- There are also two kinds of *half-open interval* defined as follows:

$$[a, b) = \{x \mid a \leqslant x < b\}$$
$$(a, b] = \{x \mid a < x \leqslant b\}$$

Question. Describe the following sets as a single interval.

1. $(1, 3) \cup (2, 4]$
2. $(\pi, 10) \cup [10, 11]$
3. $(0, 1) \cup (1, 2)$

Definition 1.7 (Unbounded intervals). There are also *unbounded intervals* defined as follows.

$$\begin{aligned} (a, \infty) &= \{x \mid a < x\} \\ (-\infty, b) &= \{x \mid x < b\} \\ [a, \infty) &= \{x \mid a \leqslant x\} \\ (-\infty, b] &= \{x \mid x \leqslant b\} \\ (-\infty, \infty) &= \mathbb{R} \end{aligned}$$

Warning: Note that the symbol ∞ only appears in the names on the left, but not as a number on the right. In fact, ∞ **is not a number**, just a symbol we use for abbreviation.

1.2 Functions

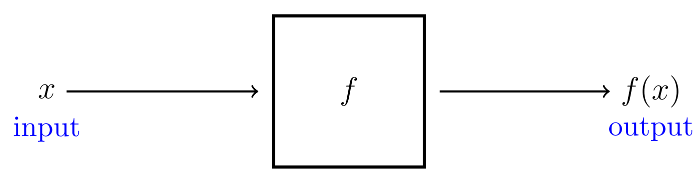
Next, we introduce functions. Functions arise when one quantity depends on another, e.g.:

- The area A of a circle depends on the radius r of the circle. For each r there is just one value of A.
- The population P of the earth depends on the time t. At each time t, the population P is a number that can in principle be determined.

Definition 1.8 (Function). A *function* f is a rule that assigns to each element x in a set D a unique element $f(x)$ in a set E. Usually, D and E will be sets of real numbers. The set D is called the *domain* of f, and the set E is called the *codomain* of f. We will sometimes write the domain of f as $\mathrm{dom}(f)$. The *range* of f is the set of all possible values of $f(x)$ as x goes through all elements of the domain (So the range is the 'smallest possible choice of E'). We write $ran(f)$ for the range of f. To put it in the set-builder notation:

$$ran(f) = \{y | y \in E \text{ and there exists an } x \in D \text{ such that } y = f(x)\}$$

We think of a function as a machine:

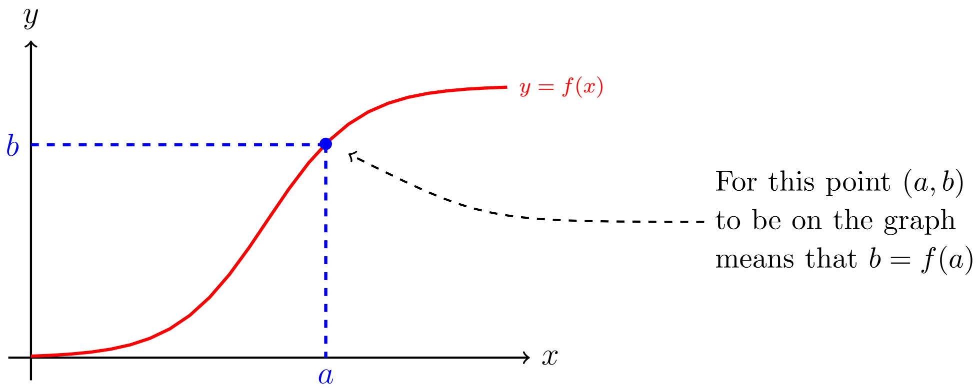


The same input x always produces the same output $f(x)$.

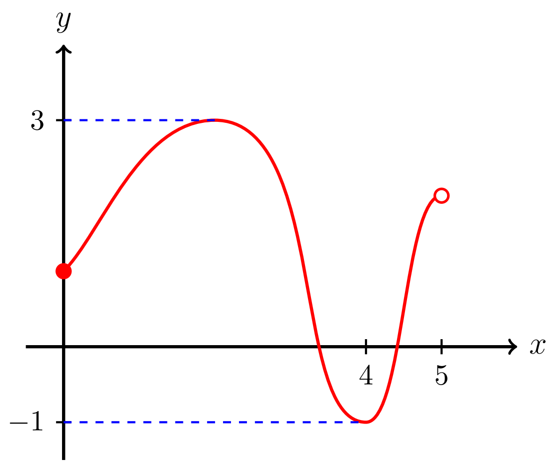
Definition 1.9 (Graphs). The *graph* of a function f is the set of all pairs (x, y) of real numbers such that x is in the domain of f and $y = f(x)$. So as a set the graph is

$$\{(x, f(x)) \mid x \text{ is in the domain of } f\}.$$

We usually plot the graph on the plane.



Question. Here is the graph of f.



In this image a solid dot means that the point is on the graph, but a hollow dot means that the point is not on the graph.

1. What is $f(4)$?
2. What is the domain of f?
3. What is the range of f?

Solution

1. -1

2. $[0, 5)$
3. $[-1, 3]$

Example 1.10. We usually define functions using formulas. For example, let f be the function with domain $D = [0, \infty)$, with $E = \mathbb{R}$, and defined by $f(x) = x^2$ for $x \in D$. The range of f is $[0, \infty)$ because $f(x) = x^2 \geqslant 0$ for all $x \in D$, and every element of $[0, \infty)$ has the form $f(x)$ for some $x \in D$.

Definition 1.11 (Domain convention). If a function is given by a formula, and the domain has not been specified, then our *domain convention* is that the domain of the function is the set of all numbers for which the formula makes sense and defines a real number. We will always use this domain convention unless otherwise specified.

Example 1.12. Let f and g be the functions defined by $f(x) = \sqrt{x}$ and $g(x) = \frac{1}{x}$. The formula $\sqrt{x}$ makes sense and defines a real number so long as $x \geqslant 0$. So the domain convention says that $\mathrm{dom}(f) = \{x \mid x \geqslant 0\} = [0, \infty)$. And the formula $\frac{1}{x}$ makes sense and defines a real number so long as $x \neq 0$. (We can *never* divide by 0!) So the convention says that $\mathrm{dom}(g) = \{x \mid x \neq 0\} = (-\infty, 0) \cup (0, \infty)$.

Question. Find the domain of the functions f and g defined as follows.

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

Write your answer as a union of intervals.

Solution

1. The formula $\sqrt{x + 2}$ produces a real number when $x + 2 \geqslant 0$, i.e. when $x \geqslant -2$. So the domain is $[-2, \infty)$.
2. The formula $\frac{1}{x^2 - x}$ makes sense so long as $x^2 - x \neq 0$. Now $x^2 - x = x(x - 1)$ is 0 exactly when $x = 0$ and $x = 1$. So the domain consists of all real numbers except 0 and 1. As a union of intervals, this is $(-\infty, 0) \cup (0, 1) \cup (1, \infty)$.

Definition 1.13 (Piecewise defined functions). A *piecewise defined function* is one that is defined in different ways on different parts of the domain. We do this by listing the different definitions and the parts of the domain on which they apply, all inside a big brace symbol. For example, we may define a function f with domain $\mathbb{R}$ as follows.

$$f(x) = \begin{cases} x & \text{if } x < 0 \\ \sin(x) & \text{if } x \geqslant 0 \end{cases}$$

So to work out $f(-1)$ we note that $-1 < 0$, so the first of the two formulas applies, and we get $f(-1) = -1$. And to work out $f(\pi/2)$ we note that $\pi/2 \geqslant 0$, so that the second of the two formulas applies, and we get $f(\pi/2) = \sin(\pi/2) = 1$.

- Note that the definition covers the whole of the domain $\mathbb{R}$, in the sense that every $x \in \mathbb{R}$ obeys either $x < 0$ or $x \geqslant 0$.
- Note that for each x, $f(x)$ is only defined once. In other words, the regions determined by $x < 0$ and $x \geqslant 0$ do not overlap.
- There can be two 'pieces' in the piecewise definition, as above, or there can be three or ten or ...
- We can sometimes change the defining regions without changing the function. For example, in this case substituting $x = 0$ into either $f(x) = x$ or $f(x) = \sin(x)$ gives $f(x) = 0$ in both cases. This means that we could have defined the two pieces on the regions $x \leqslant 0$ and $x > 0$ without changing f. (So there was a choice, and either one is fine!)

Example 1.14 (The absolute value function). The *absolute value function* sends a real number x to the distance from 0 to x. It is denoted by $|x|$, sometimes pronounced 'mod x' or 'modulus of x'. Distance is never negative, so

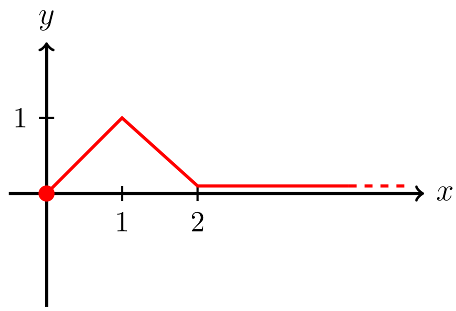
$$|\pi| = \pi, \qquad |-\pi| = \pi, \qquad |-10| = 10, \qquad |0| = 0.$$

In general, the formula is:

$$|x| = \begin{cases} x & \text{if } x \geqslant 0 \\ -x & \text{if } x < 0 \end{cases}$$

Let's check this. To work out $|-\pi|$, we observe that $x = -\pi$ satisfies $x < 0$ and so the second definition $|x| = -x$ applies, giving us $|-\pi| = -(-\pi) = \pi$.

Question. Find a formula for the function f.



Solution

$$f(x) = \begin{cases} x & \text{if } 0 \leqslant x < 1 \\ 2 - x & \text{if } 1 \leqslant x < 2 \\ 0 & \text{if } 2 \leqslant x \end{cases}$$

Definition 1.15 (Even and odd functions).

- If a function f satisfies the rule $f(-x) = f(x)$ for every x in its domain, then f is called an *even function*.
- If it satisfies $f(-x) = -f(x)$ for every x in its domain, then it is called an *odd function*.

Example 1.16. Determine whether the function is even or odd.

1. $f(x) = x^5 + x$
2. $g(x) = 1 - x^4$
3. $h(x) = 2x - x^2$

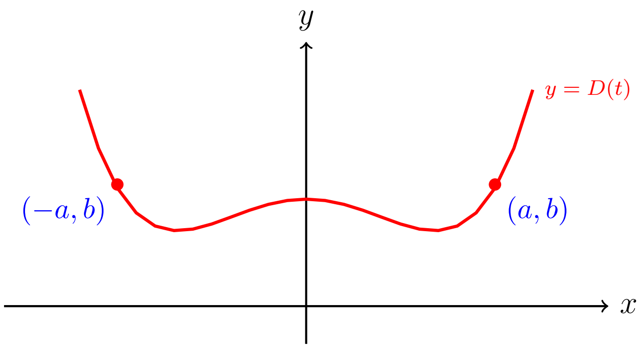
Solution. To find out whether a function f is even or odd, we start by working out $f(-x)$ and simplifying, and we then see whether it is equal to $f(x)$ or $-f(x)$.

1. $f(-x) = (-x)^5 + (-x) = -x^5 - x = -(x^5 + x) = -f(x)$ and so f is odd (and not even).
2. $g(-x) = 1 - (-x)^4 = 1 - x^4 = g(x)$ and so g is even (and not odd).
3. $h(-x) = 2(-x) - (-x)^2 = -2x - x^2$. This is not equal to $h(x)$ or to $-h(x)$, and so h is neither even nor odd.

Warning. A function does not have to be even or odd. In fact most functions are neither even nor odd. It is even possible for a function to be both even and odd.

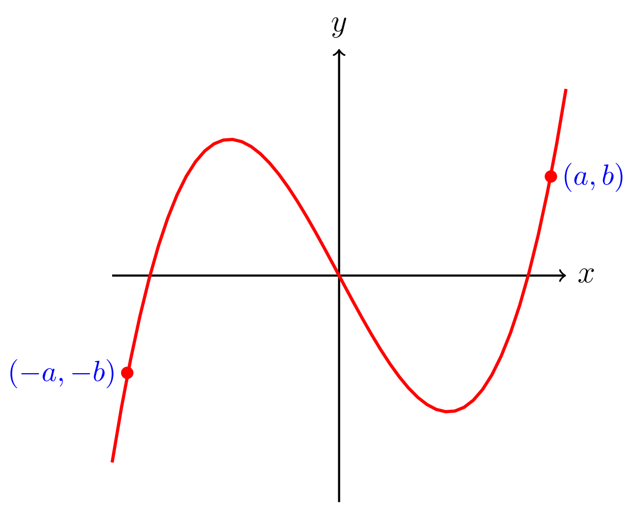
It can be helpful to understand odd and even functions in terms of their graphs.

- The graph of an even function is symmetric under reflection in the y-axis.



  Why is this? Our claim is that if (a, b) lies on the graph of f then so does $(-a, b)$. If (a, b) is on the graph then $b = f(a)$, but since f is even then $f(-a) = f(a)$, so that $b = f(-a)$, and consequently $(-a, b)$ also lies on the graph.

- The graph of an odd function is symmetric under $180°$ rotation through $(0, 0)$.



Why is this? Our claim is that if (a, b) lies on the graph of f then so does $(-a, -b)$. If (a, b) is on the graph then $b = f(a)$, but since f is odd we have $f(-a) = -f(a)$, so that $-b = -f(a) = f(-a)$, and consequently $(-a, -b)$ also lies on the graph.

Definition 1.17 (Increasing and decreasing functions). A function f is called *increasing* on an interval I if the following holds:

Whenever $x_1, x_2 \in I$ and $x_1 < x_2$, we have $f(x_1) < f(x_2)$.

And it is called *decreasing* if the following holds:

Whenever $x_1, x_2 \in I$ and $x_1 < x_2$, we have $f(x_1) > f(x_2)$.

Example 1.18.

- If g is the function defined by $g(x) = x^2$ then g is increasing on $(0, 1)$. That is because if $x_1, x_2 \in (0, 1)$ and $x_1 < x_2$, then $x_1^2 < x_2^2$, so that $g(x_1) < g(x_2)$.
- If h is the function defined by $h(x) = -x$, then h is decreasing on $(0, 1)$. That is because if $x_1, x_2 \in (0, 1)$ and $x_1 < x_2$, then $-x_1 > -x_2$, so that $h(x_1) > h(x_2)$.

In terms of graphs, a function is increasing on an interval I if the graph goes upwards as you move from left to right along I. Similarly, the function is decreasing on an interval I if the graph does downwards as you move from left to right along I.

1.3 New functions from old functions

Definition 1.19 (Sums, differences, products and quotients). Let f and g be functions. The *sum* and *difference* of f and g, denoted by $f + g$ and $f - g$, are the new functions defined by

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

and

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

The *product* and *quotient* of f and g, denoted by fg and f/g, are the new functions defined by

$$(fg)(x) = f(x)g(x)$$

and

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

Example 1.20. Let f and g be the functions defined by $f(x) = \sin(x)$ and $g(x) = 2^x$. Find formulas for $(f+g)(x)$, $(f-g)(x)$, $(fg)(x)$ and $(f/g)(x)$.

Solution.

- $(f+g)(x) = f(x) + g(x) = \sin(x) + 2^x$
- $(f-g)(x) = f(x) - g(x) = \sin(x) - 2^x$
- $(fg)(x) = f(x)g(x) = \sin(x) \cdot 2^x$
- $(f/g)(x) = f(x)/g(x) = \frac{\sin(x)}{2^x}$

Definition 1.21 (Domains of sums, differences, products and quotients). The domains of $f+g$, $f-g$, fg and f/g are defined, according to our domain convention, as the largest sets for which the given formulas make sense and produce a real number. Working this out gives us the following:

$$\begin{aligned} \mathrm{dom}(f+g) &= \mathrm{dom}(f) \cap \mathrm{dom}(g) \\ \mathrm{dom}(f-g) &= \mathrm{dom}(f) \cap \mathrm{dom}(g) \\ \mathrm{dom}(fg) &= \mathrm{dom}(f) \cap \mathrm{dom}(g) \\ \mathrm{dom}(f/g) &= \{x \in \mathrm{dom}(f) \cap \mathrm{dom}(g) \mid g(x) \neq 0\} \end{aligned}$$

For example, in the case of $f+g$, the formula $f(x)+g(x)$ makes sense whenever $f(x)$ and $g(x)$ are both defined, since we can always add any two real numbers to get another real number, and so $\mathrm{dom}(f+g) = \mathrm{dom}(f) \cap \mathrm{dom}(g)$. On the other hand, in the case of f/g, the formula $f(x)/g(x)$ makes sense whenever f and g are both defined *and* $g(x) \neq 0$, since we cannot divide by 0.

Example 1.22. Let f and g be the functions defined by $f(x) = 1/x$ and $g(x) = \sqrt{x} - 1$. What are the domains of the functions $f+g$, $f-g$, fg and f/g?

Solution. We do not need to work out the functions $f+g$, $f-g$, fg and f/g in order to answer this question! Instead, we will use the rules for their domains given in the last definition. In order to do this we first work out that $\mathrm{dom}(f) = \{x \mid x \neq 0\} = (-\infty, 0) \cup (0, \infty)$ and $\mathrm{dom}(g) = \{x \mid x \geqslant 0\} = [0, \infty)$. Thus we have:

- $\mathrm{dom}(f + g) = \mathrm{dom}(f) \cap \mathrm{dom}(g) = \{x \mid x \neq 0\} \cap [0, \infty) = (0, \infty)$.
- $\mathrm{dom}(f - g) = \mathrm{dom}(f) \cap \mathrm{dom}(g) = (0, \infty)$ as above.
- $\mathrm{dom}(fg) = \mathrm{dom}(f) \cap \mathrm{dom}(g) = (0, \infty)$ as above.
- $\mathrm{dom}(f/g) = \{x \in \mathrm{dom}(f) \cap \mathrm{dom}(g) \mid g(x) \neq 0\}$. Now $\mathrm{dom}(f) \cap \mathrm{dom}(g) = (0, \infty)$ as above, and $g(x) \neq 0$ means that $\sqrt{x} - 1 \neq 0$, i.e. that $x \neq 1$. So $\mathrm{dom}(f/g) = \{x \in (0, \infty) \mid x \neq 1\} = (0, 1) \cup (1, \infty)$.

Warning. Never answer a question like "What is the domain of fg?" by first working out a formula for $(fg)(x)$ and then investigating when the formula makes sense and produces a real number. Do it like we did above, by first working out $\mathrm{dom}(f) \cap \mathrm{dom}(g)$ step-by-step.

Question. Let f and g be the functions defined by

$$f(x) = x$$

and

$$g(x) = |x|.$$

1. What is the domain of f? Of g? Of f/g?
2. Write a 'piecewise' formula for f/g.

Solution

1. $\mathrm{dom}(f) = \mathbb{R}$, $\mathrm{dom}(g) = \mathbb{R}$, and

$$\begin{aligned} \mathrm{dom}(f/g) &= \{x \in \mathrm{dom}(f) \cap \mathrm{dom}(g) \mid g(x) \neq 0\} \\ &= \{x \in R \cap R \mid |x| \neq 0\} \\ &= \{x \in \mathbb{R} \mid x \neq 0\} \\ &= (-\infty, 0) \cup (0, \infty). \end{aligned}$$

2. Remember that
$$|x| = \begin{cases} x & \text{if } x \geqslant 0 \\ -x & \text{if } x < 0 \end{cases}$$
so then
$$(f/g)(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$
Let's check this. By the definition $(f/g)(x) = f(x)/g(x) = x/|x|$, so that if $x > 0$ then $|x| = x$ and $(f/g)(x) = x/x = 1$, and if $x < 0$ then $|x| = -x$ and $(f/g)(x) = x/(-x) = -1$.

Definition 1.23 (Composite functions). If f and g are functions, the *composite* function $f \circ g$ is the new function defined by
$$(f \circ g)(x) = f(g(x)).$$
In other words, to work out $(f \circ g)(x)$, we first apply g to work out $g(x)$, and then we apply f to the result to work out $f(g(x))$.

Example 1.24. Let f and g be the functions defined by
$$f(x) = x^2$$
and
$$g(x) = \sin(x).$$
Give formulas for $(f \circ g)(x)$ and $(g \circ f)(x)$.

Solution. We just follow the definitions and substitute in when we can, so that
$$(f \circ g)(x) = f(g(x)) = f(\sin(x)) = (\sin(x))^2 = \sin^2(x)$$
and
$$(g \circ f)(x) = g(f(x)) = g(x^2) = \sin(x^2).$$
Note that $\sin^2(x)$ and $\sin(x^2)$ are not usually equal, so that $f \circ g$ and $g \circ f$ are not the same in this case. And in general, $f \circ g \neq g \circ f$. So you must take care when working out composites: make sure you do it in the right order!

Definition 1.25 (Domains of composite functions). According to our domain convention, the domain of $f \circ g$ consists of all x for which the formula $(f \circ g)(x) = f(g(x))$ makes sense and produces a real number. Thus the domain consists of all x for which $x \in \text{dom}(g)$ and $g(x) \in \text{dom}(f)$. In other words:
$$\text{dom}(f \circ g) = \{x \mid x \in \text{dom}(g) \text{ and } g(x) \in \text{dom}(f)\}.$$

Example 1.26. Let f and g be defined by $f(x) = \sqrt{x}$ and $g(x) = \sqrt{2-x}$. Find the domains of $f \circ g$ and $g \circ g$.

Solution. To find $\operatorname{dom}(f \circ g)$ $\operatorname{dom}(g \circ g)$ we will use the rule given in the previous definition. To do this we need to know $\operatorname{dom}(f)$ and $\operatorname{dom}(g)$, which are given by $\operatorname{dom}(f) = [0, \infty)$ and $\operatorname{dom}(g) = (-\infty, 2]$. So now

$$\begin{aligned}
\operatorname{dom}(f \circ g) &= \{x \mid x \in \operatorname{dom}(g) \text{ and } g(x) \in \operatorname{dom}(f)\} \\
&= \{x \mid x \leqslant 2 \text{ and } \sqrt{2-x} \geqslant 0\} \\
&= \{x \mid x \leqslant 2\} \\
&= (-\infty, 2].
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{dom}(g \circ g) &= \{x \mid x \in \operatorname{dom}(g) \text{ and } g(x) \in \operatorname{dom}(g)\} \\
&= \{x \mid x \leqslant 2 \text{ and } \sqrt{2-x} \leqslant 2\} \\
&= \{x \mid x \leqslant 2 \text{ and } 2-x \leqslant 4\} \\
&= \{x \mid x \leqslant 2 \text{ and } x \geqslant -2\} \\
&= [-2, 2].
\end{aligned}$$

Warning. Do not try to find $\operatorname{dom}(f \circ g)$ by first computing $(f \circ g)(x)$ and then working out the domain using the resulting formula. Why not? Because you could get the wrong answer. Let's see how.

Let a and b be the functions defined by $a(t) = b(t) = \frac{1}{t}$. Then $\operatorname{dom}(a) = \operatorname{dom}(b) = (-\infty, 0) \cup (0, \infty)$. So the domain of $a \circ b$ is

$$\begin{aligned}
\operatorname{dom}(a \circ b) &= \{t \mid t \in \operatorname{dom}(b) \text{ and } b(t) \in \operatorname{dom}(a)\} \\
&= \{t \mid t \neq 0 \text{ and } \frac{1}{t} \neq 0\} \\
&= (-\infty, 0) \cup (0, \infty).
\end{aligned}$$

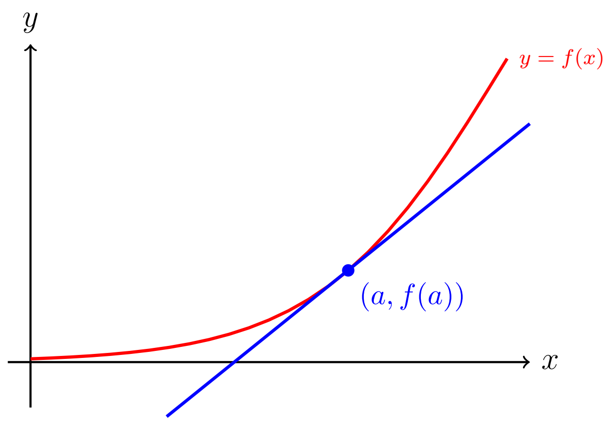
On the other hand, $(a \circ b)(t) = a(1/t) = 1/(1/t) = t$, and this formula makes sense and produces a real number for all t. So if you used this formula to find the domain you would get the wrong answer.

So what is really happening? Well, when working out $(a \circ b)(t)$ we did some *simplification*, and this showed us that $f \circ g$ could be *extended* to include 0 in its domain.

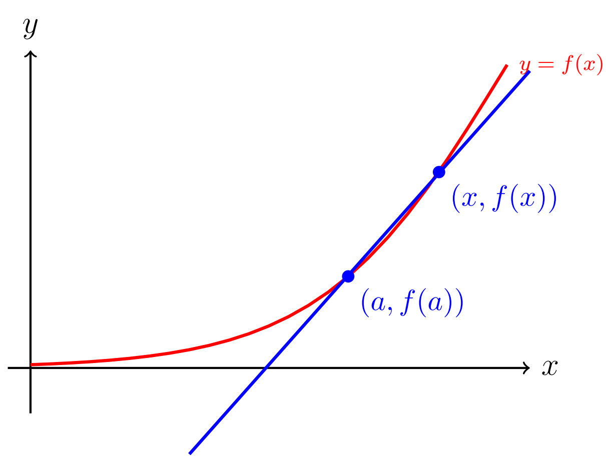
1.4 The limit of a function

Tangents

We are now going to move on to the subject of *limits*. In order to motivate this, we will talk about tangents. Suppose that we want to find the tangent to the graph $y = f(x)$ at a point $(a, f(a))$.



Here, the *tangent* is the line that passes through $(a, f(a))$ and has the same gradient as the curve at that point. The equation of the tangent line will then be $(y - f(a)) = m(x - a)$ where m is the gradient of $y = f(x)$ at $(a, f(a))$. How do we compute m? We can approximate m by choosing an x close to a and considering the *secant* line that passes through $(a, f(a))$ and $(x, f(x))$:



If x is close to a, then the secant line is a good approximation to the tangent line, and so the gradient m_x of the secant is a good approximation to m. Now

$$m_x = \frac{f(x) - f(a)}{x - a}.$$

We would like to set $x = a$ and then obtain $m_a = m$. But the formula for m_x makes no sense in the case $x = a$. Nevertheless, it seems reasonable to expect that as x gets closer to a, m_x gets closer to m, and so we say that m is the *limit* of m_x as x approaches a. (This only works if f is nice enough.) In this part of the course we are going to study and understand this idea of limit.

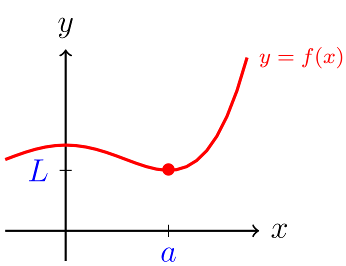
Limits

Definition 1.27 (Limits — the imprecise definition). Suppose that f is a function defined for all x near to, but not necessarily equal to, a number a. We say that the *limit of* $f(x)$ *as* x *approaches* a *is* L, and we write

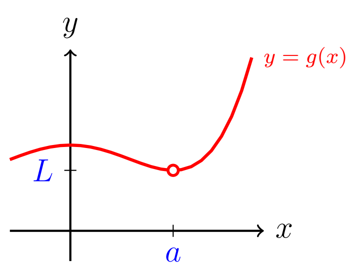
$$\lim_{x \to a} f(x) = L,$$

if we can make the values of $f(x)$ as close to L as we like by choosing x close enough to a. Here, the phrase "x near to a" means "x is in some open interval that contains a".

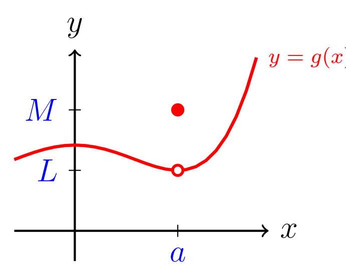
Example 1.28. $f(a)$ does not have to be defined for $\lim_{x\to a} f(x)$ to exist, and even if $f(a)$ is defined, it may not be equal to $\lim_{x\to a} f(x)$. This is shown in the following three examples.



$\mathrm{dom}(f) = (-\infty, \infty)$
$f(a) = L$
$\lim_{x\to a} f(x) = L$



$\mathrm{dom}(g) = (-\infty, 0) \cup (0, \infty)$
a not in $\mathrm{dom}(g)$
$\lim_{x\to a} g(x) = L$



$\mathrm{dom}(h) = (-\infty, \infty)$
$h(a) = M$
$\lim_{x\to a} h(x) = L$

Example 1.29. Convince yourself that for every a it holds that

$$\lim_{x\to a} 1 = 1 \text{ and } \lim_{x\to a} x = a.$$

Example 1.30. Investigate $\lim_{x\to 0} \sin(\dfrac{\pi}{x})$.

Solution. Let f be the function defined by $f(x) = \sin(\frac{\pi}{x})$. Then we have:

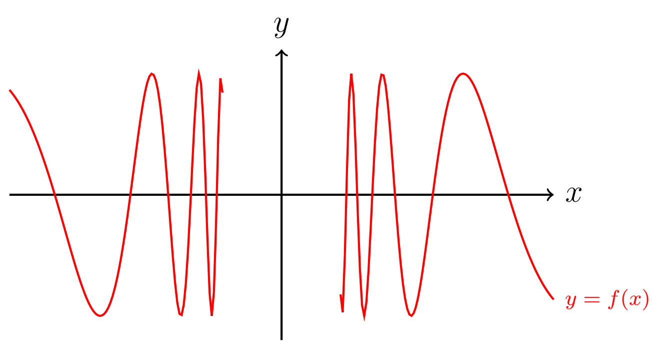
$$f\left(\frac{2}{5}\right) = \sin\left(\frac{\pi}{2/5}\right) = \sin\left(\frac{5\pi}{2}\right) = \sin\left(2\pi + \frac{\pi}{2}\right) = \sin(\pi/2) = 1$$
$$f\left(\frac{2}{9}\right) = \sin\left(\frac{\pi}{2/9}\right) = \sin\left(\frac{9\pi}{2}\right) = \sin\left(4\pi + \frac{\pi}{2}\right) = \sin(\pi/2) = 1$$
$$f\left(\frac{2}{13}\right) = \sin\left(\frac{\pi}{2/13}\right) = \sin\left(\frac{13\pi}{2}\right) = \sin\left(6\pi + \frac{\pi}{2}\right) = \sin(\pi/2) = 1$$

and so on. This shows that we can find numbers x that are as close to 0 as we wish with $f(x) = 1$. Similar calculations show that

$$f\left(\frac{2}{3}\right) = -1$$
$$f\left(\frac{2}{7}\right) = -1$$
$$f\left(\frac{2}{11}\right) = -1$$

and so on, so that we can find numbers x that are as close to 0 as we wish with $f(x) = -1$. This means that the limit *does not exist*: there is no one L for which $f(x)$ gets closer and closer to L as x gets closer and closer to 0.

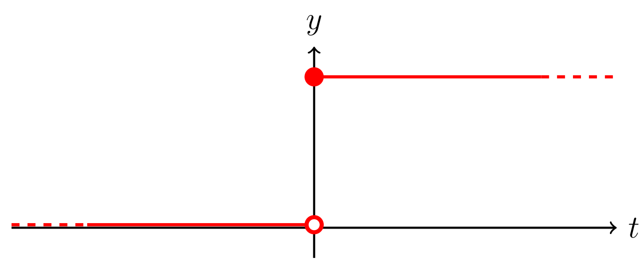
This is shown pretty clearly on the graph of $y = f(x)$. The graph oscillates ever more rapidly as you approach 0, so it does not get close any single value.



Example 1.31. The *Heaviside function* H is defined as follows:

$$H(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t \geqslant 0 \end{cases}$$

Its graph is as follows.



In this case $\lim_{t\to 0} H(t)$ does not exist. This is because there are numbers t as close to 0 as we like with $H(t) = 0$ (any negative number) and numbers t as close to 0 as we like with $H(t) = 1$ (any positive number). So there is no single number L such that we can make $H(t)$ as close to L as we like by making t sufficiently close to 0.

Definition 1.32 (One-sided limits). Let f be a function that is defined for all x close to, and less than, a number a. We write

$$\lim_{x\to a^-} f(x) = L$$

and say *the limit of $f(x)$ as x approaches a from the left is L* if we can make $f(x)$ as close to L as we wish by making x sufficiently close to, and less than, a. Here the phrase 'for all x close to and less than a' means 'for x in some open interval of the form (b, a)'.

Now let f be a function that is defined for all x close to, and greater than, a number a. We write

$$\lim_{x\to a^+} f(x) = L$$

and say *the limit of $f(x)$ as x approaches a from the right is L* if we can make $f(x)$ as close to L as we wish by making x sufficiently close to, and greater than, a. Here the phrase 'for all x close to and greater than a' means 'for x in some open interval of the form (a, b)'.

The first limit is 'from the left' and its definition involves the phrase 'less than', while the second limit is 'from the right' and its definition involves the phrase 'greater than'. That is because if we draw x and a on the number line, then $x < a$ means x is to the left of a, while $x > a$ means that x is to the right of a.

Example 1.33. Let H be the Heaviside function defined above by

$$H(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t \geqslant 0 \end{cases}$$

What are $\lim\limits_{t \to 0^-} H(t)$ and $\lim\limits_{t \to 0^+} H(t)$?

Solution. $\lim_{t \to 0^-} H(t) = 0$. This is because, however close to 0 we would like $H(t)$ to be, we can achieve that by taking $t < 0$ close enough to 0. Indeed, for any $t < 0$ we have $H(t) = 0$.

Similarly, $\lim_{t \to 0^+} H(t) = 1$. This is because, however close to 1 we would like $H(t)$ to be, we can achieve that by taking $t > 0$ close enough to 0. Indeed, for any $t > 0$ we have $H(t) = 1$.

Here is an important rule about limits.

- The limit $\lim_{x \to a} f(x)$ exists if and only if $\lim_{x \to a^-} f(x)$ and $\lim_{x \to a^+} f(x)$ both exist and are equal, in which case

$$\lim_{x \to a} f(x) = \lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x).$$

This rule means that if $\lim_{x \to a^+} f(x)$ does not exist, or if $\lim_{x \to a^-} f(x)$ does not exist, or if $\lim_{x \to a^+} f(x)$ and $\lim_{x \to a^-} f(x)$ do exist but are not equal, then $\lim_{x \to a} f(x)$ does not exist.

Example 1.34. Let f be the function defined by $f(x) = \sin(\frac{\pi}{x})$. Then $\lim_{x \to 0^+} f(x)$ does not exist, for exactly the same reasons given before. So $\lim_{x \to 0} f(x)$ does not exist.

Example 1.35. $\lim_{t \to 0^-} H(t) = 0$ and $\lim_{t \to 0^+} H(t) = 1$ do both exist, but they are not equal. So $\lim_{t \to 0} H(t)$ does not exist.

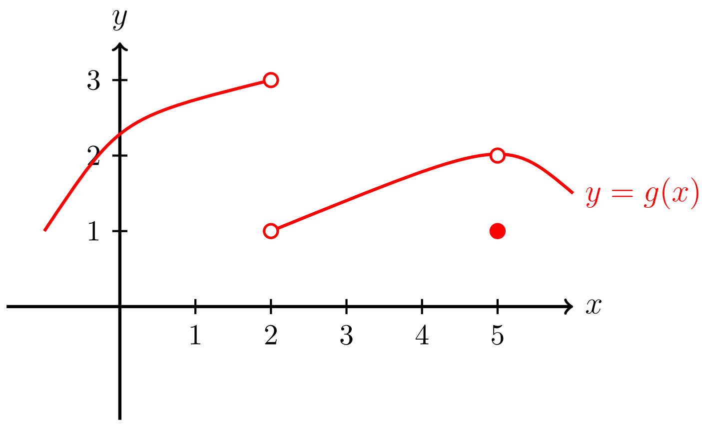
Example 1.36. Show that $\lim\limits_{x \to 0} |x| = 0$.

Solution. Remember that the absolute value function is defined by:

$$|x| = \begin{cases} x & \text{if } x \geqslant 0 \\ -x & \text{if } x < 0 \end{cases}$$

The limit $\lim_{x\to 0^-}|x|$ depends only on the values of $|x|$ for $x<0$. But if $x<0$ then $|x|=-x$, and so $\lim_{x\to 0^-}|x|=\lim_{x\to 0^-}(-x)=0$. Similarly, the limit $\lim_{x\to 0^+}|x|$ depends only on the values of $|x|$ for $x>0$. But if $x>0$ then $|x|=x|$, and so $\lim_{x\to 0^+}|x|=\lim_{x\to 0^+}x=0$. Since $\lim_{x\to 0^+}|x|$ and $\lim_{x\to 0^-}|x|$ exist and are equal, we have $\lim_{x\to 0}|x|=0$.

Example 1.37. Let g be the function whose graph is as follows.



State the values of the following, if they exist.

(a) $\lim\limits_{x\to 2^-} g(x)$ (b) $\lim\limits_{x\to 2^+} g(x)$ (c) $\lim\limits_{x\to 2} g(x)$ (d) $g(2)$

(e) $\lim\limits_{x\to 5^-} g(x)$ (f) $\lim\limits_{x\to 5^+} g(x)$ (g) $\lim\limits_{x\to 5} g(x)$ (h) $g(5)$

Solution.

(a) 3 (b) 1 (c) The limit does not exist. (d) $g(2)$ is not defined

(e) 2 (f) 2 (g) 2 (h) 1

Example 1.38. Define the function f as follows.

$$f(x)=\begin{cases} x+1 & \text{if } x\leqslant 0\\ 2x+1 & \text{if } 0<x\leqslant 1\\ x^2 & \text{if } 1<x\end{cases}$$

Which limits exist, and what are their values?

(a) $\lim\limits_{x\to 0} f(x)$ (b) $\lim\limits_{x\to 1} f(x)$ (c) $\lim\limits_{x\to \frac{1}{2}} f(x)$

Solution. 1. The piecewise definition means that we can always work out the left and right limits. For example, when we compute $\lim_{x\to 0^-} f(x)$, we can assume that x is close to 0 and less than 0, so that the first 'piece' of f applies and $f(x) = x + 1$. Thus

$$\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-} (x+1) = 1.$$

Similarly, when we compute $\lim_{x\to 0^+} f(x)$ we can assume that $x > 0$ and that x is close to 0, so that the second 'piece' of f applies and $f(x) = 2x + 1$. Thus

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} (2x+1) = 1.$$

Consequently, since the left and right limits exist and are both equal to 1, we have

$$\lim_{x\to 0} f(x) = 1.$$

2. In this case we have

$$\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} (2x+1) = 2 + 1 = 3$$

and

$$\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} x^2 = 1^2 = 1.$$

Since the left and right limits are different,

$$\lim_{x\to 1} f(x) \text{ does not exist.}$$

3. In this case we have $\lim_{x\to 1/2} f(x)$. (Here 1/2 is 'in the middle' of one of the pieces of f, rather than being a point where the definition changes.) In computing this limit we may assume that x is close to 1/2, so that the middle 'piece' of f applies and

$$\lim_{x\to \frac{1}{2}} f(x) = \lim_{x\to \frac{1}{2}} (2x+1) = 2 \times \frac{1}{2} + 1 = 2.$$

Calculating limits using the limit laws

Suppose that $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ exist. Let c be a constant. Then we have the following *limit laws*, which are rules for computing limits.

- **Sum rule.**

$$\lim_{x\to a}[f(x)+g(x)] = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$$

- **Difference rule.**

$$\lim_{x\to a}[f(x)-g(x)] = \lim_{x\to a} f(x) - \lim_{x\to a} g(x)$$

- **Scalar rule.**

$$\lim_{x\to a}[c\cdot f(x)] = c\cdot \lim_{x\to a} f(x)$$

- **Product rule.**

$$\lim_{x\to a}[f(x)g(x)] = \left[\lim_{x\to a} f(x)\right]\cdot\left[\lim_{x\to a} g(x)\right]$$

- **Quotient rule.**

$$\lim_{x\to a}\left[\frac{f(x)}{g(x)}\right] = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$$

  so long as $\lim_{x\to a} g(x) \neq 0$.

Remark 1.39. These limit laws, and the ones that follow, all also work for one-sided limits. (In other words, take a limit law, and replace every instance of '$x \to a$' with '$x \to a^-$' or '$x \to a^+$' and the result is still a true law.

Definition 1.40 (Polynomials and rational functions)*.* A *polynomial* is a function p defined by a formula of the form

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

where $a_n, \ldots, a_0$ are real numbers. A *rational function* is a function r defined by a formula of the form

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

where p and q are polynomials. The domain of any polynomial is $\mathbb{R}$, and the domain of the rational function r above is $\{x \mid q(x) \neq 0\}$, unless stated otherwise.

Using the limit laws above, we see that if $p(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$ is a polynomial, then

$$\lim_{x\to a} p(x) = b_n \lim_{x\to a} x^n + b_{n-1} \lim_{x\to a} x^{n-1} + \cdots + b_1 \lim_{x\to a} x + \lim_{x\to a} b_0 =$$

$$b_n(\lim_{x\to a} x)^n + b_{n-1}(\lim_{x\to a} x)^{n-1} + \cdots + b_1 \lim_{x\to a} x + \lim_{x\to a} b_0 =$$

$$b_n a^n + b_{n-1}a^{n-1} + \cdots b_1 a + b_0 = p(a).$$

Similarly, if q is another polynomial which satisfies $q(a) \neq 0$ then we have that

$$\lim_{x\to a} \frac{p(x)}{q(x)} = \frac{\lim_{x\to a} p(x)}{\lim_{x\to a} q(x)} = \frac{p(a)}{q(a)}.$$

We summarize this:

- **Direct Substitution Law.** Let f be a rational function and let a lie in the domain of f. Then

  $$\lim_{x\to a} f(x) = f(a).$$

  (This rule applies to many more functions than just the rational functions, as we will see later in the course.)

- Let f and g be functions such that $f(x) = g(x)$ for x close to, but not necessarily equal to, a. Then

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

  Here, "close to a" means "in an open interval containing a".

Example 1.41. Evaluate $\lim_{x\to -2} \frac{x^2 - 4}{x^2 + 6x + 8}$.

Solution. We can't use direct substitution to do this. If we tried it, we would find

$$\lim_{x\to -2} \left[\frac{x^2 - 4}{x^2 + 6x + 8}\right] = \frac{(-2)^2 - 4}{(-2)^2 + 6(-2) + 8} = \frac{0}{0}.$$

The result is nonsense, so we know that we went wrong somewhere. Indeed, -2 is not in the domain of the function, so that direct substitution was not permitted in the first place. However, our failed computation gives us a clue: substituting $x = -2$ makes x^2-4 and x^2+6x+8 equal to 0, and consequently $(x-(-2)) = (x+2)$ is a factor of both. In fact, for $x \neq -2$, we have

$$\frac{x^2-4}{x^2+6x+8} = \frac{(x-2)(x+2)}{(x+2)(x+4)} = \frac{x-2}{x+4}.$$

So by the last limit law, we have

$$\lim_{x\to-2}\left[\frac{x^2-4}{x^2+6x+8}\right] = \lim_{x\to-2}\left[\frac{(x-2)(x+2)}{(x+2)(x+4)}\right] = \lim_{x\to-2}\left[\frac{x-2}{x+4}\right] = \frac{-2-2}{-2+4} = \frac{-4}{2} = -2.$$

In this computation, we were allowed to use direct substitution to compute $\lim_{x\to-2}\left[\frac{x-2}{x+4}\right]$ because -2 lies in the domain of the function. We know that -2 lies in the domain because when we substitute $x = -2$ the denominator is not 0.

Example 1.42. Calculate the value of $\lim\limits_{x\to1}\dfrac{x^2-1}{x-1}$.

Solution. Observe that $\frac{x^2-1}{x-1} = \frac{(x-1)(x+1)}{x-1} = x+1$ for $x \neq 1$. Since the limit as x goes to 1 does not depend on what happens *at* 1, this is enough to show that $\lim_{x\to1}\frac{x^2-1}{x-1} = \lim_{x\to1}(x+1)$. Now $\lim_{x\to1}(x+1) = 1+1 = 2$, by the limit laws.

Definition 1.43 (Infinite limits). Let f be defined near to a, except possibly at a itself. We say

$$\lim_{x\to a} f(x) = \infty$$

if we can make $f(x)$ as large and positive as we wish by making x sufficiently close to, but not equal to, a. And we say

$$\lim_{x\to a} f(x) = -\infty$$

if we can make $f(x)$ as large and negative as we wish by making x sufficiently close to, but not equal to, a. We leave it to the reader to define the following similar notions.

$$\lim_{x\to a^+} f(x) = \infty \qquad \lim_{x\to a^-} f(x) = \infty \qquad \lim_{x\to a^+} f(x) = -\infty \qquad \lim_{x\to a^-} f(x) = -\infty$$

Warning. The symbol ∞ appears above only as part of the 'phrases' $\lim_{x\to a} f(x) = \infty$ and $\lim_{x\to a} f(x) = -\infty$. It does not appear elsewhere or in any other way. Remember, ∞ is not a number!

Example 1.44. $\lim_{x\to 0} \frac{1}{x^2} = \infty$. This is because if K is a large positive number, then if we want to make sure that $\frac{1}{x^2} > K$ it is enough to make sure that x is in the range $-\frac{1}{\sqrt{K}} < x < \frac{1}{\sqrt{K}}$ and $x \neq 0$.

Example 1.45. $\lim_{x\to 0^+} \frac{1}{x} = \infty$ and $\lim_{x\to 0^-} \frac{1}{x} = -\infty$. For if x is small and positive, then $\frac{1}{x}$ is large and positive, as large as we like if x is close enough to 0, so that $\lim_{x\to 0^+} \frac{1}{x} = \infty$. Similarly for the second case. Consequently, $\lim_{x\to 0} \frac{1}{x}$ does not exist, as a finite or infinite limit.

Example 1.46. Compute $\displaystyle\lim_{x\to 3^+} \frac{1}{x-3}$ and $\displaystyle\lim_{x\to 3^-} \frac{1}{x-3}$.

Solution. If x is close to 3, but greater than 3, then $x - 3$ is small and positive, so $\frac{1}{x-3}$ is large and positive, and so $\lim_{x\to 3^+} \frac{1}{x-3} = \infty$. If x is close to 3, but less than 3, then $x - 3$ is small and negative, so $\frac{1}{x-3}$ is large and negaive, and so $\lim_{x\to 3^-} \frac{1}{x-3} = -\infty$.

Example 1.47. Let f be the function defined by $f(x) = \frac{x^2-4x+3}{x^2-9}$. Think about $\lim_{x\to 3} f(x)$, $\lim_{x\to -3} f(x)$ and, if necessary, the left and right versions of these limits.

Solution. First, let's consider what happens when we make a substitution.

- When we substitute $x = 3$, we find that $x^2 - 4x + 3 = 0$ and $x^2 - 9 = 0$.
- When we substitute $x = -3$, we find that $x^2 - 4x + 3 = 24$ and $x^2 - 9 = 0$.

Remember that if a polynomial becomes 0 when we substitute $x = a$, then you know that $(x - a)$ is a factor of that polynomial. Indeed,

$$f(x) = \frac{x^2 - 4x + 3}{x^2 - 9} = \frac{(x-3)(x-1)}{(x-3)(x+3)} = \frac{x-1}{x+3}$$

so long as x is not equal to 3 or -3. But the limits in question do not depend on what happens at $x = 3$ or $x = -3$, and so

$$\lim_{x\to 3} f(x) = \lim_{x\to 3} \frac{x-1}{x+3} = \frac{3-1}{3+3} = \frac{1}{3}$$

Now as x approaches -3 from the left, $x-1$ will be approximately -4, and $x+3$ will be small and negative. So $\frac{x-1}{x+3}$ will be large and positive. Thus

$$\lim_{x\to -3^-} f(x) = \lim_{x\to -3^-} \frac{x-1}{x+3} = \infty$$

And as x approaches -3 from the right, $x-1$ will be approximately -4, and $x+3$ will be small and positive. So $\frac{x-1}{x+3}$ will be large and negative. Thus

$$\lim_{x\to -3^+} f(x) = \lim_{x\to -3^+} \frac{x-1}{x+3} = -\infty.$$

And finally $\lim_{x\to -3} f(x)$ does not exist, either as a finite or an infinite limit.

Definition 1.48. Let f be a function which is defined for all x which is positive and large enough. In other words: there exists an $r \in \mathbb{R}$ such that $f(x)$ is defined for all $x > r$. We say

$$\lim_{x\to\infty} f(x) = L$$

if we can make $f(x)$ close to L as we wish by taking x to be large enough.

We define $\lim_{x\to -\infty} f(x) = L$ in a similar fashion. We next define infinte limits at infinity:

Definition 1.49. Let f be a function which is defined for all x which is positive and large enough. We say

$$\lim_{x\to\infty} f(x) = \infty$$

if we can make $f(x)$ as large as we wish by taking x to be large enough.

We can similarly define

$$\lim_{x\to -\infty} f(x) = \infty, \lim_{x\to -\infty} f(x) = -\infty, \text{ and } \lim_{x\to \infty} f(x) = -\infty.$$

Example 1.50. Convince yourself that $\lim_{x\to\infty} 1 = 1$ and that $\lim_{x\to\infty} x = \infty$.

Remark 1.51. The limit laws still work if one replaces a with ∞ or $-\infty$, assuming that the limits exist and are finite. Thus, if $\lim_{x\to\infty} f(x)$ and $\lim_{x\to\infty} g(x)$ exist and finite, then $lim_{x\to\infty} f(x) + g(x)$ exist and is equal to $\lim_{x\to\infty} f(x) + \lim_{x\to\infty} g(x)$.

Example 1.52. Calculate the limit

$$\lim_{x\to\infty} \frac{x+1}{2x+3}.$$

Solution. Since $\lim_{x\to\infty}(x+1) = \lim_{x\to\infty} 2x+3 = \infty$ we can not just use the quotient rule. Indeed, we will receive the phrase $\frac{\infty}{\infty}$ which has no clear meaning. Instead, we will rewrite the function $f(x) = \frac{x+1}{2x+3}$ as

$$f(x) = \frac{\frac{1}{x}(1+\frac{1}{x})}{\frac{1}{x}(2+\frac{3}{x})}.$$

Notice that in order to write f in this way we need to assume that $x \neq 0$. Since we are studying the limit for $x \to \infty$ this does not matter for us, since we only care about the values of $f(x)$ for large values of x. The function f can thus be written as $f(x) = \frac{1+\frac{1}{x}}{2+\frac{3}{x}}$. So we wrote f as a quotient of two functions which do have finite limits at ∞. It holds that $\lim_{x\to\infty} 1 + \frac{1}{x} = 1$ and $\lim_{x\to\infty}(2+\frac{3}{x}) = 2$. It holds by the quotient rule that $\lim_{x\to\infty} f(x) = \frac{1}{2}$.

How to approach limit questions

Let's imagine that we have the following imaginary question about a function f.

> **Question:** Does $\lim_{x\to a} f(x)$ exist as a finite or infinite limit, and if so, what is its value?

Here is how we try to solve this question:

- If f is given by a single formula, begin by substituting $x = a$.
- If $f(a)$ is defined, and f is sufficiently nice (see the Direct Substitution rule later; quotients of polynomials are always sufficiently nice) then $\lim_{x\to a} f(x) = f(a)$.
- If $f(a)$ is not defined, then attempt to simplify, and start the process again. (Note: If substituting $x = a$ into a polynomial produces 0, then $(x - a)$ is a factor of the polynomial.
- If f is defined piecewise and a lies at the 'join' of two pieces, then first consider $\lim_{x\to a^-} f(x)$ and $\lim_{x\to a^+} f(x)$.

There is not always a systematic approach to computing limits, for example $\lim_{x\to 0} \frac{\sin(x)}{x}$, which we will compute later from first principles. Often you may have to try different approaches before finding the one that works. However, you will only be asked questions that you are capable of solving!

Example 1.53. Suppose that f and g are functions defined near to, but not necessarily at, the number a, and suppose that $\lim_{x\to-2} f(x) = 1$ and $\lim_{x\to-2} g(x) = -1$. Use limit laws to evaluate the following.

1. $\lim_{x\to-2}[f(x) + 5g(x)]$

2. $\lim_{x\to-2}\left[\frac{f(x)}{g(x)}\right]$

Solution. 1.

$$\begin{aligned}
\lim_{x\to-2}[f(x) + 5g(x)] &= \lim_{x\to-2}[f(x) + [5g(x)]] \\
&= \lim_{x\to-2} f(x) + \lim_{x\to-2}[5g(x)] && \text{by sum law} \\
&= \lim_{x\to-2} f(x) + 5 \cdot \lim_{x\to-2} g(x) && \text{by scalar law} \\
&= 1 + 5 \cdot (-1) \\
&= -4.
\end{aligned}$$

2.

$$\begin{aligned}
\lim_{x\to-2}\left[\frac{f(x)}{g(x)}\right] &= \frac{\lim_{x\to-2} f(x)}{\lim_{x\to-2} g(x)} && \text{by quotient law} \\
&= \frac{1}{-1} \\
&= -1
\end{aligned}$$

Remark 1.54. In the last solution, we should really have checked that at each step, when we used a limit law, the limits in question existed. However, in each case this was checked by the remaining steps of the computation. So all is well, and we allow ourselves to tackle the problems in this way.

Warning. If you are computing a limit using limit laws, and arrive at an expression that makes no sense, or a limit that does not exist, then discard the computation and try something else.

Here are a further set of limit laws, some of which you have already seen before. We again assume that $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ exist and are finite, and that c is any number.

- **Power law.** Let n be a positive integer. Then:

$$\lim_{x\to a}[f(x)]^n = \left[\lim_{x\to a} f(x)\right]^n$$

- **Constant.**

$$\lim_{x\to a} c = c.$$

- **Identity law.**

$$\lim_{x\to a} x = a.$$

- **Consequently**

$$\lim_{x\to a} x^n = a^n$$

  for any positive integer n.

- **And similarly**

$$\lim_{x\to a} \sqrt[n]{x} = \sqrt[n]{a}$$

  for any positive integer n, where $a > 0$ if n is even.

- **An more generally**

$$\lim_{x\to a} \sqrt[n]{f(x)} = \sqrt[n]{\lim_{x\to a} f(x)}$$

  where n is a positive integer and, if n is even, then $\lim_{x\to a} f(x) > 0$.

Remark 1.55. The purpose of the previous computation was to show how the various different limit laws work. However, if we compare the first and fifth steps of the above computation we have the following.

$$\lim_{x\to -2} \frac{x^3 + 2x^2 - 1}{5 - 3x} \qquad \text{and} \qquad \frac{[-2]^3 + 2[-2]^2 - 1}{5 - 3[-2]}$$

Looking at these two expressions, it seems that we should be able to go from one to the other by directly substituting $x = -2$. That is the message of the direct substitution law below.

Example 1.56. Evaluate $\lim_{t\to 0} \dfrac{\sqrt{t^2+9} - 3}{t^2}$.

Solution. Substituting $t = 0$ into the formula gives us $\frac{0}{0}$, so we need to do something else. Factorizing top and bottom of the fractions also would not help. Instead we use the following trick. **Try to understand and remember the trick!** For $t \neq 0$ we have:

$$\begin{aligned}
\frac{\sqrt{t^2+9}-3}{t^2} &= \frac{\sqrt{t^2+9}-3}{t^2} \cdot \frac{\sqrt{t^2+9}+3}{\sqrt{t^2+9}+3} \\
&= \frac{(\sqrt{t^2+9}-3)(\sqrt{t^2+9}+3)}{t^2(\sqrt{t^2+9}+3)} \\
&= \frac{(t^2+9)-9}{t^2(\sqrt{t^2+9}+3)} \\
&= \frac{t^2}{t^2(\sqrt{t^2+9}+3)} \\
&= \frac{1}{\sqrt{t^2+9}+3}
\end{aligned}$$

This means that

$$\lim_{t\to 0} \frac{\sqrt{t^2+9}-3}{t^2} = \lim_{t\to 0} \frac{1}{\sqrt{t^2+9}+3} = \frac{1}{\sqrt{9}+3} = \frac{1}{3+3} = \frac{1}{6}.$$

Here the last step follows from a version of direct substitution that we will see shortly.

Theorem 1.57. *Suppose that $f(x) \leqslant g(x)$ for all x close to, but not necessarily equal to, a. Then $\lim_{x\to a} f(x) \leqslant \lim_{x\to a} g(x)$, assuming that both limits exist.*

Warning. You cannot replace the two instances of $\leqslant$ in the theorem with $<$, because then the theorem fails. Can you see an example of this?

Theorem 1.58. *(Squeeze Theorem)*
Suppose that

$$f(x) \leqslant g(x) \leqslant h(x)$$

for x close to, but not necessarily equal to, a. Suppose also that

$$\lim_{x\to a} f(x) = L = \lim_{x\to a} h(x).$$

Then $\lim_{x\to a} g(x)$ exists and is equal to L.

Example 1.59. Show that $\lim_{x\to 0} x^2 \sin\left(\frac{1}{x}\right) = 0$.

Solution. Define $g(x) = x^2 \sin(1/x)$. (The squeeze theorem tells us about the function g, and we want to know about $x^2 \sin(1/x)$, which is why we made this choice.) Define $f(x) = -x^2$ and $h(x) = x^2$. Since $-1 \leqslant \sin(1/x) \leqslant 1$ and $x^2 \geqslant 0$, we have

$$-x^2 \leqslant x^2 \sin(1/x) \leqslant x^2.$$

In other words,

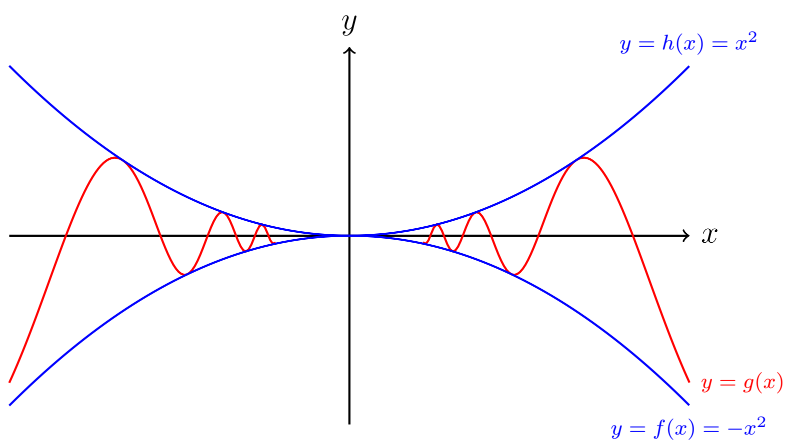
$$f(x) \leqslant g(x) \leqslant h(x).$$

Also, $\lim_{x\to 0} f(x) = \lim_{x\to 0}(-x^2) = 0$ and $\lim_{x\to 0} h(x) = \lim_{x\to 0}(x^2) = 0$. So the squeeze theorem applies (with f, g and h as specified above, with $a = 0$, and with $L = 0$), and tells us that $\lim_{x\to 0} x^2 \sin(1/x) = 0$.

Example 1.60. Suppose you had been asked to find $\lim_{x\to 0} |x|\cos(1/x)$ using the squeeze theorem. You would put $g(x) = |x|\cos(1/x)$, because the squeeze theorem tells us about the limit of g. But what f and h would you choose?

Solution. We would choose $f(x) = -|x|$ and $h(x) = |x|$. Then the squeeze theorem could be applied, exactly as in the previous example, but this time using the inequalities $-1 \leqslant \cos(1/x) \leqslant 1$ and $|x| \geqslant 0$.

Example 1.61. Here is a 'picture' of the squeeze theorem at work, in the case where $g(x) = x^2 \sin(1/x)$ as in Example 1.59. For this we took $f(x) = -x^2$ and $h(x) = x^2$. Here are the graphs of the three functions, with $y = f(x)$ at the bottom in blue and $y = h(x)$ at the top in blue, and $y = g(x)$ between them in red.



The inequality $f(x) \leqslant g(x) \leqslant h(x)$ translates to the fact that the red graph is squeezed between the two blue graphs. The conclusion of the squeeze theorem, that $\lim_{x\to 0} g(x) = 0$, is now immediately clear from the graphs.

Example 1.62. Show that $\lim_{x\to 0} x\sin(1/x) = 0$.

Solution. (Note: This example is harder than the previous two. The reason for this is that in the previous examples we had $x^2 \geqslant 0$ and $|x| \geqslant 0$ respectively, whereas now the analogous inequality $x \geqslant 0$ is not true.)

Define g by $g(x) = x\sin(1/x)$, and define f and h as follows.

$$f(x) = \begin{cases} -x & \text{if } x \geqslant 0 \\ x & \text{if } x < 0 \end{cases} \qquad \text{and} \qquad h(x) = \begin{cases} x & \text{if } x \geqslant 0 \\ -x & \text{if } x < 0 \end{cases}$$

Since $-1 \leqslant \sin(1/x) \leqslant 1$, we have:

- If $x \geqslant 0$, then $-x \leqslant x\sin(1/x) \leqslant x$.
- If $x < 0$, then $-x \geqslant x\sin(1/x) \geqslant x$, or in other words, $x \leqslant x\sin(1/x) \leqslant -x$.

The first bullet point says that $f(x) \leqslant g(x) \leqslant h(x)$ when $x \geqslant 0$, and the second bullet point says that $f(x) \leqslant g(x) \leqslant h(x)$ when $x < 0$. So $f(x) \leqslant g(x) \leqslant h(x)$ holds for all x. But note that $\lim_{x\to 0} f(x) = 0$ and $\lim_{x\to 0} h(x) = 0$, as we see for example by inspecting the left and right hand limits of each one. So the squeeze theorem applies and tells us that $\lim_{x\to 0} g(x) = 0$ as required.

Limits: rules of thumb

To conclude, we give here a summary of what is allowed, and what is not allowed, to do when calculating limits involving infinity. We write this here in an a slightly imprecise way. As we mentioned, we cannot treat ∞ as a number, so for each rule we say exactly what we mean:

- If $\lim_{x\to a} f(x) = \infty$ and $\lim_{x\to a} g(x) = \infty$ then
$$\lim_{x\to a} f(x) + g(x) = \lim_{x\to a} f(x)g(x) = \infty.$$
  We write this informally as $\infty + \infty = \infty$ and $\infty \cdot \infty = \infty$.

- If $\lim_{x\to a} f(x) = \infty$ and $\lim_{x\to a} g(x) = c > 0$, then
$$\lim_{x\to a} f(x)g(x) = \lim_{x\to a} \frac{f(x)}{g(x)} = \infty.$$

- If $\lim_{x\to a} f(x) = c$ and $\lim_{x\to a} g(x) = \infty$ then $\lim_{x\to a} \frac{f(x)}{g(x)} = 0$.

- If $\lim_{x\to a} f(x) = c > 0$ and $\lim_{x\to a} g(x) = 0$, and $g(x) > 0$ for x close enough to a, then $\lim_{x\to a} \frac{f(x)}{g(x)} = \infty$. Similar results hold when $c < 0$, or $g(x) < 0$, or both. We just need to take care of the sign.

- Assume that $f(x) \leq g(x)$ when x is close enough to a number a. If $\lim_{x\to a} f(x) = \infty$ then $\lim_{x\to a} g(x) = \infty$.

Now for a list of problematic situations. In these situations we cannot calculate the limit directly using the limit rules, but we have to simplify the function somehow first.

- The limit $\lim_{x\to a} \frac{f(x)}{g(x)}$ when $\lim_{x\to a} f(x) = \infty = \lim_{x\to a} g(x) = \infty$ or $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$. We say informally that $\frac{\infty}{\infty}$ and $\frac{0}{0}$ are not defined.

- The limit $\lim_{x\to a} f(x) - g(x)$ where $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \infty$. We say informally that $\infty - \infty$ is not defined.

- The limit $\lim_{x\to a} f(x)g(x)$ where $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = \infty$. We say informally that $0 \cdot \infty$ is not defined.

1.5 The precise definition of the limit

Definition 1.63 (The precise definition of limit). Let f be a function defined on some open interval containing a, except possibly at a itself. Then we say that *the limit of f as x approaches a is L*, and write
$$\lim_{x\to a} f(x) = L,$$
if the following condition holds.

> For every number $\epsilon > 0$, there is a number $\delta > 0$ such that $|f(x) - L| < \epsilon$ whenever $0 < |x - a| < \delta$.

Question. The following questions are supposed to help you to understand the definition of the limit.

1. The set of all y satisfying $|y - L| < \epsilon$ is an interval. Which interval?
2. Think of ϵ as a small number. Describe in words what it means for y to satisfy the condition $|y - L| < \epsilon$. Do not use ϵ in your answer.
3. Now think of δ as a small number. Describe in words what it means for x to satisfy $0 < |x - a| < \delta$. Do not use δ in your answer.

Solution. Please try to think these through before looking at the answers!

1. It is the interval $(L - \epsilon, L + \epsilon)$.
2. It means that y is close to L.
3. It means that x is close to a but not equal to a.

So now let's try to understand the definition of $\lim_{x\to a} f(x) = L$. The original version is:

> For every number $\epsilon > 0$, there is a number $\delta > 0$ such that $|f(x) - L| < \epsilon$ whenever $0 < |x - a| < \delta$.

And we think of it as saying:

> We can make $f(x)$ as close to L as we like by making x close enough to a, but not necessarily equal to a.

Here, the red part of the second sentence corresponds to the red part of the original version, and so on.

Let us understand what exactly we need to show when we want to prove that

$$\lim_{x\to a} f(x) = L.$$

The definition of the limit has the following form:
For every number $\epsilon > 0$ there is a number $\delta > 0$ such that SOMETHING HAPPENS.
At first glance, it looks as if there are infinitely many statement that we have to prove here, one statement for each possible value of ϵ: We have to prove that

For $\epsilon = 0.5$ there is a number $\delta > 0$ such that Something happens
and that
For $\epsilon = 0.1$ there is a number $\delta > 0$ such that Something happens
and that
For $\epsilon = 0.02$ there is a number $\delta > 0$ such that Something happens
and so on. However, we are going to prove all these statements at once. Here is an example:

Example 1.64. Use the precise definition of the limit to show that $\lim_{x\to 5} x = 5$.

Solution. We write first exactly what we need to prove:
For every number $\epsilon > 0$ there is a number $\delta > 0$ such that $|f(x) - L| < \epsilon$ whenever $0 < |x - a| < \delta$.
Here $f(x) = x$, $a = 5$ and $L = 5$. So we need to prove the following statement:
For every number $\epsilon > 0$ there is a number $\delta > 0$ such that $|x - 5| < \epsilon$ whenever $0 < |x - 5| < \delta$.
We need to prove this statement for every $\epsilon > 0$. So for $\epsilon = 0.1$ we need to find a number $\delta > 0$ such that if $0 < |x - 5| < \delta$ then $|x - 5| < 0.1 = \epsilon$. Such a number is $\delta = \epsilon = 0.1$. For $\epsilon = 0.01$ we need to fine a number $\delta > 0$ such that if $0 < |x - 5| < \delta$ then $|x - 5| < 0.01 = \epsilon$. Such a number is $\delta = \epsilon = 0.01$. We see that we do not have to go through all possible values of ϵ. We just need to say what δ we choose, given ϵ. In this case, we see that we can always choose $\delta = \epsilon$. Notice that usually the value of δ depends on ϵ.

So to conclude this, let us see what we need to do in order to show that $\lim_{x\to 5} x = 5$: We start with a number $\epsilon > 0$. We choose $\delta = \epsilon$. Then we have that $|x - 5| < \epsilon$ whenever $0 < |x - 5| < \delta$. This finishes the proof.

Example 1.65. Use the precise definition of the limit to show that $\lim_{x\to 0} 1 = 1$.

Solution. In this case we have $a = 0$, $L = 1$ and $f(x) = 1$ the constant function. Notice that $f(x) - L = 1 - 1 = 0$ is constant zero, and does not depend on x. So to prove the limit is 1, we need to prove the following:
For every number $\epsilon > 0$ there is a number $\delta > 0$ such that $0 < \epsilon$ whenever $|x| < \delta$.
But it is always true that $0 < \epsilon$, and therefore we can choose whatever value of $\delta > 0$ we want. The statement will be true.

Let us see a bit more complicated example:

Example 1.66. Use the precise definition of the limit to show that $\lim_{x\to 6}(3x-4) = 14$.

Solution. We give here a general outline of how to write the solution to the question "Use the precise definition of the limit to show that $\lim\limits_{x\to a} f(x) = L$."

> Let $\epsilon > 0$. Choose $\delta = \cdots$. Suppose that $0 < |x - a| < \delta$. Then
>
> $$\begin{aligned} |f(x) - L| &= \cdots \\ &= \cdots \\ &= \cdots \\ &< \cdots \\ &= \cdots \\ &= \epsilon. \end{aligned}$$
>
> Thus $|f(x) - L| < \epsilon$ as required.

This is what a general answer should look like, at least for a simple question like this one. (Probably it doesn't feel simple right now, but you'll get used to it!) Now let's start to make the answer our own by putting in the specifics. We are answering the question "Use the precise definition of the limit to show that $\lim_{x\to 6}(3x - 4) = 14$" and so we have $f(x) = 3x - 4$, $a = 6$ and $L = 14$. So now let's write out the next version of the answer.

> Let $\epsilon > 0$. Choose $\delta = \cdots$. Suppose that $0 < |x - 6| < \delta$. Then
>
> $$\begin{aligned} |(3x - 4) - 14| &= \cdots \\ &= \cdots \\ &= \cdots \\ &< \cdots \\ &= \cdots \\ &= \epsilon. \end{aligned}$$
>
> Thus $|(3x - 4) - 14| < \epsilon$ as required.

Now what we have to do is fill in all of the blanks, which are currently denoted by $\cdots$. We start, perhaps surprisingly, by filling in the blanks after $|(3x - 4) - 14| = \cdots$. Here is how it goes:

Let $\epsilon > 0$. Choose $\delta = \cdots$. Suppose that $0 < |x - 6| < \delta$. Then

$$\begin{aligned} |(3x-4) - 14| &= |3x - 18| \\ &= |3(x-6)| \\ &= 3|x-6| \\ &< 3\delta \\ &= \cdots \\ &= \epsilon. \end{aligned}$$

Thus $|(3x - 4) - 14| < \epsilon$ as required.

What did we just do? We simplified the expression $|(3x - 4) - 14|$ until it contained the expression $|x - 6|$, and then we used the assumption that $|x - 6| < \delta$. Now we have to fill in the rest. Ask yourself: what value of δ can we choose so that we can fill in the rest? We want to fill in the equations $3\delta = \cdots = \epsilon$. What value of δ gives us $3\delta = \epsilon$? It is $\delta = \epsilon/3$. So now we can complete the solution:

Let $\epsilon > 0$. Choose $\delta = \epsilon/3$. Suppose that $0 < |x - 6| < \delta$. Then

$$\begin{aligned} |(3x-4) - 14| &= |3x - 18| \\ &= |3(x-6)| \\ &= 3|x-6| \\ &< 3\delta \\ &= 3(\epsilon/3) \\ &= \epsilon. \end{aligned}$$

Thus $|(3x - 4) - 14| < \epsilon$ as required.

This is the final answer!

Remark 1.67. Here are some very important points for you to remember.

- If an exam or test question wants you to use the precise definition of the limit, it will say so. Otherwise you should stick to the methods we have used so far.
- When you are asked a question like this in an exam, you are not expected to write out multiple versions like we did in the last solution.

That was just my attempt to show you the steps. Instead, you will learn to write out your solution in one go by simply leaving things blank until you know what to write there.

- You will quickly learn that the hardest part of solving a question like this is to find or choose δ. However, your task when answering this question is to write out the whole solution.

- How you present your answer, and the order you present it, is important. Exam and test questions have marks for that. The simplest and best approach is to present your answer exactly as we have done here.

We would like to see some more examples of calculation of the limit using the precise definition. For this, we first go through properties of the absolute value functions and inequalities.

Properties of the absolute value function. Here are some useful properties of the absolute value function that should come in useful through the course, especially in this section.

- $|a \cdot b| = |a| \cdot |b|$.
- $|a \cdot b| = a \cdot |b|$ if $a \geqslant 0$.
- $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$ assuming that $b \neq 0$.
- $|a + b| \leqslant |a| + |b|$.
- $|x| < M$ if and only if $-M < x < M$.
- More generally, $|x - a| < M$ if and only if $a - M < x < a + M$.

Properties of inequalities. And here are some useful properties of inequalities.

- If $a < b < c$ then $a < c$.
- If $a \leqslant b < c$ then $a < c$.
- If $a < b \leqslant c$ then $a < c$.
- If $a \leqslant b \leqslant c$ then $a \leqslant c$.

- If $a < b$ and $c > 0$ then $ac < bc$.
- If $a < b$ and $c < 0$ then $bc < ac$. In particular $-b < -a$.
- If $a < b$ and the sign of a and b is the same (that is: they are both positive or both negative) , then $\frac{1}{b} < \frac{1}{a}$.
- If a and b are positive numbers, $a < b$, and n is a natural number, then $a^n < b^n$ and $\sqrt[n]{a} < \sqrt[n]{b}$.
- If a and b are negative numbers, $a < b$, and n is a natural number, then $a^n < b^n$ whenever n is odd, and $b^n < a^n$ whenever a is even (Why is there a difference between even and odd values of n?).

Example 1.68. Use the precise definition of the limit to show that $\lim_{x\to 3}(4x+5) = 17$.

Solution. Let's follow the same process as before. I am going to write out what my solution looks like at each step, so that you can see what I do. But when you write out your answer, you should simply leave the blanks as gaps to fill in, and you should only write out the solution once! I start with my skeleton answer.

> Let $\epsilon > 0$. Define $\delta =$ blank 1. Suppose that $0 < |x - 3| < \delta$. Thus
>
> $$\begin{aligned}|(4x + 5) - 17| &= \text{blank 3}\\ &\vdots\\ &= \epsilon.\end{aligned}$$
>
> Thus $|(4x + 5) - 17| < \epsilon$ as required.

Now I will fill in as much of blank 3 as I can.

> Let $\epsilon > 0$. Define $\delta =$ blank 1. Suppose that $0 < |x - 3| < \delta$.

Thus

$$\begin{aligned}|(4x+5)-17| &= |4x-12| \\ &= |4(x-3)| \\ &= 4|x-3| \\ &= \text{blank 3} \\ &\vdots \\ &= \epsilon.\end{aligned}$$

Thus $|(4x+5)-17| < \epsilon$ as required.

I see now that in blank 3 I have an expression involving only $|x-3|$. So I substitute in the inequality $|x-3| < \delta$.

Let $\epsilon > 0$. Define $\delta =$ blank 1. Suppose that $0 < |x-3| < \delta$. Thus

$$\begin{aligned}|(4x+5)-17| &= |4x-12| \\ &= |4(x-3)| \\ &= 4|x-3| \\ &< 4\delta \\ &= \text{blank 3} \\ &\vdots \\ &= \epsilon.\end{aligned}$$

Thus $|(4x+5)-17| < \epsilon$ as required.

Now I am in a position to choose my value of δ so that indeed $4\delta = \epsilon$, which is what we would need to complete blank 3. The correct choice is $\delta = \epsilon/4$. So now I can complete the solution:

Let $\epsilon > 0$. Define $\delta = \epsilon/4$. Suppose that $0 < |x-3| < \delta$. Thus

$$\begin{aligned}|(4x+5)-17| &= |4x-12| \\ &= |4(x-3)| \\ &= 4|x-3| \\ &< 4\delta \\ &= 4(\epsilon/4) \\ &= \epsilon.\end{aligned}$$

Thus $|(4x+5)-17| < \epsilon$ as required.

And that's it! **Please remember**, you should only write out the final answer, and you should do it by leaving the blanks as actual spaces on the page.

Example 1.69. Use the precise definition of the limit to show that $\lim_{x\to 0} x^2 = 0$.

Solution. This example is a lot like the last one, so I won't spell out all of the steps. Here's the answer.

Let $\epsilon > 0$. Define $\delta = \sqrt{\epsilon}$. Suppose that $0 < |x| < \delta$. Then

$$\begin{aligned} |x^2 - 0| &= |x|^2 \\ &< \delta^2 \\ &= (\sqrt{\epsilon})^2 \\ &= \epsilon \end{aligned}$$

so that $|x^2 - 0| < \epsilon$ as required.

Now we're going to move on to some harder examples. These will need a slightly more elaborate solution. So let's imagine that we've been given the following typical question.

Use the precise definition of the limit to show that $\lim\limits_{x\to a} f(x) = L$. (Here f, L and a will be specified.)

The typical solution to a question like this looks like the following.

Let $\epsilon > 0$. Define $\delta =$ blank 1. Suppose that $0 < |x-a| < \delta$.

Then blank 2.

Thus

$$\begin{aligned} |f(x) - L| &= \textit{blank 3} \\ &\vdots \\ &= \epsilon. \end{aligned}$$

So $|f(x) - L| < \epsilon$ as required.

Here blank 1 will be your chosen δ, usually a formula involving ϵ. And blank 2 will be a paragraph involving some preliminary computations. And finally, blank 3 will be a series of simplifications and substitutions like before.

How do you go about turning the typical question and typical answer into an actual solution? In particular, how do you fill in the blanks? Here is what you should do:

1. Write out the skeleton solution, filling in the values of $f(x)$, L and a, and leaving the blanks empty. Blank 2 may have to be an entire paragraph, and blank 3 may require a lot of lines.

2. Start on blank 3, where you work out and simplify $|f(x) - L|$, aiming for an expression involving $|x - a|$.

3. If the resulting expression involves x in some other way, you will need to fill in blank 2, which will consist of some preliminary working out, and then continue filling in blank 3. (We will see about this later on.)

4. Blank 3 will now express $|f(x)-L|$ in terms of $|x-a|$. Now substitute δ in place of $|x-a|$, making sure that you include the relevant inequality.

5. Look at blank 3 and make a good choice to fill in blank 1. It should be an expression of the form $\delta = \cdots$ where the right hand side is a formula involving ϵ.

6. Now complete blank 3, using your choice of δ.

7. You have finished!

Let's see several examples of this in action. Before we do, we need a new definition.

Definition 1.70 (The minimum). Let p and q be real numbers. Then $\min(p, q)$ denotes p or q, whichever is the smaller. So for example $\min(\pi, 3) = 3$ since $3 < \pi$, and $\min(\pi, 4) = \pi$ since $\pi < 4$. Note that

$$\min(p, q) \leqslant p$$

and

$$\min(p, q) \leqslant q.$$

(I won't give a proof of these inequalities. Think about the definition of $\min(p, q)$ and it will hopefully become clear.)

Example 1.71. Use the precise definition of the to show that $\lim_{x \to 2}(x^2+x+1) = 7$.

Solution. I'll write out my solution several times, adding more detail every time. However in lectures you will see me write everything out just once, and that is what you should do! I start with my skeleton answer.

> Let $\epsilon > 0$. Define $\delta =$ blank 1. Suppose that $0 < |x-2| < \delta$.
>
> Then blank 2.
>
> So
>
> $$|(x^2+x+1)-7| = \text{blank 3}$$
> $$\vdots$$
> $$= \epsilon.$$
>
> Thus $|(x^2+x+1)-7| < \epsilon$ as required.

Now I will fill in as much of blank 3 as I can.

> Let $\epsilon > 0$. Define $\delta =$ blank 1. Suppose that $0 < |x-2| < \delta$.
>
> Then blank 2.
>
> Thus
>
> $$\begin{aligned} |(x^2+x+1)-7| &= |x^2+x-6| \\ &= |(x+3)(x-2)| \\ &= |x+3| \cdot |x-2| \\ &= \text{blank 3} \\ &\ \vdots \\ &= \epsilon. \end{aligned}$$
>
> Thus $|(x^2+x+1)-7| < \epsilon$ as required.

Now here we see a problem: we've obtained an expression involving $|x-2|$ as we want, but there's an annoying factor $|x+3|$ that we can't control. Now we will complete blank 2 and also edit blank 1.

Let $\epsilon > 0$. Define $\delta = \min(1, \text{blank 1})$. Suppose that $0 < |x-2| < \delta$.

Then since $\delta \leqslant 1$ and $|x-2| < \delta$, we have $|x-2| < 1$. That means $-1 < (x-2) < 1$, so by adding 5 to every term we see that $4 < (x+3) < 6$, so that $|x+3| < 6$.

Thus

$$\begin{aligned}
|(x^2+x+1)-7| &= |x^2+x-6| \\
&= |(x+3)(x-2)| \\
&= |x+3| \cdot |x-2| \\
&< 6|x-2| \\
&= \text{blank 3} \\
&\vdots \\
&= \epsilon.
\end{aligned}$$

Thus $|(x^2+x+1)-7| < \epsilon$ as required.

Now we can continue much as before. Here is the final answer.

Let $\epsilon > 0$. Define $\delta = \min(1, \epsilon/6)$. Suppose that $0 < |x-2| < \delta$.

Then since $\delta \leqslant 1$ and $|x-2| < \delta$, we have $|x-2| < 1$. That means $-1 < (x-2) < 1$, so by by adding 5 to every term we see that $4 < (x+3) < 6$, so that $|x+3| < 6$.

Thus

$$\begin{aligned}
|(x^2+x+1)-7| &= |x^2+x-2| \\
&= |(x+3)(x-2)| \\
&= |x+3| \cdot |x-2| \\
&< 6|x-2| \\
&< 6\delta \\
&\leqslant 6(\epsilon/6) \\
&= \epsilon.
\end{aligned}$$

Thus $|(x^2+x+1)-7| < \epsilon$ as required.

At this stage you should study our proof carefully and ask why each step follows from the next. Here are some pointers.

- since $\delta \leqslant 1$: This follows because by $\delta = \min(1, \epsilon/6)$, and we always have $\min(p, q) \leqslant p$ and $\min(p, q) \leqslant q$.

- we have $|x-2| < 1$ so $-1 < (x-2) < 1$: This is because saying $|a| < M$ is the same as saying that $-M < a < M$. (This was one of our useful facts about the absolute value function.)

- $4 < (x+3) < 6$, so $|x+2| < 6$: This follows because if $4 < (x+2) < 6$, then in particular $-6 < (x+2) < 6$, so that $|x+2| < 6$. (Again using one of our useful facts about the absolute value function.)

- $6\delta \leqslant 6(\epsilon/6)$: Remember that $\delta = \min(1, \epsilon/6)$, so that $\delta \leqslant \epsilon/6$, so that $6\delta \leqslant 6(\epsilon/6)$.

Also, we made a lot of choices when we wrote out the proof. Why did we make these choices, and what choice did we have?

- Why did we choose $\delta = \min(1, \epsilon/6)$? Defining δ this way means that we get two facts, namely $\delta \leqslant 1$ and $\delta \leqslant \epsilon/6$. We used these two different facts in two different places — try to see where.

- Why did we choose the $\epsilon/6$ in $\delta = \min(1, \epsilon/6)$? This was chosen to make the final "tower" of working out correct, where we replace an expression involving δ with one involving ϵ. You choose $\epsilon/6$ to make sure that you end that calculation with an ϵ.

- Why did we choose the 1 in $\delta = \min(1, \epsilon/6)$? Well, the 1 told us that $\delta \leqslant 1$, which we used in the second paragraph to put a bound on $|x+3|$. In fact, it didn't matter what number we chose here. It could have been any positive number. To see why, replace the 1 with a 9 and change the rest of the proof accordingly. What happens?

- Why did we add 5 to every term of $-1 < (x-2) < 1$? In this paragraph we were trying to understand the quantity $|x+3|$. What we already knew was that $-1 < (x-2) < 1$, and to make this tell us something about $(x+3)$ we added 5, which turns the $(x-2)$ into the $(x+3)$.

Example 1.72. Use the precise definition of the limit to show that $\lim_{x\to 2}\left(\frac{x+3}{x-1}\right) = 5$.

Solution. Again, I will go through the example in many steps, writing each one out in full so that you can see the changes. In the lectures I will just do it in one go, leaving blanks until I know what goes where, and that is what you should do when working out your own answers. I will start by writing out the typical solution, filling in the values of a, $f(x)$ and L, and also working out as much of the final computation as I can. Again, my δ will be a minimum, so I'll put that in as well.

Let $\epsilon > 0$. Define $\delta = \min(-, -)$. Suppose that $0 < |x-2| < \delta$.

Then blank 2.

Thus

$$
\begin{aligned}
\left|\left(\frac{x+3}{x-1}\right) - 5\right| &= \left|\frac{(x+3)-5(x-1)}{x-1}\right| \\
&= \left|\frac{-4x+8}{x-1}\right| \\
&= \left|\frac{-4(x-2)}{x-1}\right| \\
&= 4\cdot\left|\frac{x-2}{x-1}\right| \\
&= 4\cdot\frac{|x-2|}{|x-1|} \\
&= 4\cdot\frac{1}{|x-1|}\cdot|x-2| \\
&= \text{blank 3} \\
&\ \ \vdots \\
&= \epsilon.
\end{aligned}
$$

So $\left|\left(\frac{x+3}{x-1}\right) - 5\right| < \epsilon$ as required.

The situation looks good, because we expressed $\left|\left(\frac{x+3}{x-1}\right)-5\right|$ in terms of $|x-2|$, but now there is a factor of $\frac{1}{|x-1|}$ making things complicated. Like in our previous example, we need to make sure that this factor of $\frac{1}{|x-1|}$ is not too large. That means that we need to make sure that $|x-1|$ itself is not too *small*. Here is how we do it.

> Let $\epsilon > 0$. Define $\delta = \min(1/2, -)$. Suppose that $0 < |x-2| < \delta$.
>
> Then since $\delta \leqslant 1/2$, we have $-1/2 < (x-2) < 1/2$. Adding 1 to both sides gives us $1/2 < (x-1) < 3/2$. From this we can see that $|x-1| > 1/2$. This rearranges to tell us that $\frac{1}{|x-1|} < 2$.
>
> Thus
>
> $$\begin{aligned} \left|\left(\frac{x+3}{x-1}\right)-5\right| &= \left|\frac{(x+3)-5(x-1)}{x-1}\right| \\ &= \left|\frac{-4x+8}{x-1}\right| \\ &= \left|\frac{-4(x-2)}{x-1}\right| \\ &= 4\cdot\left|\frac{x-2}{x-1}\right| \\ &= 4\cdot\frac{|x-2|}{|x-1|} \\ &= 4\cdot\frac{1}{|x-1|}\cdot|x-2| \\ &< 4\cdot 2\cdot|x-2| \\ &= 8\cdot|x-2| \\ &< 8\cdot\delta \\ &= \text{blank 3} \\ &\vdots \\ &= \epsilon. \end{aligned}$$
>
> So $\left|\left(\frac{x+3}{x-1}\right)-5\right| < \epsilon$ as required.

Now all we have to do is fill in the second part of the minimum and complete our proof.

> Let $\epsilon > 0$. Define $\delta = \min(1/2, \epsilon/8)$. Suppose that $0 < |x-2| < \delta$.
>
> Then since $\delta \leqslant 1/2$, we have $-1/2 < (x-2) < 1/2$. Adding 1 to both sides gives us $1/2 < (x-1) < 3/2$. From this we can see that $|x-1| > 1/2$. This rearranges to tell us that $\frac{1}{|x-1|} < 2$.
>
> Thus
>
> $$\begin{aligned} \left|\left(\frac{x+3}{x-1}\right) - 5\right| &= \left|\frac{(x+3)-5(x-1)}{x-1}\right| \\ &= \left|\frac{-4x+8}{x-1}\right| \\ &= \left|\frac{-4(x-2)}{x-1}\right| \\ &= 4 \cdot \left|\frac{x-2}{x-1}\right| \\ &= 4 \cdot \frac{|x-2|}{|x-1|} \\ &= 4 \cdot \frac{1}{|x-1|} \cdot |x-2| \\ &< 4 \cdot 2 \cdot |x-2| \\ &= 8 \cdot |x-2| \\ &< 8 \cdot \delta \\ &\leqslant 8 \cdot (\epsilon/8) \\ &= \epsilon. \end{aligned}$$
>
> So $\left|\left(\frac{x+3}{x-1}\right) - 5\right| < \epsilon$ as required.

Like in the previous example, we made a lot of choices, and you might ask how and why we made those choices. The only point now that is different from before is this one:

- Why did we choose $1/2$ in $\delta = \min(1/2, \epsilon/8)$? The $1/2$ was used in the second paragraph in order to show that $1/2 < (x-1) < 3/2$. The key thing here was to make sure that in the final inequality the two 'ends', namely 1/2 and 3/2, were both positive (or at least have the same sign). For this purpose, we needed 1/2 to be small enough. Check this: Replace 1/2 with 1/4 and see what happens to the proof. Then replace it with 1 and see what happens.

Precise definition of infinite limits

We gave the precise definition of $\lim_{x\to a} f(x) = L$. We will also give now the precise definition of $\lim_{x\to a} f(x) = \infty$ and $\lim_{x\to\infty} f(x) = L$. We will the give a few examples.

Definition 1.73. Let f be a function defined close to a point a. We say that $\lim_{x\to a} f(x) = \infty$ if the following condition holds: For every natural number N there exists a $\delta > 0$ such that $f(x) > N$ whenever $|x - a| < \delta$.

Let us compare again this definition with the imprecise definition we gave before. The imprecise definition was the following:
We say that $lim_{x\to a} f(x) = \infty$ if we can make $f(x)$ as large and positive as we wish by making x sufficiently close to, but not equal to, a.
Now let us write the above definition with colors:
We say that $\lim_{x\to a} f(x) = \infty$ if the following condition holds: For every natural number N there exists a $\delta > 0$ such that $f(x) > N$ whenever $0 < |x - a| < \delta$. The phrase "we can make $f(x)$ as large and positive as we wish" translates to saying that $f(x) > N$. The "by making x sufficiently close to.." part translates to $|x - a| < \delta$. Let us see an example:

Example 1.74. Prove that $\lim_{x\to 0} \frac{1}{x^2} = \infty$.

Solution. In the proof here we will follow the pattern of the previous proofs. The only difference is that now we need to show $f(x) > N$ instead of $|f(x) - L| < \epsilon$. We use again the blanks as before. In first step we write:
Let $N > 0$ be a natural number. Choose $\delta = \cdots$. Assume that $0 < |x| < \delta$. We have :

$$\begin{aligned} f(x) = \frac{1}{x^2} &= \cdots \\ &= \\ &\vdots \\ &> N. \end{aligned}$$

In our case, we know that $|x| < \delta$. By the inequality rules we have we know that:
$x^2 = |x|^2 < \delta^2$ and therefore
$f(x) = \frac{1}{x^2} > \frac{1}{\delta^2}$. We want to get $f(x) > N$. This will be true if $N = \frac{1}{\delta^2}$, or, in other words, if $\delta = \frac{1}{\sqrt{N}}$. So the proof of $\lim_{x\to 0} f(x) = \infty$ is this:
Let $N > 0$ be a natural number. Choose $\delta = \frac{1}{\sqrt{N}}$. Assume that $0 < |x| < \delta$. Then we have:

$$f(x) = \frac{1}{x^2} > \frac{1}{\delta^2} = N$$

and we are done.

In a similar way, we can also give precise definition for all the other limits, like the one sided limits, limits such as $\lim_{x\to\infty} f(x)$ and so on. The importance of this part in the course was for you to see that there is a precise definition for the intuitive notion of a limit.

The example above is the last example of computing a limit with the precise definition that we will do in class. Please study the examples in detail, and follow the hints, to try to understand how they work and how you yourself could answer them. Now we will see how the precise definition of the limit can be used to prove the limit laws, in this case the sum law. All of the limit laws we have seen so far can be proved in this way, though some proofs are harder than others.

Theorem 1.75 (The sum law for limits)**.** *Let f and g be functions defined close to a, but not necessarily at a itself. Suppose that $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ exist. Then $\lim_{x\to a}[f(x) + g(x)]$ exists and*

$$\lim_{x\to a}[f(x) + g(x)] = \lim_{x\to a} f(x) + \lim_{x\to a} g(x).$$

Proof. Suppose that $\lim_{x\to a} f(x) = L$ and that $\lim_{x\to a} g(x) = M$. Then we must show that $\lim_{x\to a}[f(x) + g(x)] = L + M$.

Let $\epsilon > 0$. Since $\lim_{x\to a} f(x) = L$, there is $\delta_1 > 0$ such that $|f(x) - L| < \epsilon/2$ whenever $0 < |x - a| < \delta_1$. And since $\lim_{x\to a} g(x) = M$, there is $\delta_2 > 0$ such that $|g(x) - M| < \epsilon/2$ whenever $0 < |x - a| < \delta_2$. Define $\delta = \min(\delta_1, \delta_2)$ and suppose that $0 < |x - a| < \delta$. Then since $\delta \leqslant \delta_1$, we have $0 < |x-a| < \delta_1$, and so $|f(x) - L| < \epsilon/2$. And since $\delta \leqslant \delta_2$, we have $0 < |x - a| < \delta_2$, and so

$|g(x) - M| < \epsilon/2$. So finally

$$\begin{aligned} |(f(x)+g(x)) - (L+M)| &= |(f(x)-L) + (g(x)-M)| \\ &\leqslant |f(x)-L| + |g(x)-M| \\ &< \epsilon/2 + \epsilon/2 \\ &= \epsilon \end{aligned}$$

so that $|(f(x)+g(x)) - (L+M)| < \epsilon$ as required. □

1.6 Continuity

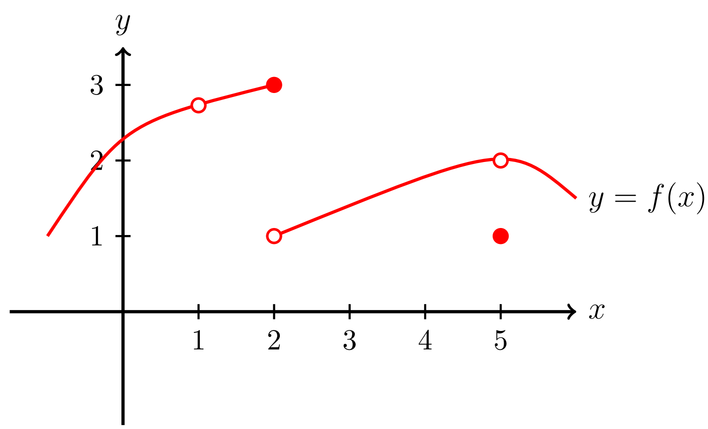
Now we will study a certain class of 'nice' functions that are called 'continuous'. These are the functions for which limits can be worked out by direct substitution. Intuitively, they are the functions whose graph can be drawn without taking the pen from the page.

Definition 1.76 (Continuity of a function at a point). A function f is called *continuous at a number a* if $\lim_{x\to a} f(x) = f(a)$. In practice, for f to be continuous requires the following three things.

1. $f(a)$ is defined, or in other words, a is in the domain of f.
2. $\lim_{x\to a} f(x)$ exists and is finite.
3. $\lim_{x\to a} f(x) = f(a)$.

If f is not continuous at a, then we say that f is *discontinuous at a*, or that it has a *discontinuity at a*.

Example 1.77. Let f be the function with the following graph.



At which of the following numbers a is f discontinuous? In each case, say which of properties 1, 2, and 3 fails.

- $a = 1$
- $a = 2$
- $a = 5$

Solution.

- At $a = 1$ the function is discontinuous because $f(1)$ is not defined, so that property 1 fails.
- At $a = 2$ the function is discontinuous because, although $f(2)$ is defined, $\lim_{x\to 2} f(x)$ does not exist, so that property 2 fails.
- At $a = 5$ the function is discontinuous because, although $f(5)$ is defined, and although $\lim_{x\to 5} f(x)$ exists, the two are not equal. Indeed, $f(5) = 1$ but $\lim_{x\to 5} f(x) = 2$. So property 3 fails.

Example 1.78. At which numbers are the functions f, g and h defined as follows discontinuous?

- $f(x) = \dfrac{x^2 - x - 2}{x - 2}$
- $g(x) = \begin{cases} 1/x^2 & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$
- $h(x) = \begin{cases} \frac{x^2-x-2}{x-2} & \text{if } x \neq 2 \\ 1 & \text{if } x = 2 \end{cases}$
- $$H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geqslant 0 \end{cases}$$

Solution.

- f is not defined at 2, and so 2 is a discontinuity. But f is continuous everywhere else.
- g is discontinuous at 0 because, although $g(0)$ is defined, $\lim_{x\to 0} g(x)$ does not exist as a finite limit. But g is continuous everywhere else.

- h is continuous at every number except possibly 2. At $a = 2$, we see that $h(2)$ is defined and equal to 1, and that

$$\begin{aligned}\lim_{x\to 2} h(x) &= \lim_{x\to 2} \frac{x^2 - x - 2}{x - 2} \\ &= \lim_{x\to 2} \frac{(x-2)(x+1)}{x-2} \\ &= \lim_{x\to 2}(x+1) \\ &= 3\end{aligned}$$

  so that $\lim_{x\to 2} h(x)$ does exist, but is not equal to $h(2)$, so that $x = 2$ is in fact a discontinuity.

- For $a < 0$ we have $H(a) = 0 = \lim_{x\to a} H(x)$ and H is continuous there. For $a > 0$ we have $H(a) = 1 = \lim_{x\to a} H(x)$ and H is also continuous there. We have seen that the one sided limits of $H(x)$ at $x = 0$ do not agree, and therefore $\lim_{x\to 0} H(x)$ does not exist, and $H(x)$ is therefore not continuous there.

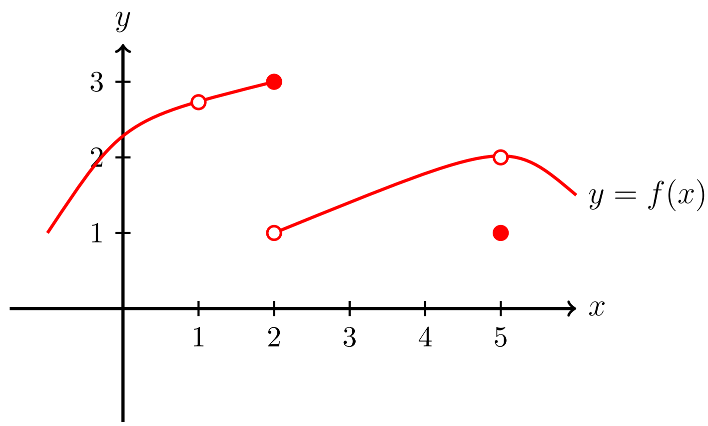
Definition 1.79 (Continuity of a function, at a point, from one side). We say that a function f is *continuous from the left* at a number a if $\lim_{x\to a^-} f(x) = f(a)$. In other words, for f to be continuous from the left at a we require the following three properties:

1. $f(a)$ is defined.
2. $\lim\limits_{x\to a^-} f(x)$ exists and is finite.
3. $\lim\limits_{x\to a^-} f(x) = f(a)$.

And we say that a function f is *continuous from the right* at a number a if $\lim_{x\to a^+} f(x) = f(a)$. In other words, for f to be continuous from the right at a we require the following three properties:

1. $f(a)$ is defined.
2. $\lim\limits_{x\to a^+} f(x)$ exists and is finite.
3. $\lim\limits_{x\to a^+} f(x) = f(a)$.

Example 1.80. Let f again be the function with the following graph.



Then

- f is not continuous from the right or the left at $a = 1$ since $f(1)$ is not defined.
- f is continuous from the left at $a = 2$ because $\lim_{x\to 2^-} f(x) = 3 = f(2)$.
- f is not continuous from the right at $a = 2$ because $\lim_{x\to 2^+} f(x) = 1 \neq 3 = f(2)$.
- f is not continuous from the right or the left at $a = 5$ because $\lim_{x\to 5^-} f(x) = \lim_{x\to 5^+} f(x) = 2$ but $f(5) = 1$.

Definition 1.81 (Continuity of a function on an interval). A function f is *continuous on an interval* I if it is continuous at a for all $a \in I$. If a is an endpoint of the interval, then we only require continuity from the left (if a is the right-hand end of the interval) or from the right (if a is the left-hand end of the interval).

Example 1.82. Show that the function f defined by $f(x) = 1 - \sqrt{1-x^2}$ is continuous on $[-1, 1]$.

Solution. If a lies in the range $-1 < a < 1$, then by the limit laws

$$\begin{aligned}
\lim_{x\to a}\left(1-\sqrt{1-x^2}\right) &= 1 - \lim_{x\to a}\sqrt{1-x^2} \\
&= 1 - \sqrt{\lim_{x\to a}(1-x^2)} \\
&= 1 - \sqrt{1-a^2} \\
&= f(a)
\end{aligned}$$

so that f is continuous at a as required. Here, the use of the square root law required the fact that $1 - x^2 > 0$ for x close to, but not equal to, a. Similar computations with left and right limits show continuity from the right and left at -1 and 1 respectively. So f is continuous on $[-1, 1]$.

Now we will see how to construct new continuous functions from existing ones.

Theorem 1.83. *Suppose that f and g are continuous at a, and that c is a constant. Then the following functions are also continuous at a:*

$f+g$ $\quad$ $f-g$ $\quad$ fg $\quad$ cf $\quad$ f/g, *as long as* $g(a) \neq 0$.

Similarly, if f and g are continuous on an interval I, then the following are also continuous on I.

$f+g$ $\quad$ $f-g$ $\quad$ fg $\quad$ cf $\quad$ f/g, *as long as* $g(x) \neq 0$ *for all* $x \in I$.

Proof for $f+g$. Let us show that if f and g are continuous at a, then so is $f+g$. Indeed, we have the following,

$$\begin{aligned}\lim_{x\to a}(f+g)(x) &= \lim_{x\to a}[f(x)+g(x)] \\ &= \lim_{x\to a} f(x) + \lim_{x\to a} g(x) \\ &= f(a)+g(a) \\ &= (f+g)(a)\end{aligned}$$

so that $f+g$ is continuous at a. □

Using this theorem, we are able to construct a large supply of continuous functions.

Theorem 1.84. *Any polynomial function is continuous on $\mathbb{R} = (-\infty, \infty)$. Any rational function is continuous at any point in its domain, and so in particular is continuous on any interval in its domain.*

Proof. The function f_0 defined by $f_0(x) = 1$ is continuous on $\mathbb{R}$, as is the function f_1 defined by $f_1(x) = x$. Then by the last theorem the function $f_1 f_1$, which is defined by $(f_1 f_1)(x) = x^2$ is also continuous on $\mathbb{R}$. Similarly, it follows that all of the functions f_n defined by $f_n(x) = x^n$, are continuous on $\mathbb{R}$. But then an arbitrary polynomial p with formula

$$p(x) = c_k x^k + c_{k-1} x^{k-1} + \cdots + c_1 x + c_0$$

is nothing other than

$$p = c_k f_k + \cdots + c_1 f_1 + c_0 f_0,$$

which is continuous on $(-\infty, \infty)$, again using the previous theorem. Finally, an arbitrary rational function r defined by $r(x) = p(x)/q(x)$ where p and q are polynomials is nothing other than the quotient function p/q, which again by the previous theorem is continuous at all a for which $q(a) \neq 0$, or in other words, at any a that lies in its domain. □

Example 1.85. On which intervals is the function p defined by $p(x) = \frac{x^3+2x^2-1}{5-3x}$ continuous?

Solution. p is a rational function, and so is continuous on any interval inside its domain. The domain of p contains every number for which $5 - 3x \neq 0$, i.e. for which $x \neq 5/3$, hence is $(-\infty, 5/3) \cup (5/3, \infty)$. So the largest possible intervals on which p is continuous are $(-\infty, 5/3)$ and $(5/3, \infty)$, and it is also continuous on any interval inside these.

In fact, we can do better than the previous theorem about continuity of polynomials and rational functions.

Theorem 1.86. *The following classes of functions are continuous at every point of their domains, and on every interval within their domains.*

- *Polynomials*
- *Rational functions*
- *Root functions*
- *Trigonometric functions (namely sin, cos, tan, sec, cosec, cot)*
- *Exponential functions*
- *Logarithmic functions*

(We will learn more about the last three classes of functions later on in the course.)

Example 1.87. On which intervals is the function continuous?

1. $f(x) = x^{100} - 2x^{37} + 75$

2. $g(x) = \dfrac{x^2 + 2x + 17}{x^2 - 1}$

3. $h(x) = \sqrt{x} + \dfrac{x+1}{x-1} - \dfrac{x+1}{x^2+1}$

Solution. 1. f is polynomial, so is continuous on $(-\infty, \infty)$.

2. g is rational, so is continuous at every point on its domain, which is $(-\infty, -1) \cup (-1, 1) \cup (1, \infty)$, so is continuous on the intervals $(-\infty, -1)$, $(-1, 1)$ and $(1, \infty)$.

3. h is a sum of:

   - A root function with domain $[0, \infty)$.
   - A rational function with domain $(-\infty, 1) \cup (1, \infty)$.
   - A rational function with domain $(-\infty, \infty)$.

   So h is continuous at any point a that lies in all three domains, and is continuous on the intervals in all three domains, i.e. $[0, 1)$ and $(1, \infty)$.

Here we have specified the largest possible intervals on which the functions are continuous. It follows that they are also continuous on any smaller intervals. For example, g is continuous on $(-10, 1)$ or $(-10, 1.5]$.

Now we will see one last way of constructing new continuous functions from old ones.

Theorem 1.88. *1. Suppose that* $\lim_{x\to a} g(x) = b$ *and that* $\lim_{x\to b} f(x) = c$.
Then $\lim_{x\to a}(f \circ g)(x)$ *exists and is equal to* c.

2. Suppose that $\lim\limits_{x\to a} g(x) = b$ *and that* f *is continuous at* b. *Then*

$$\lim_{x\to a}(f \circ g)(x) = f\left(\lim_{x\to a} g(x)\right).$$

3. *If g is continuous at a and f is continuous at $g(a)$, then $f \circ g$ is continuous at a.*

Example 1.89. Let F be the function defined by $F(x) = \dfrac{1}{\sqrt{x} - 4}$. At what points is F continuous?

Solution. $F = p \circ q$, where p and q are defined by $q(x) = \sqrt{x}$ and $p(x) = \frac{1}{x-4}$. These are continuous on their domains, and so the same is true of F. Now

$$\mathrm{dom}(F) = \{x \mid x \geq 0 \text{ and } \sqrt{x} - 4 \neq 0\},$$

and

$$\sqrt{x} - 4 \neq 0 \iff \sqrt{x} \neq 4 \iff x \neq 16$$

So F is continuous at all non-negative numbers a except for $a = 16$.

Example 1.90. Let f be the function defined as follows.

$$f(x) = \begin{cases} x & \text{if } x \leqslant 1 \\ 1/x & \text{if } 1 < x < 3 \\ \sqrt{x-3} & \text{if } 3 \leqslant x \end{cases}$$

At which numbers is f discontinuous?

Solution. First of all we show that f is continuous away from the 'joins' in the definition of f. More precisely, we show that f is continuous on each of the intervals $(-\infty, 1)$, $(1, 3)$ and $(3, \infty)$.

- On the interval $(-\infty, 1)$, f is given by $f(x) = x$. This is a polynomial, so f is continuous at every point of $(-\infty, 1)$.
- On the interval $(1, 3)$, f is given by $f(x) = 1/x$. This is a rational function and $(1, 3)$ lies in its domain, so f is continuous at every point of $(1, 3)$.
- On the interval $(3, \infty)$, f is given by $f(x) = \sqrt{x-3}$, so $f = p \circ q$ where p and q are defined by $q(x) = x - 3$ and $q(x) = \sqrt{x}$. Now if $x \in (3, \infty)$ then q is continuous at x, and p is continuous at $q(x) > 0$, so that f is continuous at x.

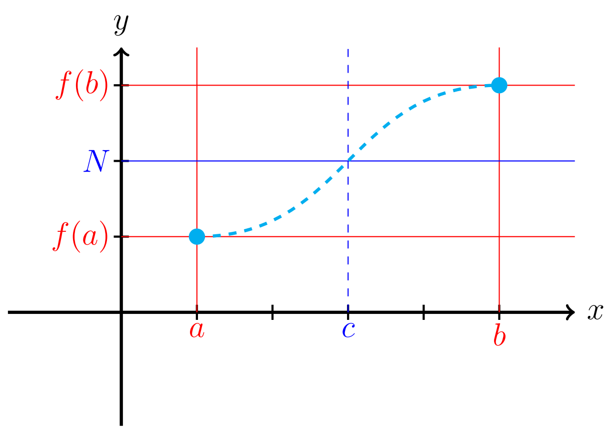
It remains to see whether f is continuous at $a = 1$ and $a = 3$. In these cases we understand $\lim_{x \to a} f(x)$ by looking at the one-sided limits, which we can compute using the relevant 'piece' of the definition of f.

- $\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} x = 1$, and $\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+}(1/x) = 1$, so $\lim_{x\to 1} f(x) = 1$. But $f(1) = 1$ as well, so that f is continuous at 1.
- $\lim_{x\to 3^-} f(x) = \lim_{x\to 3^-}(1/x) = \frac{1}{3}$, and $\lim_{x\to 3^+} f(x) = \lim_{x\to 3^+} \sqrt{x-3} = 0$, so f is not continuous at 3.

Now we will see an important result about continuous functions. It is the result that says that the graph of a continuous function does not contain any vertical 'gaps' or 'jumps'.

Theorem 1.91 (The Intermediate Value Theorem)**.** *Suppose that f is continuous on the closed interval $[a, b]$, that $f(a) \neq f(b)$, and that N is a number lying between $f(a)$ and $f(b)$, but not equal to either $f(a)$ or $f(b)$. Then there is $c \in (a, b)$ such that $f(c) = N$.*

To try to explain this theorem, we depict the information on the following graph



Here the only thing we know about the function f is that it is continuous, and that it passes through $(a, f(a))$ and $(b, f(b))$. What the theorem guarantees is that the graph actually crosses the line $y = N$ at (c, N), as opposed to somehow magically jumping over the line $y = N$. The theorem is not entirely straightforward, and we will not give the proof.

Example 1.92. Show that the equation $4x^3 - 6x^2 + 3x - 2 = 0$ has a solution between 1 and 2.

Solution. Remember that a *solution* of an equation in a variable x is a choice of x that makes that equation true. So here, a solution is a choice of x for which $4x^3 - 6x^2 + 3x - 2$ really equals 0.

We will apply the intermediate value theorem with the function f defined by $f(x) = 4x^3 - 6x^2 + 3x - 2$, with $a = 1$ and $b = 2$, and with $N = 0$. Let's check that the conditions are satisfied. First, f is a polynomial, and so is continuous on $(-\infty, \infty)$, and in particular is continuous on $[a, b] = [1, 2]$. Next, $f(a) = f(1) = 4-6+3-2 = -1$ and $f(b) = f(2) = 32-24+6-2 = 12$, so that $N = 0$ does lie between $f(a)$ and $f(b)$, but is not equal to either of them. So the intermediate value theorem applies and shows that there is $c \in (a, b) = (1, 2)$ for which $f(c) = N$, i.e. $f(c) = 0$. This c is the required solution of the original equation.

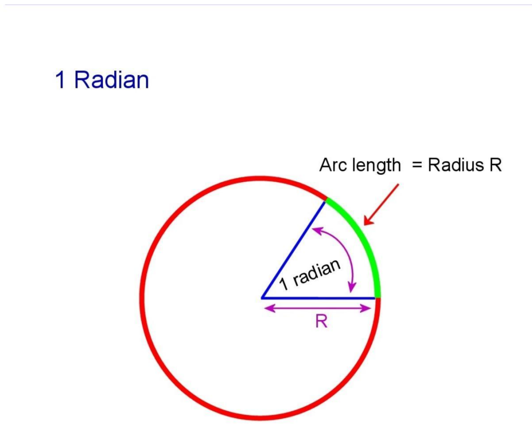
Example 1.93. Show that the equation $\sin^2(x) = \frac{1}{5}$ has a solution between 0 and $\frac{\pi}{4}$.

Solution. Again, a solution to this equation is a choice of x that makes the equation true. Define f by $f(x) = \sin^2(x)$. Now sin is continuous on its domain $(-\infty, \infty)$, and $f(x) = \sin^2(x) = (\sin(x))^2$, so that f is also continuous on $(-\infty, \infty)$, and in particular f is continuous on $[0, \pi/4]$. Now $f(0) = 0$ and $f(\pi/4) = 1/2$. Let $N = 1/5$, then N lies between $f(0)$ and $f(\pi/4)$, and is not equal to either of them. So the intermediate value theorem applies and shows that there is $c \in (0, \pi/4)$ such that $f(c) = 1/5$, or in other words that $\sin^2(c) = 1/5$. Thus c is the required solution to the original equation.

Continuity of the trigonometric functions

We will give now a concrete proof of the fact that the trigonometric functions $\sin(x)$ and $\cos(x)$ are continuous. Since $\tan(x) = \frac{\sin(x)}{\cos(x)}$ it will also be continuous in its domain of definition. A similar statement holds for $\cot(x)$.

One important thing to mention here: when we write $\sin(x)$, the angle for the sine function is measured in *radians*, and not in *degrees*. An angle of x radians will have $\frac{180x}{\pi}$ degrees, and an angle of r degrees will have $\frac{r\pi}{180}$ radians. The radians have the following interpretation: In a circle of radius 1, an angle of x radians corresponds to an arc of length x. See the picture below. Use the picture also to see why $\sin(x) < x$ when $0 < x < \pi/2$. This inequality depends on measuring in radians and not in degrees.



We prove the continuity of $\sin(x)$ and $\cos(x)$ in a few steps:

1. $\sin(x)$ is continuous at $x = 0$. For $0 < x < \pi/2$ it holds that $0 < \sin(x) < x$. By the squeeze Theorem it follows that $\lim_{x\to 0^+} \sin(x) = 0 = \sin(0)$. In a similar way we can also show that $\lim_{x\to 0^-} \sin(x) = 0$, or simply by using the fact that $\sin(x)$ is an odd function.

2. $\cos(x)$ is continuous at $x = 0$. This is true because for $-\pi/2 < x < \pi/2$ it holds that $\cos(x) = \sqrt{1 - \sin(x)^2}$, and we already know that $\sin(x)$ is continuous at 0.

3. $\sin(x)$ is continuous everywhere. Let a be any number. It holds that $\lim_{x\to a} \sin(x) = \lim_{x\to 0} \sin(a + x)$ (explain why!). But $\sin(a + x) = \sin(a)\cos(x) + \sin(x)\cos(a)$. By using now the continuity of $\sin(x)$ and $\cos(x)$ at zero, and using the fact that $\sin(0) = 0$ and $\cos(0) = 1$, we get that this limit is $\sin(a)$. This implies that sin is continuous on the entire $\mathbb{R}$.

Chapter 2

Differentiation

2.1 Derivatives and rates of change

Tangents

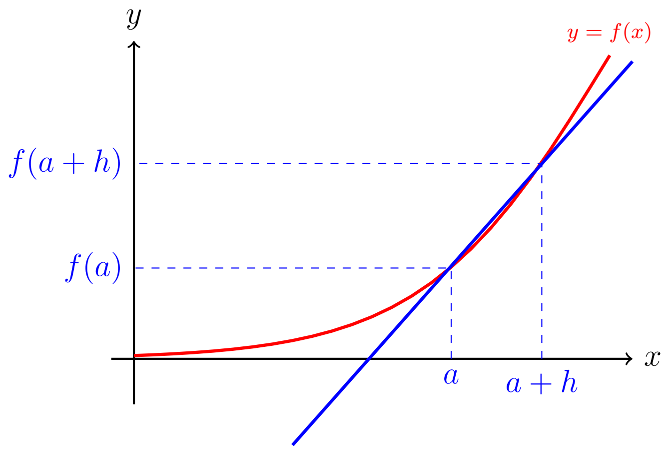
Definition 2.1. The *tangent line* to the curve $y = f(x)$ at the point $(a, f(a))$ is the line through $(a, f(a))$ with gradient

$$m = \lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

if this limit exists. Notice that this limit is the same as

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

This is pictured in the graph below, which shows that $\frac{f(a+h)-f(a)}{h}$ is the gradient of the line that crosses $y = f(x)$ at $(a, f(a))$ and $(a + h, f(a + h))$.



Example 2.2. Find the equation of the tangent line to the curve $y = x^2$ through the point $(1, 1)$.

Solution. Let f be the function defined by $f(x) = x^2$, so that our curve is given by $y = f(x)$. Then the gradient of the line is

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

And so the equation of the tangent line is $(y - 1) = 2(x - 1)$, or in other words $y = 2x - 1$.

Velocities

Suppose that an object is moving along a line according to the equation $s = f(t)$ where s is the *displacement*, i.e. position along the line, t is tthe time, and $f(t)$ is the *position function*. The *average velocity* of the object between times a and $a+h$ is then

$$\text{average velocity} = \frac{\text{distance travelled}}{\text{time taken}} = \frac{f(a+h)-f(a)}{(a+h)-a} = \frac{f(a+h)-f(a)}{h}.$$

And the *instantaneous velocity* at time a is

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

This is the gradient of the graph of $y=f(x)$ at $(a, f(a))$.

Derivatives

Definition 2.3 (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*.

Example 2.4. Let f be the function defined by $f(x) = x^2 - 8x + 9$. Using the definition of the derivative, find the derivative of f at a.

Solution. We start the question by simply writing out the definition of the derivative.

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

Next, we write out $f(a+h)$ and $f(a)$ using the definition of f. Remember, when you write down $f(a)$, you do it by taking the definition of $f(x)$ and replacing every x with a. And when you write down $f(a+h)$, do it by replacing every x with $(a+h)$ — remember to include the brackets, as it will save you from making a lot of mistakes.

$$\begin{aligned} f'(a) &= \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \\ &= \lim_{h\to 0} \frac{[(a+h)^2 - 8(a+h) + 9] - [a^2 - 8a + 9]}{h} \end{aligned}$$

And now we expand, simplify, and try to work out the limit.

$$
\begin{aligned}
f'(a) &= \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \\
&= \lim_{h\to 0} \frac{[(a+h)^2 - 8(a+h)+9] - [a^2-8a+9]}{h} \\
&= \lim_{h\to 0} \frac{[a^2+2ah+h^2-8a-8h+9]-[a^2-8a+9]}{h} \\
&= \lim_{h\to 0} \frac{a^2+2ah+h^2-8a-8h+9-a^2+8a-9}{h} \\
&= \lim_{h\to 0} \frac{2ah+h^2-8h}{h} \\
&= \lim_{h\to 0}(2a+h-8) \\
&= 2a-8.
\end{aligned}
$$

So $f'(a) = 2a - 8$.

Example 2.5. Let f be the function defined by $f(x) = 2x^2 + x - 3$. Find $f'(2)$.

Solution. We begin, as always, by writing out the definition of $f'(2)$. This is of course just the same as the definition of $f'(a)$, but with 2 substituted in place of a.

$$
\begin{aligned}
f'(2) &= \lim_{h\to 0} \frac{f(2+h)-f(2)}{h} \\
&= \lim_{h\to 0} \frac{[2(2+h)^2 + (2+h) - 3] - [2\cdot 2^2 + 2 - 3]}{h} \\
&= \lim_{h\to 0} \frac{[2\cdot 2^2 + 8h + 2h^2 + 2 + h - 3] - [2\cdot 2^2 + 2 - 3]}{h} \\
&= \lim_{h\to 0} \frac{2\cdot 2^2 + 8h + 2h^2 + 2 + h - 3 - 2\cdot 2^2 - 2 + 3}{h} \\
&= \lim_{h\to 0} \frac{9h+2h^2}{h} \\
&= \lim_{h\to 0}(9+2h) \\
&= 9.
\end{aligned}
$$

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.

Example 2.6. Let g be the function defined by $g(x) = \frac{1}{x+2}$. Use the definition of the derivative to find a formula for $g'(a)$.

Solution.

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

Example 2.7. Let f be the function defined by

$$f(x) = \begin{cases} x^3 & \text{if } x \geqslant 0 \\ -x^3 & \text{if } x < 0. \end{cases}$$

Show that $f'(0) = 0$.

Solution. Let us start our working out as usual.

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

We would now like to substitute in the definition of $f(h)$ and then work out the limit, but the formula for $f(h)$ depends on whether $h \geqslant 0$ or $h < 0$, and when we are working out the limit we do not know which of these applies. However, we can easily work out the left and right handed limits, as follows.

$$\begin{aligned} &\lim_{h \to 0^+} \frac{f(0+h) - f(0)}{h} \\ &= \lim_{h \to 0^+} \frac{f(h)}{h} \\ &= \lim_{h \to 0^+} \frac{h^3}{h} \\ &= \lim_{h \to 0^+} h^2 \\ &= 0. \end{aligned}$$

Here, we were able to replace $f(h)$ with h^3 since it is a limit as h approaches 0 from the right, so that we know $h > 0$ and consequently $f(h) = h^3$. And now we do the left-handed limit.

$$\begin{aligned} &\lim_{h \to 0^-} \frac{f(0+h) - f(0)}{h} \\ &= \lim_{h \to 0^-} \frac{f(h)}{h} \\ &= \lim_{h \to 0^-} \frac{-h^3}{h} \\ &= \lim_{h \to 0^-} (-h^2) \\ &= 0. \end{aligned}$$

Again, since this is a limit as h approaches 0 from the left, we knew that $h < 0$, and so were able to replace $f(h)$ with $-h^3$. Now, since

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

we can conclude that

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

as required.

Example 2.8. Define f by $f(x) = |x|$. Does $f'(0)$ exist?

Solution. Recall the definition of the absolute value:

$$|x| = \begin{cases} x & \text{if } x \geqslant 0 \\ -x & \text{if } x < 0 \end{cases}$$

We start working out the derivative as follows.

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

Now we see that, since the definition of $|h|$ depends on whether h is positive or negative, we must examine the left and right handed limit separately. This gives us

$$\lim_{h\to 0^+} \frac{|h|}{h} = \lim_{h\to 0^+} \frac{h}{h} = \lim_{h\to 0^+} 1 = 1.$$

Here we were able to replace $|h|$ with h since we are looking at a limit as h approaches 0 *from the right*, so that $h > 0$ and consequently $|h| = h$. And

$$\lim_{h\to 0^-} \frac{|h|}{h} = \lim_{h\to 0^-} \frac{-h}{h} = \lim_{h\to 0^-} -1 = -1.$$

Here we were able to replace $|h|$ with $-h$ since were looking at a limit as h approaches 0 *from the left*, so that $h < 0$ and consequently $|h| = -h$. Since $\lim_{h\to 0^+} \frac{|h|}{h}$ and $\lim_{h\to 0^-} \frac{|h|}{h}$ are not equal, it follows that $\lim_{h\to 0} \frac{|h|}{h}$ does not exist. Consequently $f'(0)$ does not exist.

2.2 The derivative as a function

Definition 2.9 (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

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

Example 2.10. Let f be the function defined by $f(x) = x^3 - x$. Use the definition of the derivative to find a formula for $f'(x)$.

Solution.

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

Example 2.11. Let $f(x) = \sqrt{x}$. Use the definition of the derivative to find a formula for $f'(x)$.

Solution.

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

The domain of $f'(x)$ is $(0, \infty)$.

Example 2.12. Let f be the function defined by $f(x) = |x|$. What is the domain of f'?

Solution. We already saw in Example 2.8 that $f'(0)$ does not exist. We will show here that $f'(x)$ *does* exist if $x \neq 0$, so that $\mathrm{dom}(f) = \{x \mid f'(x) \text{ exists}\} = \{x \mid x \neq 0\} = (-\infty, 0) \cup (0, \infty)$.

Case 1: If $x > 0$, then

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

and in particular $f'(x)$ does indeed exist. Here, we were able to replace $|x|$ with x since $x > 0$. And we were able to replace $|x+h|$ with $(x+h)$ for the following reason: we know that $x > 0$, and since we are looking at a limit as h approaches 0, we can assume that h is small — small enough that $(x+h) > 0$ also.

Case 2: If $x < 0$, then

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

and in particular $f'(x)$ does indeed exist. Here, we were able to replace $|x|$ with $-x$ since $x < 0$. And we were able to replace $|x+h|$ with $-(x+h)$ for the following reason: we know that $x < 0$, and since we are looking at a limit as h approaches 0, we can assume that h is small — small enough that $(x+h) < 0$ also.

So we have confirmed that $f'(x)$ exists when $x \neq 0$, and in fact we have the following formula for f'.

$$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

Example 2.13. Let k be the function defined by $k(x) = \sqrt[3]{x}$. Show that $k'(0)$ does not exist.

Solution. $k'(0)$, if it exists, is given by the following limit.

$$k'(0) = \lim_{h\to 0} \frac{k(0+h)-k(0)}{h} = \lim_{h\to 0} \frac{\sqrt[3]{h}-\sqrt[3]{0}}{h} = \lim_{h\to 0} \frac{\sqrt[3]{h}}{h} = \lim_{h\to 0} \frac{h^{\frac{1}{3}}}{h} = \lim_{h\to 0} \frac{1}{h^{\frac{2}{3}}}$$

But this limit is not finite. Therefore, the derivative does not exist at the point 0.

Example 2.14. Let H be the Heaviside function defined by

$$H(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t \geqslant 0 \end{cases}$$

Show that $H'(0)$ does not exist.

Solution. We must show that

$$\lim_{h \to 0} \frac{H(h) - H(0)}{h}$$

does not exist. Now, $H(0) = 1$, but the value of $H(h)$ will depend on whether $h > 0$ or $h < 0$. So to understand this limit we will look at the right and left handed limits. The right-handed limit does exist:

$$\lim_{h \to 0^+} \frac{H(h) - H(0)}{h} = \lim_{h \to 0^+} \frac{1 - 1}{h} = \lim_{h \to 0^+} \frac{0}{h} = \lim_{h \to 0^+} 0 = 0.$$

But the left-handed limit does not:

$$\lim_{h \to 0^-} \frac{H(h) - H(0)}{h} = \lim_{h \to 0^-} \frac{0 - 1}{h} = \lim_{h \to 0^-} \frac{-1}{h}$$

So the two-sided limit $\lim_{h \to 0} \frac{H(h)-H(0)}{h}$ does not exist either, and consequently $H'(0)$ does not exist.

In fact, the underlying reason that $H'(0)$ does not exist is that H is not continuous at 0, as in the following theorem.

Theorem 2.15. *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.*

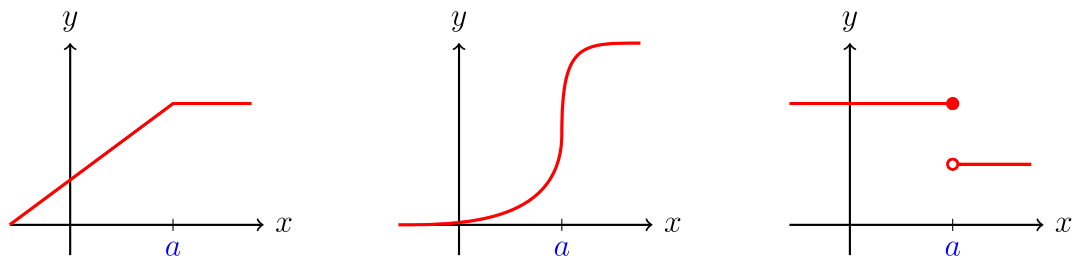
Proof. Assume that f is differentiable at a. We calculate:

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

This shows us that $\lim_{x \to a} f(x) = f(a)$, and so f is continuous at a. Notice that in order to write the limit of the product as product of the limits, we have used the fact that both limits exists and are finite. □

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

Example 2.16. Let f be the function defined as follows:

$$f(x) = \begin{cases} -x^2 & \text{if } x < 0 \\ x^2 & \text{if } x \geq 0 \end{cases}$$

Then $f'(x)$ exists for all x, and is given by the following formula:

$$f(x) = \begin{cases} -2x & \text{if } x < 0 \\ 2x & \text{if } x > 0 \\ 0 & \text{if } x = 0 \end{cases}$$

Or in other words, $f(x) = 2|x|$. We have seen that the absolute value function is continuous but not differentiable in zero. This shows us that $f'(x)$ does not have to be differentiable everywhere, even if f is.

The next example shows that the derivative does not even have to be continuous:

Example 2.17. Let f be the function defined as follows.

$$f(x) = \begin{cases} x^2 \sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Then $f'(x)$ exists for all x, and is given by the following formula:

$$f'(x) = \begin{cases} 2x\sin(1/x) - \cos(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

(The formula for $f'(x)$ when $x \neq 0$ is an exercise in differentiation rules, which we will see later. The formula for $f'(0)$ requires you to use the precise definition of the derivative together with the squeeze theorem.) Now, observe that f' is not continuous at $a = 0$. So we see that a function may be differentiable everywhere, but that its derivative may not be continuous.

Definition 2.18 (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 2.19 (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 *nth derivative of* f, which is denoted by $f^{(n)}$.

In Leibniz notation, if $y = f(x)$, then we would write

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

Remark 2.20. As was mentioned before, if $f(t)$ is a function which described the location of an object at time t, then the derivative $f'(t)$ gives us the velocity of the object at time t. The second derivative $f''(t)$ gives us the *acceleration* of the object at time t, since the acceleration measures the rate of change of the velocity with respect to the time.

Example 2.21. If f is defined by $f(x) = x^3 - x$, then find f' and f'' and f'''.

Solution. First,

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

Next,

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

Finally,

$$\begin{aligned}
f'''(x) &= \lim_{h\to 0} \frac{f''(x+h)-f''(x)}{h} \\
&= \lim_{h\to 0} \frac{6(x+h)-6x}{h} \\
&= \lim_{h\to 0} \frac{6h}{h} \\
&= \lim_{h\to 0} 6 \\
&= 6.
\end{aligned}$$

2.3 Differentiation formulas

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 positive integer, then

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

Proof. Let f be defined by $f(x) = x^n$. Then we must show that $f'(x) = nx^{n-1}$. And indeed,

$$\begin{aligned}
f'(x) &= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} \\
&= \lim_{h\to 0}\frac{(x+h)^n - x^n}{h} \\
&= \lim_{h\to 0}\frac{x^n + nx^{n-1}h + \cdots + \binom{n}{r}x^{n-r}h^r + \cdots + h^n - x^n}{h} \\
&= \lim_{h\to 0}\frac{nx^{n-1}h + \cdots + \binom{n}{r}x^{n-r}h^r + \cdots + h^n}{h} \\
&= \lim_{h\to 0} nx^{n-1} + \cdots + \binom{n}{r}x^{n-r}h^{r-1} + \cdots + h^{n-1} \\
&= nx^{n-1}.
\end{aligned}$$

□

Example 2.22.

- If $f(x) = x^6$ then $f'(x) = 6x^{6-1} = 6x^5$.
- If $y = t^4$ then $\frac{dy}{dt} = 4t^{4-1} = 4t^3$.
- $\frac{d}{dr}r^3 = 3r^{3-1} = 3r^2$.

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

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

Example 2.23.
- $\frac{d}{dx}(3x^4) = 3\frac{d}{dx}(x^4) = 3 \times 4x^3 = 12x^3$.
- $\frac{d}{dx}(-x) = -\frac{d}{dx}(x) = -\frac{d}{dx}(x^1) = -1 \times 1 \times x^{1-1} = -1$.

The sum rule. If f and g are both differentiable, then

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

Proof. We have

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

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

This implies that the limit exists and is finite, and therefore $(f+g)'(x) = f'(x) + g'(x)$. □

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.

Example 2.24. Compute $\frac{d}{dx}[3x^2 + 2x + 1]$.

Solution.

$$\begin{aligned}\frac{d}{dx}[3x^2 + 2x + 1] &= 3\frac{d}{dx}x^2 + 2\frac{d}{dx}x + \frac{d}{dx}1 \\ &= 3 \cdot 2x^{2-1} + 2x^{1-1} + 0 \\ &= 6x^1 + 2x^0 \\ &= 6x + 2\end{aligned}$$

Example 2.25. Compute $\frac{d}{dx}[2x^5 + 4x^3 - 3x^2 + 2]$.

Solution.

$$\begin{aligned}\frac{d}{dx}[2x^5 + 4x^3 - 3x^2 + 2] &= 2\frac{d}{dx}x^5 + 4\frac{d}{dx}x^3 - 3\frac{d}{dx}x^2 + \frac{d}{dx}2 \\ &= 2 \cdot 5x^{5-1} + 4 \cdot 3x^{3-1} - 3 \cdot 2x^{2-1} + 0 \\ &= 10x^4 + 12x^2 - 6x\end{aligned}$$

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,

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

Proof. We calculate:

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

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

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

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

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

Notice that we have used here the fact that $\lim_{h\to 0} g(x+h) = g(x)$. This follows from the fact that g is continuous at the point x, because g is differentiable at the point x. □

Example 2.26. Find $F'(x)$ if $F(x) = (4x^3)(7x^4)$.

Solution. Observe that $F(x)$ is the product of the functions given by $4x^3$ and $7x^4$. So by the product rule,

$$\begin{aligned} F'(x) &= \frac{d}{dx}[(4x^3) \cdot (7x^4)] \\ &= (4x^3) \cdot \frac{d}{dx}(7x^4) + \frac{d}{dx}(4x^3) \cdot (7x^4) \\ &= (4x^3) \cdot (28x^3) + (12x^2) \cdot (7x^4) \\ &= 112x^6 + 84x^6. \\ &= 196x^6. \end{aligned}$$

In this case it would have been quicker to first simplify the function and then differentiate. Indeed, $F(x) = (4x^3)(7x^4) = 28x^7$ so that $F'(x) = 28 \times 7x^6 = 196x^6$. However, when we come to use the product rule later we will not be able to simplify in this way.

Example 2.27. Suppose that f and g are functions and that $f(x) = x^2 \cdot g(x)$. Suppose also that we know that $g(2) = 1$ and $g'(2) = 3$. Find $f'(2)$.

Solution. Even though we don't know what f and g actually are, we can still use the product rule by regarding $f(x)$ as the product of x^2 with $g(x)$:

$$f'(x) = \frac{d}{dx}[x^2 \cdot g(x)] = x^2 \cdot \frac{d}{dx}g(x) + \frac{d}{dx}[x^2] \cdot g(x) = x^2 \cdot g'(x) + 2x \cdot g(x).$$

Now we can substitute $x = 2$ to find that

$$f'(2) = 2^2 \cdot g'(2) + (2 \times 2) \cdot g(2) = 4 \cdot g'(2) + 4 \cdot g(2) = 4 \cdot 3 + 4 \cdot 1 = 16.$$

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,

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

Proof. Similar to the product rule, we wll use the fact that $\lim_{h\to 0} g(x+h) = g(x)$. We calculate

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

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

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

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

$$\frac{f'(x)g(x)}{g(x)^2}+\frac{-f(x)g'(x)}{g(x)^2} =$$

$$\frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}.$$

□

Example 2.28. Let $y = \dfrac{x^2+2x-1}{x^3-2}$. Find y'.

Solution.

$$\begin{aligned}
y' &= \frac{d}{dx}\left[\frac{x^2+2x-1}{x^3-2}\right] \\
&= \frac{(x^3-2)\frac{d}{dx}(x^2+2x-1)-(x^2+2x-1)\frac{d}{dx}(x^3-2)}{(x^3-2)^2} \\
&= \frac{(x^3-2)(2x+2)-(x^2+2x-1)(3x^2)}{(x^3-2)^2} \\
&= \frac{(2x^4+2x^3-4x-4)-(3x^4+6x^3-3x^2)}{(x^3-2)^2} \\
&= \frac{-x^4-4x^3+3x^2-4x-4}{(x^3-2)^2}
\end{aligned}$$

Let's do the same example, but with some advice attached to it:

$$
\begin{aligned}
y' &= \frac{d}{dx}\left[\frac{x^2+2x-1}{x^3-2}\right] \\
&= \frac{(x^3-2)\frac{d}{dx}(x^2+2x-1)-(x^2+2x-1)\frac{d}{dx}(x^3-2)}{(x^3-2)^2} && \text{always write this out in full} \\
&= \frac{(x^3-2)(2x+2)-(x^2+2x-1)(3x^2)}{(x^3-2)^2} && \text{differentiate } \textit{before} \text{ multiplying out} \\
&= \frac{(2x^4+2x^3-4x-4)-(3x^4+6x^3-3x^2)}{(x^3-2)^2} && \text{keep the second part in a bracket} \\
&= \frac{-x^4-4x^3+3x^2-4x-4}{(x^3-2)^2} && \textit{now} \text{ subtract}
\end{aligned}
$$

finally, don't expand the bottom

Derivatives of root functions

We recall some relevant notions on root functions. The nth root of a is defined in case n is even and a is non-negative, or in case n is odd and a is any number. When n is odd, the nth root of a is the unique number b which satisfies $b^n = a$. When n is even, the nth root of a is the unique non-negative number b which satifsies $b^n = a$. We write

$$\sqrt[n]{a} = b \text{ or } a^{\frac{1}{n}} = b.$$

For a rational number (that is: a number which we can write as a quotient of two integers $\frac{m}{n}$ we write

$$a^{\frac{m}{n}} = \sqrt[n]{a}^m.$$

The usual rules for manipulating powers hold here:

Rules for manipulating powers.

- $\dfrac{1}{a^n} = a^{-n}$
- $a^p \cdot a^q = a^{p+q}$
- $a^{1/p} = \sqrt[p]{a}$
- $a^p b^p = (ab)^p$

- $\dfrac{a^p}{a^q} = a^{p-q}$
- $(a^p)^q = a^{pq}$

The function $f(x) = x^{\frac{m}{n}}$ is differentiable for $x \neq 0$, and it is also differentiable at $x = 0$ if $\frac{m}{n} \geq 1$. We have the following rule, which generalizes the previous rule for the derivative of x^n: **General power rule.**

$$\frac{d}{dx}(x^{\frac{m}{n}}) = \frac{m}{n}x^{\frac{m}{n}-1}.$$

Or: if we just write $\frac{m}{n} = r$ then:

$$\frac{d}{dx}(x^r) = rx^{r-1}.$$

Remark 2.29.

- Notice that this can also be used for negative r: if $r < 0$, then $x^r = \frac{1}{x^{-r}}$, and the rules of derivatives work the same for negative powers (and the function is differentiable for $x \neq 0$ in case the denominator in r is even, and for $x > 0$ in case the denominator in r is odd).
- This rule of derivation works the same in case the exponent r is not a rational number, but a real number. We will explain later what we mean by expression such as $2^{\sqrt{2}}$ or 3^{π}.

Example 2.30.

- If $y = \dfrac{1}{x}$, then $\dfrac{dy}{dx} = \dfrac{d}{dx}\left[\dfrac{1}{x}\right] = \dfrac{d}{dx}[x^{-1}] = (-1)x^{-1-1} = -x^{-2}$.
- $\dfrac{d}{dt}\left[\dfrac{6}{t^3}\right] = 6\dfrac{d}{dt}\left[\dfrac{1}{t^3}\right] = 6\dfrac{d}{dt}[t^{-3}] = 6 \times (-3) \times t^{-3-1} = -18t^{-4}$.
- If $f(x) = x^{0.8}$, then $f'(x) = 0.8x^{-0.2}$.
- If $y = \dfrac{1}{\sqrt[3]{x^2}}$, then

$$y = \frac{1}{(x^2)^{1/3}} = \frac{1}{x^{2/3}} = x^{-\frac{2}{3}}$$

so

$$\frac{dy}{dx} = \frac{d}{dx}(x^{-\frac{2}{3}}) = -\frac{2}{3}x^{-\frac{2}{3}-1} = -\frac{2}{3}x^{-\frac{5}{3}}.$$

Example 2.31. Differentiate the function f defined by $f(t) = \sqrt{t}(a + bt)$.

Solution.

$$\begin{aligned}
f'(t) &= \frac{d}{dt} f(t) \\
&= \frac{d}{dt}\left(\sqrt{t}(a + bt)\right) \\
&= \frac{d}{dt}\left(t^{\frac{1}{2}}(a + bt)\right) \\
&= \frac{d}{dt}\left(at^{\frac{1}{2}} + bt^{\frac{3}{2}}\right) \\
&= a\frac{d}{dt}\left(t^{\frac{1}{2}}\right) + b\frac{d}{dt}\left(t^{\frac{3}{2}}\right) \\
&= \frac{1}{2}at^{-\frac{1}{2}} + \frac{3}{2}bt^{\frac{1}{2}}.
\end{aligned}$$

Example 2.32. Differentiate $y = \dfrac{\sqrt{x}}{1 + x^2}$.

Solution.

$$\begin{aligned}
\frac{dy}{dx} &= \frac{(1 + x^2)\frac{d}{dx}\left(\sqrt{x}\right) - \sqrt{x}\frac{d}{dx}(1 + x^2)}{(1 + x^2)^2} \\
&= \frac{(1 + x^2)\left(\frac{1}{2}x^{-\frac{1}{2}}\right) - x^{\frac{1}{2}}(2x)}{(1 + x^2)^2} \\
&= \frac{\frac{1}{2}x^{-\frac{1}{2}} + \frac{1}{2}x^{\frac{3}{2}} - 2x^{\frac{3}{2}}}{(1 + x^2)^2} \\
&= \frac{\frac{1}{2}x^{-\frac{1}{2}} - \frac{3}{2}x^{\frac{3}{2}}}{(1 + x^2)^2}
\end{aligned}$$

2.4 Derivatives of trigonometric functions

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

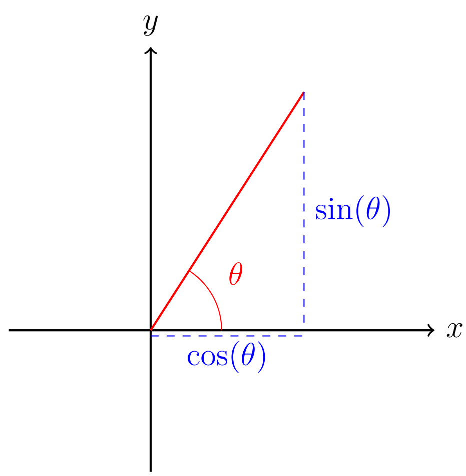
Definition 2.33 (The trigonometric functions). The *trigonometric functions* are as follows.

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

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

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

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:



In order to differentiate the trigonometric functions, we will need some more facts about them.

Sum-of-angles formulas.

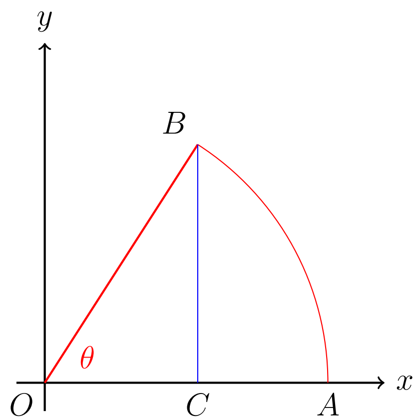
$$\sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b)$$
$$\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$$

Two special limits.

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

The sum of angles formulas should be familiar to you, but the two special limits may not be.

In the next couple of pages we will compute the first of these two limits, namely $\lim_{\theta\to 0}\frac{\sin(\theta)}{\theta} = 1$. The computation of the other one is quite similar. To begin, we assume that $\theta > 0$ and we draw a diagram depicting $\sin(\theta)$ and θ as lengths.



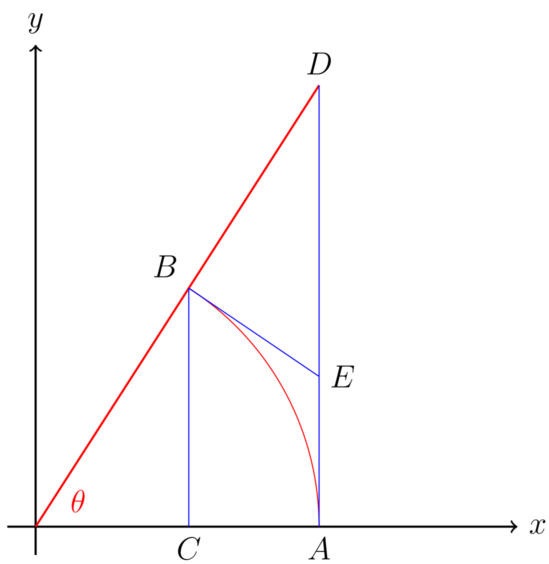
Here, the arc AB is a segment of a circle with radius 1 and making angle θ with the positive x-axis. Since the length of an arc of angle α in a circle of radius r is αr, we have:

$$|AB| = \theta \times 1 = \theta.$$

And by the definition of $\sin(\theta)$ and $\cos(\theta)$ we have:

$$|BC| = \sin(\theta)$$

Now we will extend our diagram to obtain a bit more information.



In this picture we found D by extending OB until its endpoint was directly above A. And we found E by drawing the line segment from B that makes a right-angle with BD, until we meet AD. So by considering the triangle OAD we see that

$$|AD| = \frac{|AD|}{|OA|} = \tan(\theta).$$

Now we write down some inequalities. We clearly have

$$|BC| < |AB|,$$

and since AE and EB form part of a polygon bounding the entire circle, we have

$$|AB| < |AE| + |EB| < |AE| + |ED| = |AD|.$$

So altogether we have

$$|BC| < |AB| < |AD|.$$

Substituting our computations of $|AB|$, $|BC|$ and $|AD|$ into these inequalities gives

$$\sin(\theta) < \theta < \tan(\theta)$$

and a little rearrangement gives

$$\cos(\theta) < \frac{\sin(\theta)}{\theta} < 1.$$

Now, we know that

$$\lim_{\theta\to 0^+} \cos(\theta) = 1 = \lim_{\theta\to 0^+} 1.$$

So by (a one-sided version of) the squeeze theorem, we find that

$$\lim_{\theta\to 0^+} \frac{\sin(\theta)}{\theta} = 1.$$

Since $\frac{\sin(\theta)}{\theta}$ is even, we know that

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

as well, so that

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

as required.

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

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

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

Proof that $\frac{d}{d\theta}\sin(\theta) = \cos(\theta)$*.*

$$\begin{aligned}
\frac{d}{d\theta}\sin(\theta) &= \lim_{h\to 0} \frac{\sin(\theta+h) - \sin(\theta)}{h} \\
&= \lim_{h\to 0} \frac{\sin(\theta)\cos(h) + \cos(\theta)\sin(h) - \sin(\theta)}{h} \\
&= \lim_{h\to 0} \left[\sin(\theta)\cdot\frac{\cos(h)-1}{h} + \cos(\theta)\cdot\frac{\sin(h)}{h}\right] \\
&= \sin(\theta)\cdot\lim_{h\to 0}\frac{\cos(h)-1}{h} + \cos(\theta)\cdot\lim_{h\to 0}\frac{\sin(h)}{h} \\
&= \sin(\theta)\cdot 0 + \cos(\theta)\cdot 1 \\
&= \cos(\theta)
\end{aligned}$$

□

The proof that $\frac{d}{d\theta}\cos(\theta) = -\sin(\theta)$ is similar.
Derivatives of trigonometric functions.

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

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

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

Warning. The symbol $\sec^2(\theta)$ means $[\sec(\theta)]^2$, and similarly for $\sin^2(\theta)$, $\cos^2(\theta)$ and so on. On the other hand, $\sin^{-1}(\theta)$ does *not* denote $\frac{1}{\sin(\theta)}$, but instead denots the inverse function, also called $\arcsin(\theta)$, and similarly for the other trigonometric functions.

Example 2.34. We can derive the differentiation formulas for tan, sec, csc and cot from the known differentiation formulas for sin and cos. For example,

$$\begin{aligned}
\frac{d}{d\theta}\cot(\theta) &= \frac{d}{d\theta}\left[\frac{\cos(\theta)}{\sin(\theta)}\right] \\
&= \frac{\sin(\theta)\frac{d}{d\theta}\cos(\theta) - \cos(\theta)\frac{d}{d\theta}\sin(\theta)}{\sin^2(\theta)} \\
&= \frac{-\sin(\theta)\sin(\theta) - \cos(\theta)\cos(\theta)}{\sin^2(\theta)} \\
&= -\frac{\sin^2(\theta) + \cos^2(\theta)}{\sin^2(\theta)} \\
&= -\frac{1}{\sin^2(\theta)} \\
&= -\csc^2(\theta).
\end{aligned}$$

Example 2.35. Find the 57th derivative of $\cos(x)$.

Solution.

$$\begin{aligned}
\frac{d}{dx}\cos(x) &= -\sin(x) \\
\frac{d^2}{dx^2}\cos(x) &= \frac{d}{dx}\left(-\sin(x)\right) = -\cos(x) \\
\frac{d^3}{dx^3}\cos(x) &= \frac{d}{dx}\left(-\cos(x)\right) = \sin(x) \\
\frac{d^4}{dx^4}\cos(x) &= \frac{d}{dx}\left(\sin(x)\right) = \cos(x)
\end{aligned}$$

So differentiating $\cos(x)$ four times gives us back $\cos(x)$. That means that the same is true if we differentiate it four times, or eight, or twelve, or ..., or 56 times. (Since $56 = 14 \times 4$.) Thus

$$\frac{d^{56}}{dx^{56}}\cos(x) = \cos(x)$$

and consequently

$$\frac{d^{57}}{dx^{57}}\cos(x) = \frac{d}{dx}\cos(x) = -\sin(x).$$

Example 2.36. Calculate $\lim_{x\to 0}\frac{\sin(5x)}{3x}$.

Solution. We'll use the fact that $\lim_{t\to 0}\frac{\sin(t)}{t} = 1$.

$$\begin{aligned}
\lim_{x\to 0}\frac{\sin(5x)}{3x} &= \lim_{x\to 0}\frac{5}{3}\cdot\frac{\sin(5x)}{5x} \\
&= \frac{5}{3}\cdot\lim_{x\to 0}\frac{\sin(5x)}{(5x)} \\
&= \frac{5}{3}\cdot\lim_{t\to 0}\frac{\sin(t)}{t} \\
&= \frac{5}{3}\cdot 1 \\
&= \frac{5}{3}.
\end{aligned}$$

Here, we made a 'substitution' of t in place of $5x$, since if x approaches 0, then so does $5x = t$.

Example 2.37. Calculate $\lim_{x\to 0}[x^2 \cot(x)]$.

Solution.

$$\begin{aligned}
\lim_{x\to 0}[x^2 \cot(x)] &= \lim_{x\to 0}\left[x^2 \frac{\cos(x)}{\sin(x)}\right] \\
&= \lim_{x\to 0}\left[x \cdot \cos(x) \cdot \frac{x}{\sin(x)}\right] \\
&= \lim_{x\to 0}\left[x \cdot \cos(x) \cdot \left[\frac{\sin(x)}{x}\right]^{-1}\right] \\
&= \lim_{x\to 0}[x] \cdot \lim_{x\to 0}[\cos(x)] \cdot \lim_{x\to 0}\left[\frac{\sin(x)}{x}\right]^{-1} \\
&= 0 \cdot 1 \cdot 1^{-1} \\
&= 0.
\end{aligned}$$

2.5 The Chain Rule

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

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

Proof. Write $g(a) = b$. We would like to calculate the derivative of $f \circ g$ at a. If $g(x) \neq g(a)$ for x close enough to a but not equal to a, we calculate:

$$\lim_{x\to a} \frac{f(g(x)) - f(g(a))}{x - a} = \lim_{x\to a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \frac{g(x) - g(a)}{x - a} =$$

$$\lim_{x\to a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \lim_{x\to a} \frac{g(x) - g(a)}{x - a} = \lim_{y\to b} \frac{f(y) - f(b)}{y - b} \lim_{x\to a} \frac{g(x) - g(a)}{x - a} =$$

$$f'(b)g'(a) = f'(g(a))g'(a)$$

and we are done. We have used here the fact that since g is differentiable at a, g is also continuous at a. Therefore, when $x \to a$ it holds that $g(x) \to b$.

The problem is that we may have many points in which $g(x) = g(a)$ even when $x \neq a$. For this, we define a new function:

$$Q(y) = \begin{cases} \frac{f(y)-f(b)}{y-b} & \text{if } y \neq b \\ f'(b) & \text{if } y = b \end{cases}$$

Notice that Q is a continuous function. It follows that $Q \circ g$ is also a continuous function, since it is the composition of two continuous functions. We define another function:

$$H(x) = \begin{cases} \frac{g(x)-g(a)}{x-a} & \text{if } x \neq a \\ g'(a) & \text{if } x = a \end{cases}$$

Notice that H is defined for all x in the domain of g, and is also continuous. We define a new function F to be $F(x) = (Q \circ g)H$. We claim now that

$$F(x) = ((Q \circ g)H)(x) = \frac{f(g(x)) - f(g(a))}{x - a} \text{ if } x \neq a.$$

Indeed, by using the formula for Q and for H we get that if $g(x) \neq g(a)$ then

$$F(x) = \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \frac{g(x) - g(a)}{x - a} = \frac{f(g(x)) - f(g(a))}{x - a}.$$

If $g(x) = g(a)$ then $H(x) = 0$ and therefore $F(x) = 0$. On the other hand $\frac{f(g(x))-f(g(a))}{x-a} = 0$ because $f(g(x)) = f(g(a))$. The function F is continuous because it is the product of two continuous functoins. We then have that

$$f'(g(x))g'(x) = F(a) = \lim_{x \to a} F(x) = \lim_{x \to a} \frac{f(g(x)) - f(g(a))}{x - a} = (f \circ g)'(a)$$

and we are done. □

Example 2.38. Find $F'(x)$ if F is defined by $F(x) = \sqrt{x^2 + 1}$.

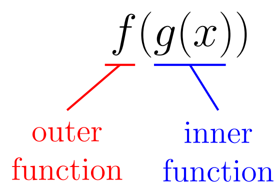
Solution. $F = f \circ g$ where $g(x) = x^2 + 1$ and $f(u) = \sqrt{u}$. Thus by the chain rule,

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

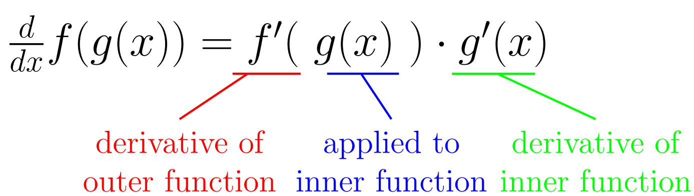
Now, $g'(x) = 2x$ and $f'(u) = \frac{d}{du}\sqrt{u} = \frac{d}{du}u^{\frac{1}{2}} = \frac{1}{2}u^{\frac{1}{2}-1} = \frac{1}{2}u^{-\frac{1}{2}}$. Thus

$$\begin{aligned} F'(x) &= f'(g(x)) \cdot g'(x) \\ &= f'(x^2+1) \cdot 2x \\ &= \frac{1}{2}(x^2+1)^{-\frac{1}{2}} \cdot 2x \\ &= \frac{x}{\sqrt{x^2+1}}. \end{aligned}$$

In applying the chain rule, we think of $f(g(x))$ as an *outer function* f applied to an *inner function* g.



Then the chain rule says:



Example 2.39. Differentiate $\sin(x^2)$ and $\sin^2(x)$.

Solution. Let us differentiate $\sin(x^2)$. In this case we take the outer function to be $\sin(\ )$ and the inner function to be x^2. Then the derivative of the outer function is $\cos(\ )$, and applying this to the inner function gives $\cos(x^2)$. And the derivative of the inner function is $2x$. So altogether we have

$$\frac{d}{dx}\sin(x^2) = \cos(x^2) \cdot 2x = 2x\cos(x^2).$$

Now let us differentiate $\sin^2(x) = [\sin(x)]^2$. In this case we take the outer function to be $[\ \ ]^2$ and the inner function to be $\sin(x)$. So the derivative of the outer function is $2[\ \ ]$ and applying this to the inner function gives $2[\sin(x)]$. And the derivative of the inner function is $\cos(x)$. So altogether we have

$$\frac{d}{dx}\sin^2(x) = \frac{d}{dx}[\sin(x)]^2 = 2[\sin(x)] \cdot \cos(x) = 2\sin(x)\cos(x).$$

The chain rule with the power rule.

$$\frac{d}{dx}[g(x)]^n = n[g(x)]^{n-1} \cdot g'(x).$$

Example 2.40. Let f be the function defined by $f(x) = \dfrac{1}{\sqrt[3]{x^2+x+1}}$. Find f'.

Solution. First, we write f as a power of another function.

$$f(x) = \frac{1}{\sqrt[3]{x^2+x+1}} = \frac{1}{(x^2+x+1)^{1/3}} = (x^2+x+1)^{-1/3}.$$

Thus, by the rule above, we have

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

Example 2.41. Differentiate $y = (2x+1)^5(x^3+x-1)^4$.

Solution. Here, y is given to us as the product of the two functions $(2x+1)^5$ and $(x^3+x-1)^4$. So we start by using the product rule. Then on the next line, we use the chain-and-powers rule.

$$\begin{aligned}\frac{dy}{dx} &= \frac{d}{dx}\left[(2x+1)^5\right](x^3-x+1)^4 + (2x+1)^5\frac{d}{dx}\left[(x^3-x+1)^4\right]\\ &= 5(2x+1)^4 \cdot (2+0) \cdot (x^3-x+1)^4 + (2x+1)^5 \cdot 4(x^3-x+1)^3 \cdot (3x^2-1+0)\\ &= 2 \cdot (2x+1)^4 \cdot (x^3-x+1)^3\left[5 \cdot (x^3-x+1) + 2 \cdot (2x+1) \cdot (3x^2-1)\right]\\ &= 2 \cdot (2x+1)^4 \cdot (x^3-x+1)^3 \cdot (17x^3-9x+6x^2+3)\end{aligned}$$

Example 2.42. Differentiate $y = \sin(\sin(\sin(x)))$.

Solution. We will have to use the chain rule twice, as follows.

$$\begin{aligned}\frac{dy}{dx} &= \cos(\sin(\sin(x))) \cdot \frac{d}{dx}\left[\sin(\sin(x))\right]\\ &= \cos(\sin(\sin(x))) \cdot \left[\cos(\sin(x)) \cdot \frac{d}{dx}\sin(x)\right]\\ &= \cos(\sin(\sin(x))) \cdot \cos(\sin(x)) \cdot \cos(x)\end{aligned}$$

2.6 Implicit differentiation

The functions we have met so far have been described by expressing one variable explicitly in terms of another variable, for example

$$y = \sqrt{x^3 + 1} \text{ or } y = x \cos(x).$$

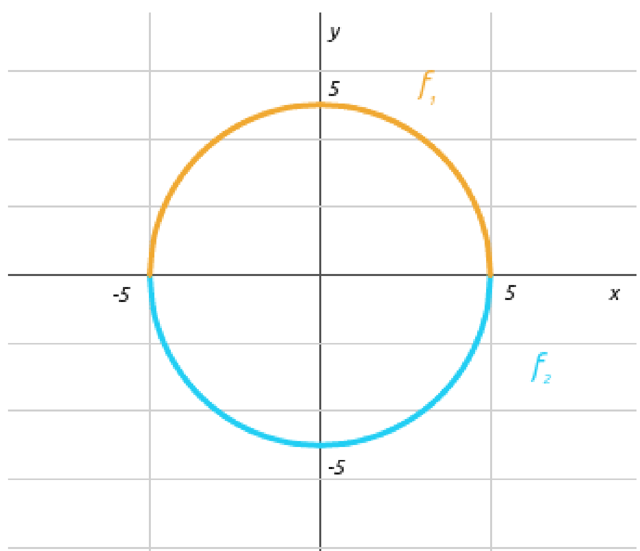
Some functions, however, are described instead by a relation between two variables.

For example,

$$x^2 + y^2 = 25.$$

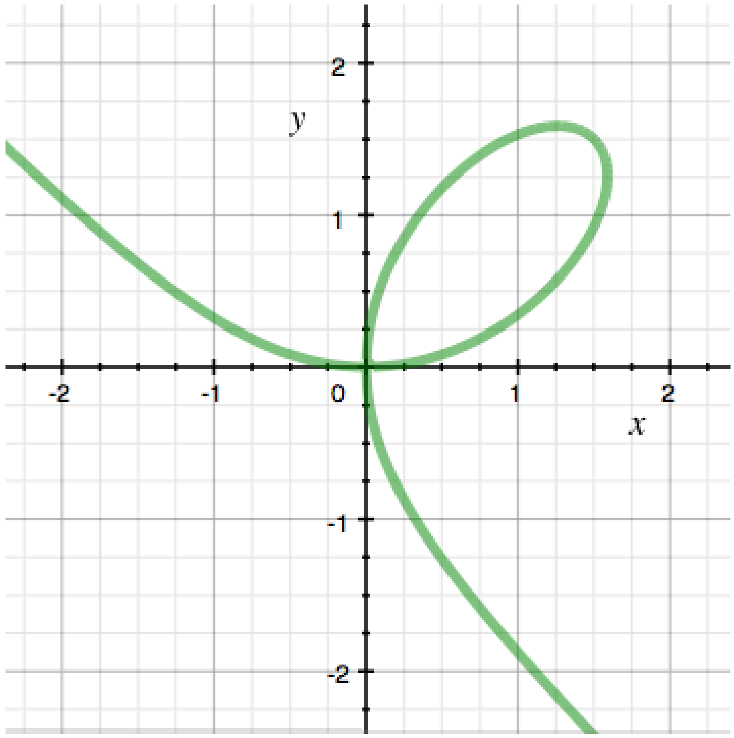
The set of all points (x, y) satisfying this equation determines a curve — in this case it is the circle centered at $(0, 0)$ and with radius 5.

However, this is not the graph of a single function, since there are vertical lines which intersect the curve at two points (and not at one point). However, we say that the equation *implicitly* defines a function f if, when we substitute $f(x)$ in place of y, the equation holds true for all values of x in the domain of f. For example, if we define f_1 and f_2 by $f_1(x) = \sqrt{25 - x^2}$ and $f_2(x) = -\sqrt{25 - x^2}$, then f_1 and f_2 are both implicitly defined by our equation.



Let us take another example. It is called the *folium of Descartes*.

$$x^3 + y^3 = 6xy$$



Here there are many functions implicitly defined by the equation.

Now, even though it is not possible to express y *globally* as a function $f(x)$, it is nevertheless possible to express y *locally* as a function $f(x)$, and to find a formula for $\frac{dy}{dx}$ in terms of x and y.

We have the following theorem, which we shall not prove now. For this theorem we write $F_y(x, y)$ for the function resulting from deriving $F(x, y)$ with respect to y (we think of x as a constant and of y as a variable and derive it accordingly). Similarly, $F_x(x, y)$ is the function resulting from deriving $F(x, y)$ with respect to x.

Theorem 2.43. *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 2.44.
- 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 $\frac{-F_x(x,y)}{F_y(x,y)}$.

Example 2.45. Find an equation of the tangent line to the curve $x^2 + y^2 = 25$ at $(x, y) = (3, 4)$.

Solution. Differentiating both sides of $x^2 + y^2 = 25$ gives

$$\begin{aligned} \frac{d}{dx}\left[x^2 + y^2\right] &= \frac{d}{dx}[25] \\ 2x + \frac{d}{dx}y^2 &= 0 \\ 2x + 2y\frac{dy}{dx} &= 0 \end{aligned}$$

so that $\frac{dy}{dx} = -\frac{x}{y}$. For $x = 3$, $y = 4$, this gives $\frac{dy}{dx} = -3/4$. So the equation of the tangent is

$$(y - 4) = -\frac{3}{4}(x - 3).$$

What $\frac{dy}{dx}$ means here is that, if f is implicitly defined by our equation, then substituting $y = f(x)$ gives a valid formula for $f'(x)$.

Example 2.46. Find the points on the folium of Descartes $x^3 + y^3 = 6xy$ where $y' = 0$.

Solution. Differentiating both sides of $x^3 + y^3 = 6xy$ gives

$$3x^2 + 3y^2 \cdot \frac{d}{dx}(y) = 6\frac{d}{dx}(x) \cdot y + 6x \cdot \frac{d}{dx}(y)$$

where we have used the chain rule combined with the power law to differentiate y^3, and where we have used the product rule to differentiate $6xy$. This gives

$$3x^2 + 3y^2y' = 6y + 6xy'$$

and rearranging gives

$$y' = \frac{2y - x^2}{y^2 - 2x}.$$

This is true whenever $2x \neq y^2$. Thus $y' = 0$ if and only if $y = x^2/2$. However, not all points that satisfy the equation $y = x^2/2$ actually lie on our curve. To find those points, we substitute $y = x^2/2$ into the original equation $x^3 + y^3 = 3xy$ to get

$$x^3 + \left(\frac{x^2}{2}\right)^3 = 6x\left(\frac{x^2}{2}\right).$$

Rearranging this gives us

$$x^3(x^3 - 16) = 0$$

so that $x = 0$ or $x = 2^{4/3}$. The corresponding y values are $y = 0^2/2 = 0$ and $y = (2^{4/3})^2/2 = 2^{5/3}$. But when $y = 0$ it holds that $x = 0$, and the derivative of the equation just gives us $0 = 0$, which contains no information on y'.

So the only required point is

$$(2^{4/3}, 2^{5/3}).$$

Example 2.47. Find y'' if $x^4 + y^4 = 16$.

Solution. Differentiating both sides of

$$x^4 + y^4 = 16$$

gives

$$4x^3 + 4y^3y' = 0$$

or equivalently

$$y' = -\frac{x^3}{y^3}.$$

We may now differentiate both sides of this equation to obtain a formula for y'.

$$\begin{aligned} y'' &= \frac{d}{dx}\left[-\frac{x^3}{y^3}\right] \\ &= -\frac{3x^2y^3 - x^3 \cdot 3y^2 \cdot y'}{(y^3)^2} \end{aligned}$$

This formula can be simplified but, more importantly, we can also substitute our known value for y'. This gives:

$$\begin{aligned} y'' &= -\frac{3x^2y^3 - x^3 \cdot 3y^2 \cdot \left(-\frac{x^3}{y^3}\right)}{y^6} \\ &= -\frac{3x^2y^4 + 3x^6}{y^7} \end{aligned}$$

This is a formula for y'' in terms of just x and y, which is good, but it happens that we can simplify it further, by using the original equation $x^4 + y^4 = 16$.

$$\begin{aligned} y'' &= -3x^2\frac{y^4 + x^4}{y^7} \\ &= -3x^2\frac{y^4 + x^4}{y^7} \\ &= -3x^2\frac{16}{y^7} \\ &= -48\frac{x^2}{y^7} \end{aligned}$$

2.7 Maximum and minimum values

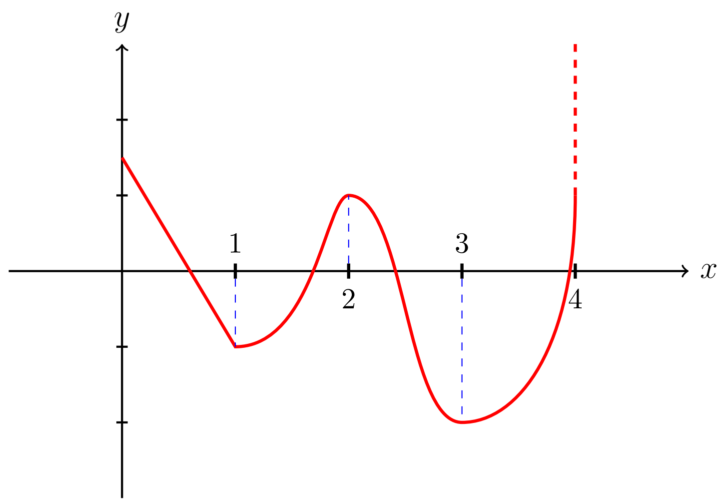
Many real-life problems involve computing the maximum or minimum value (the 'extreme values') of a function. For example, when we throw an object in the air, what height does it reach? Or what is the maximum acceleration of a space shuttle? If the acceleration is too great it may have profound consequences for the astronauts on board.

Definition 2.48 (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 $[a, b]$.)

Example 2.49. Consider the function f with the following graph. Observe that the domain of f is $[0, 4)$.



- f has absolute minimum -2 at 3.
- f has no absolute maximum.
- f has a local maximum of $f(2) = 1$ at 2.
- f has a local minimum of $f(1) = -1$ at 1.

Example 2.50. Let f be the function with domain $\mathbb{R}$ defined by $f(x) = \cos(x)$. Then f has absolute maximum 1, attained at the points $c = 0, 2\pi, 4\pi, \ldots, -2\pi, -4\pi, \ldots$. These points are all the locations of local maxima.

Example 2.51. Let h be the function with domain $\mathbb{R}$ defined by $h(x) = x^2$. This function has no absolute or local maximum, but it has an absolute minimum, at $x = 0$.

Example 2.52. Let g be the function with domain $\mathbb{R}$ defined by $g(x) = x^3$. This function has no absolute or local maxima or minima.

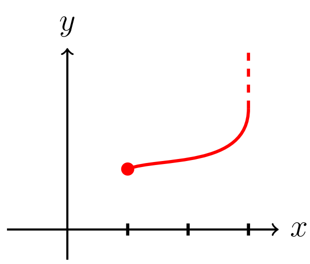
The above examples demonstrate several important points:

- Absolute maxima and minima in the interior of the domain are always local maxima and minima.
- However, local maxima and minima are not necessarily the absolute maximum or minimum.
- There may not be an absolute maximum or minimum at all.
- The absolute maximum and minimum may be attained at several different points.

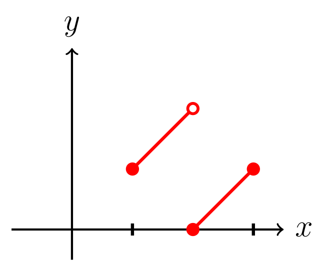
The following theorem guarantees that there are absolute maximum and minimum values under certain conditions.

Theorem 2.53 (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 $c, d \in [a, b]$.*

In the statement of the theorem, the conditions that f is *continuous* and that the interval is *closed* are important. Indeed, the theorem is false if they are not included, as the following examples show.



domain $[1, 3)$
continuous
absolute minimum
no absolute maximum



domain $[1, 3]$
not continuous
absolute minimum
no absolute maximum

In order to study local maxima and minima more carefully, we introduce the notion of critical number. As the theorem below shows, the critical numbers

Definition 2.54 (Critical number)*.* 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.

Theorem 2.55. *(Fermat) If f has a local minimum or maximum at c, then c is a critical number of f.*

Proof. One option is that $f'(c)$ is just not defined. In this case c is a critical number, by definition of a critical number. The second option is that $f'(c)$ is defined. We will show that in this case it is zero.

Assume that c is a local maximum of f. The derivative $f'(c)$ can be calculated using one-sided limits. So on the one hand, we have that

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

Since f has a local maximum at c, the expressoin $f(c+h) - f(c)$ is ≤ 0 for h small enough, while h is positive This means that the expressoin in the limit is non-positive, and therefore the resulting limit is non-positive: $f'(c) \leq 0$.

On the other hand, we have the other one-sided limit:

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

Now the expression $f(c+h) - f(c)$ is still ≤ 0, again, because f has a local maximum at c. But h is negative, so the resulting expression is non-negative. This means that the limit is non-negative as well, so we get $f'(c) \geq 0$.

The only way the two one sided limits can agree, is when $f'(c) = 0$. □

Example 2.56. Let f be the function defined by $f(x) = x^3 - x$. Then $f'(x) = 3x^2 - 1$ is defined for all x, so the critical numbers are just the c for which $f'(c) = 0$, or in other words $c = \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}$.

Example 2.57. Let g be the function defined by $g(x) = |x|$. Then we have seen that

$$g'(x) = \begin{cases} -1 & \text{if } x < 0, \\ 1 & \text{if } x > 0. \end{cases}$$

Since $g'(0)$ is not defined, 0 is a critical number of g. For all other x, $g'(x)$ is defined but $g'(x) \neq 0$, and that means that there are no more critical numbers of g.

Example 2.58. Find the critical numbers of f where $f(x) = x^{\frac{3}{5}}(4 - x)$.

Solution. First we work out $f'(x)$.

$$\begin{aligned} f'(x) &= \frac{d}{dx}(x^{\frac{3}{5}})(4-x) + x^{\frac{3}{5}}\frac{d}{dx}(4-x) \\ &= \tfrac{3}{5}x^{-\frac{2}{5}}(4-x) - x^{\frac{3}{5}} \\ &= \tfrac{12}{5}x^{-\frac{2}{5}} - \tfrac{8}{5}x^{\frac{3}{5}} \end{aligned}$$

Thus $f'(x)$ exists for all x except $x = 0$. And for $x \neq 0$ we have $f'(x) = 0$ when $\frac{12}{5}x^{-\frac{2}{5}} - \frac{8}{5}x^{\frac{3}{5}}$. Multiplying both sides of this equation by $\frac{5}{4}x^{\frac{2}{5}}$ gives $3 - 2x = 0$, i.e. $x = \frac{3}{2}$. So the critical numbers of f are 0 and $\frac{3}{2}$.

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.

Example 2.59. Find the absolute maximum and minimum values of the function f defined by $f(x) = x^3 - 3x^2 + 1$ in the interval $[1, 3]$.

Solution. First, f is a polynomial, so is continuous on any interval, in particular on the closed interval $[1, 3]$. This means that we may apply the closed interval method.

1. First we find the critical numbers of f in $(1, 3)$. Differentiating gives

   $$f'(x) = 3x^2 - 6x = 3x(x - 2).$$

   Thus $f'(x)$ exists for all x, so the critical numbers of f are simply those c for which $f'(c) = 0$, i.e. $x = 0$ and $x = 2$. The only one of these which lies in $(1, 3)$ is $x = 2$, and $f(2) = 8 - 3 \cdot 4 + 1 = -3$.

2. Next, we find the values of f at the endpoints of the interval. We get that $f(1) = 1 - 3 + 1 = -1$ and $f(3) = 27 - 3 \cdot 9 + 1 = 1$.

3. We now need to take the largest and smallest of the numbers from parts 1 and 2. These are $f(2) = -3$, $f(1) = -1$ and $f(3) = 1$. So the absolute maximum of f on $[1, 3]$ is 1 attained at 3, and the absolute minimum of f on $[1, 3]$ is -3 attained at 2.

Example 2.60. Find the absolute maximum and absolute minimum of the function $f(x) = 2x - \sin(x)$ in the interval $[0, \pi/2]$.

Solution. We calculate the derivative. It is $f'(x) = 2 - \cos(x)$. In particular, it is always defined and positive. This means that there are no critical numbers in $(0, \pi/2)$, and we just need to calculate the value of f at the end points. We have $f(0) = 2 \cdot 0 - \sin(0) = 0$ and $f(\pi/2) = 2 \cdot \pi/2 - \sin(\pi/2) = \pi - 1 > 0$. So $f(0) = 0$ is the absolute minimum of the function, and $f(\pi/2) = \pi - 1$ is the absolute maximum of the function.

Example 2.61. Find the absolute maximum and minimum values of the function f defined by $f(x) = x^3 + 2x$ in the interval $[a, b]$ where $a < b$.

Solution. At first glance, this exercise seems a bit strange, because we do not know what a and b are. We can still solve this! For this, we consider the derivative $f'(x) = 3x^2 + 2$. Since $x^2 \geq 0$ for every x and $2 > 0$, we see that $f'(x) > 0$ for every x. Also, we see that $f'(x)$ is defined for every x. This implies that the function is increasing on $[a, b]$, and therefore the absolute minimum is at $x = a$ and the absolute maximum is at $x = b$.

2.8 The Mean Value Theorem

Imagine you are driving a car from point A to point B. The velocity of your car might change throughout the drive: you start slow, then you accelerate, then you slow down et cetera. If the distance between point A and point B is X, and the amount of time the drive took you is T, then the average velocity of your drive was $\frac{X}{T}$. The mean value theorem is the following:

Theorem 2.62 (Mean Value Theorem, imprecise statement)**.** *The velocity of your car at some point, between A and B, was exactly the average velocity $\frac{X}{T}$.*

The precise statement of the mean value theorem is the following:

Theorem 2.63 (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 2.64 (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 $f'(c) = 0$.

Proof. We have the following three cases:

Case I : Assume that the function is constant: $f(x) = f(a) = f(b)$ for every x in $[a, b]$. Then the derivative is zero, and we can take c to be any number in (a, b).

Case II : Assume that $f(x) > f(a)$ for some x in (a, b). We know, by the extreme value theorem, that f has a maximum in $[a, b]$. Since $f(x) > f(a) = f(b)$, this maximum is not a nor b. So the maximum is attained at some c in (a, b). But then c is also a local maximum, and it is therefore a crtical number. The derivative $f'(c)$ exists, and by the Theorem of Fermat we have $f'(c) = 0$.

Case III : Assume that $f(x) < f(a)$ for some x in (a, b). This is similar to Case II. Fill in the details!

□

We now use the theorem of Rolle to prove the Mean Value Theorem:

Proof of MVT. : Write $s = \frac{f(b)-f(a)}{b-a}$. Consider the function $g(x) = f(x) - sx$. We calculate: $g(a) = f(a) - sa$ and $g(b) = f(b) - sb$. This implies:

$$g(a) - g(b) = f(a) - f(b) - sa + sb =$$

$$f(a) - f(b) + s(b-a) =$$

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

$$f(a) - f(b) + f(b) - f(a) = 0.$$

This means that $g(a) = g(b)$. The function g satisfies the condition of the theorem of Rolle: Since f and sx are continuous in $[a,b]$ and differentiable in (a,b) the same is true for the difference $g(x) = f(x) - sx$. Therefore, by the theorem of Rolle, there is a point $c \in (a,b)$ such that $g'(c) = 0$. But $g'(c) = f'(c) - s$. This means that

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

which is what we wanted to prove. □

The mean value theorem has the following corollary:

Theorem 2.65. *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 2.66. *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! □

Example 2.67. prove that $f(x) = x^3 + x - 1$ has exactly one real root.

Solution. First, notice that $f(0) = -1 < 0$ and $f(1) = 1 > 0$. This means that f has a root between 0 and 1. Next we show that there are no two roots. We do this by contradiction: Assume that a and b are two roots of f, and $a < b$. Then by the theorem of Rolle, there is a point $a < c < b$ such that $f'(c) = 0$. But $f'(x) = 3x^2 + 1$ is always positive, so this is impossible. We thus have only one root.

Example 2.68. Suppose that $f(0) = -3$ and $f'(x) \leq 5$ for all x. How large can $f(2)$ be?

Solution. The MVT gives us that $\frac{f(2)-f(0)}{2-0} = f'(c)$ for some $c \in (0, 2)$. So $f(2) = 2f'(c) - 3 \leq 2 \cdot 5 - 3 = 7$. This means that $f(2)$ is bounded above by 7.

2.9 How derivatives affect the shape of a graph

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.

Example 2.69. On what open intervals is the function f defined by $f(x) = 3x^4 - 4x^3 - 12x^2 + 5$ increasing? On what open intervals is it decreasing?

Solution. We first work out the derivative of f.

$$f'(x) = 12x^3 - 12x^2 - 24x = 12x(x^2 - x - 2) = 12x(x - 2)(x + 1).$$

This vanishes for $x = -1, 0, 2$, and we now work out what happens on the intervals obtained by deleting these points.

- On $(-\infty, -1)$ we have $(x+1) < 0$, $x < 0$, $(x-2) < 0$, so that $f'(x) < 0$. Consequently f is decreasing on $(-\infty, -1)$.
- On $(-1, 0)$ we have $(x + 1) > 0$, $x < 0$, $(x - 2) < 0$, so that $f'(x) > 0$. Consequently f is increasing on $(-1, 0)$.
- On $(0, 2)$ we have $(x + 1) > 0$, $x > 0$, $(x - 2) < 0$, so that $f'(x) < 0$. Consequently f is decreasing on $(0, 2)$.
- On $(2, \infty)$ we have $(x + 1) > 0$, $x > 0$, $(x - 2) > 0$, so that $f'(x) > 0$. Consequently f is increasing on $(2, \infty)$.

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

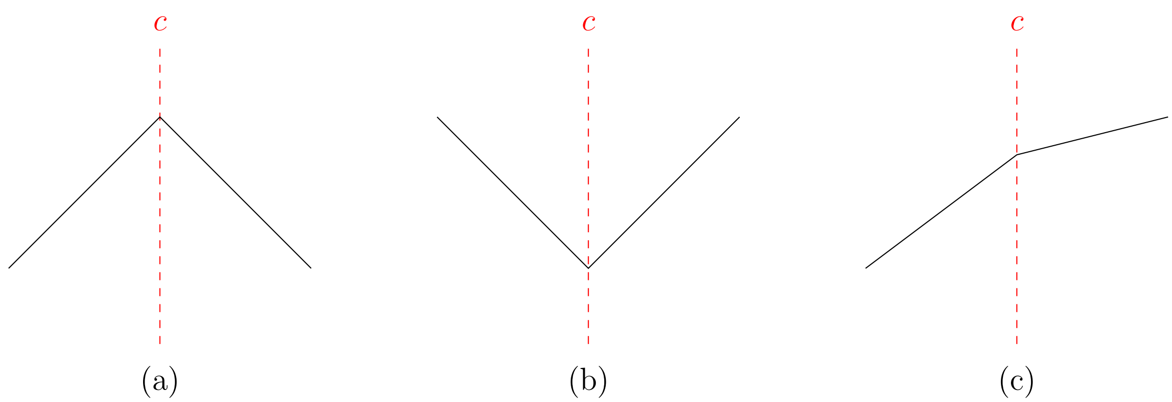
The first derivative test. Suppose that c is a critical number of a continuous function f.

(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 $f''(c) = 0$.

Example 2.70. Classify the local maxima and minima of the function f defined by $f(x) = x^5 - 5x^4$.

Solution. First we find f' and f''.

$$f'(x) = 5x^4 - 20x^3 = 5x^3(x-4), \qquad f''(x) = 20x^3 - 60x^2.$$

So the critical numbers of f are $x = 0$ and $x = 4$.

Now $f''(4) = 20 \cdot 4^4 - 60 \cdot 4^2 = 4160 > 0$, and f'' is defined and continuous near 4, so that the second derivative test applies and tells us that f has a local minimum at 4.

Next, $f''(0) = 0$, so that the second derivative test tells us nothing. We must instead use the first derivative test. Observe that if x is close to 0, i.e. small, then in the expression $f'(x) = 5x^3(x - 4)$ the term 5 is positive, $(x - 4)$ is negative, and x^3 has the same sign as x. So if x is close to 0 and negative, then $f'(x)$ is positive, and if x is close to 0 and positive, then $f'(x)$ is negative. So $f'(x)$ changes from positive to negative at $x = 0$, and hence f has a local maximum at 0.

Example 2.71. For the function f of Example 2.17, 0 is a critical number, but neither the first nor the second derivative tests apply. Indeed, f has neither a local maximum nor a local minimum at 0.

2.10 Exponential functions and their derivatives

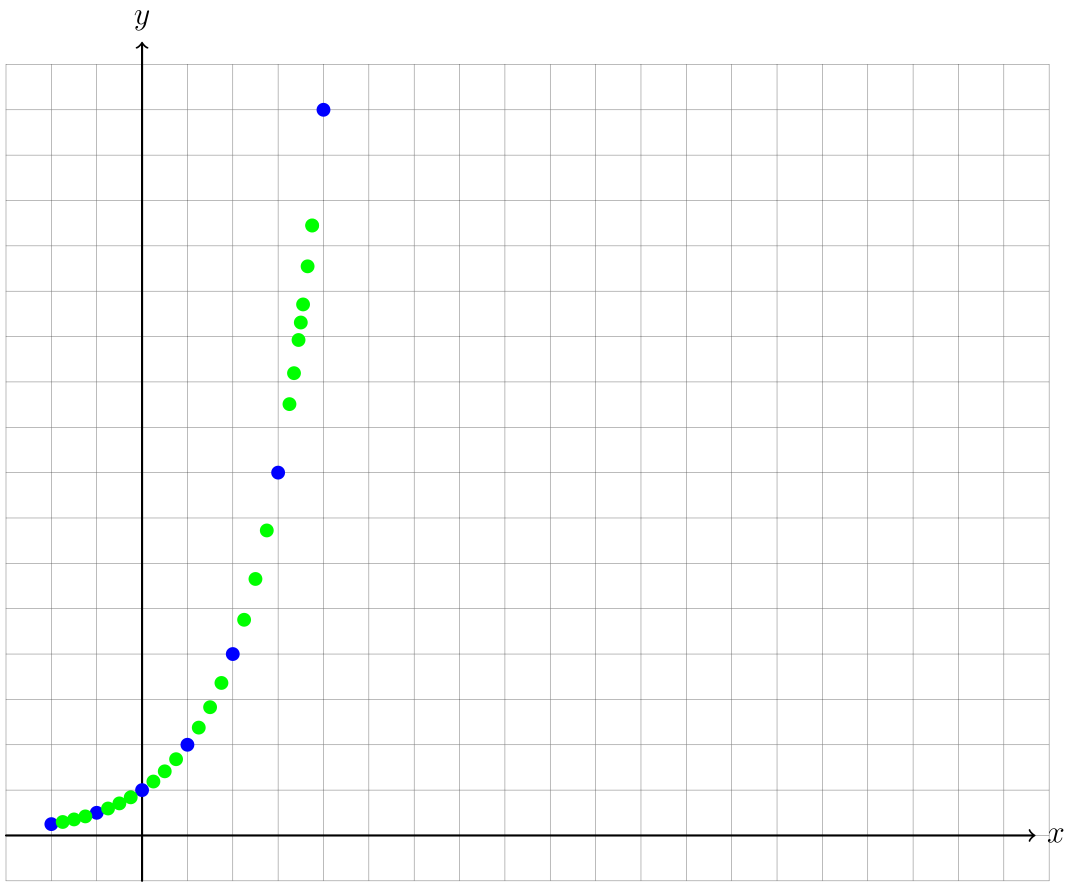
Definition 2.72 (Exponential functions). Let $a > 1$ be a real number. We want to define the associated *exponential function*

$$f(x) = a^x.$$

What do we actually mean when we speak about a^x? When x is rational, i.e. when $x = \frac{p}{q}$ with p, q integers and $q > 0$, then

$$f(x) = a^{\frac{p}{q}} = (\sqrt[q]{a})^p.$$

However, we want to define a^x for all real numbers, not only for rational ones. The next drawing shows some values of 2^x for x rational. The integer values of x are marked with blue, and the non-integer values are marked with green.

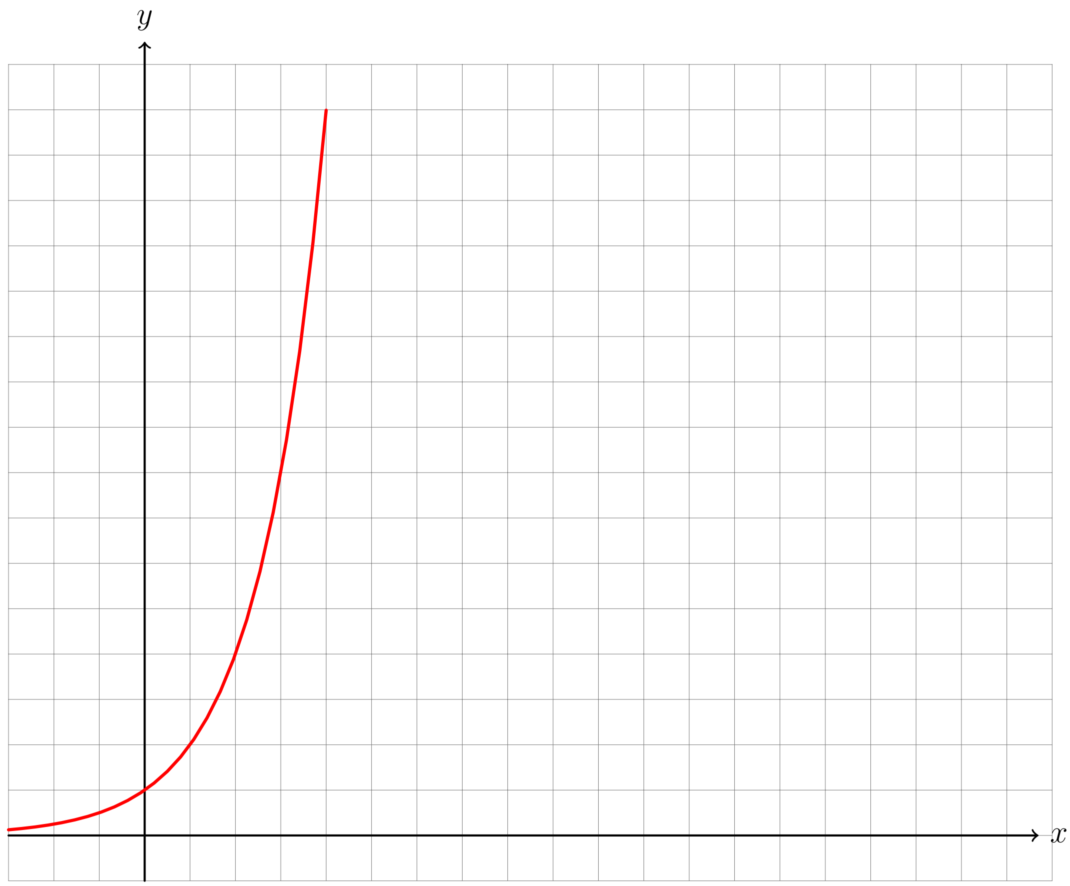


We notice the following two properties of the drawing :

1. If $x < y$ are rational numbers, then $a^x < a^y$.
2. The points on the drawing look like values of a function, one which goes up really fast!

The points in the drawings are in fact part of a graph of a function. We use the following fact, without proving it:
FACT: Assume that $a > 1$. For every real number x there is a unique real number a^x such that: if $r_1 < x < r_2$ and r_1 and r_2 are rational numbers, then $a^{r_1} < a^x < a^{r_2}$.

Here is a sketch of the graph of a^x:



In order to give a precise definition of a^x when x is real we need some more advanced tools, like the limit of a sequence. We shall not give the precise definition in this course. Regardless of the precise definition, it is enough for us to know that all of the rules for manipulating power functions listed on page 83 apply.

Let us differentiate $f(x) = a^x$:

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

It is a fact, which we are not in a position to prove here, that $\lim_{h\to 0} \frac{a^h - a^0}{h}$ always exists, and moreover, that there is a unique choice of a for which $\lim_{h\to 0} \frac{a^h - a^0}{h} = 1$.

Definition 2.73. We define e to be the unique number for which $\lim_{h\to 0} \frac{e^h - e^0}{h} = 1$.

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 e^x.

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

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

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

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

Example 2.74. Let k be defined by $k(x) = e^{x^2}$. Find $k'(x)$.

Solution. $\frac{d}{dx} e^{x^2} = e^{x^2} \cdot \frac{d}{dx} x^2 = e^{x^2} \cdot 2x.$

Example 2.75. Let y be defined by $y = e^{-5x} \sin(4x)$. Find y'.

Solution.

$$\begin{aligned} y' &= \frac{d}{dx}\left[e^{-5x} \cdot \sin(4x)\right] \\ &= \frac{d}{dx}\left[e^{-5x}\right] \cdot \sin(4x) + e^{-5x} \cdot \frac{d}{dx}\left[\sin(4x)\right] \\ &= -5e^{-5x} \cdot \sin(4x) + e^{-5x} \cdot 4\cos(4x). \end{aligned}$$

Example 2.76. Consider the function $f(x) = e^{2x} - 2e^x$. Find where f is increasing and decreasing, and where f has a local maximum/ minimum.

Solution. We derive f and get

$$f'(x) = 2e^{2x} - 2e^x = 2e^x(e^x - 1).$$

Since e^x is always positive, f is decreasing when $e^x < 1$, that is, when $x < 0$, and f is increasing when $e^x > 1$, that is, when $x > 0$. By the first derivative test, f has a local minimum at $x = 0$.

2.11 Logarithmic functions

Definition 2.77 (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 2.78.

- $e^0 = 1$, 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

Let us now work out the derivative of ln. Let f be the function defined by $f(t) = e^t$. Then it holds that $\ln(f(t)) = t$. Deferentiating and using the chain rule we get:

$$\ln'(f(t))f'(t) = 1$$
$$\ln'(f(t)) = \frac{1}{f'(t)} = \frac{1}{e^t} = \frac{1}{f(t)}.$$

Substituing $f(t) = x$ we get

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

So we have:
The derivative of $\ln$**.**

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

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

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

Example 2.79. Find $\frac{d}{dx}\ln(\cos(x))$

Solution. Since $\ln(\cos(x)) = \ln g(x)$ where $g(x) = \cos(x)$, we have

$$\frac{d}{dx}[\ln(\cos(x))]] = \frac{\frac{d}{dx}\cos(x)}{\cos(x)} = \frac{-\sin(x)}{\cos(x)} = -\tan(x).$$

Example 2.80. Differentiate the function f defined by $f(x) = \sqrt[3]{\ln(x)}$.

Solution.

$$\frac{d}{dx}\left(\sqrt[3]{\ln(x)}\right) = \frac{d}{dx}(\ln(x))^{\frac{1}{3}} = \frac{1}{3}(\ln(x))^{\frac{-2}{3}} \cdot \frac{d}{dx}\ln(x) = \frac{1}{3}(\ln(x))^{\frac{-2}{3}} \cdot \frac{1}{x} = \frac{1}{3x\ln(x)^{\frac{2}{3}}}$$

Example 2.81. Find $\frac{d}{dx}\ln\left(\frac{x+1}{2x+1}\right)$.

Solution.

$$\begin{aligned}
\frac{d}{dx}\ln\left(\frac{x+1}{2x+1}\right) &= \frac{1}{\frac{x+1}{2x+1}}\cdot\frac{d}{dx}\left(\frac{x+1}{2x+1}\right)\\
&= \frac{2x+1}{x+1}\cdot\frac{(2x+1)\cdot\frac{d}{dx}(x+1)-(x+1)\cdot\frac{d}{dx}(2x+1)}{(2x+1)^2}\\
&= \frac{2x+1}{x+1}\cdot\frac{(2x+1)\cdot 1-(x+1)\cdot 2}{(2x+1)^2}\\
&= \frac{2x+1}{x+1}\cdot\frac{-1}{(2x+1)^2}\\
&= \frac{-1}{(x+1)(2x+1)}
\end{aligned}$$

Question. What is wrong with the following solution to the above example? Write first

$$\ln\left(\frac{x+1}{2x+1}\right) = \ln(x+1) - \ln(2x+1).$$

Derive this and get

$$\frac{1}{x+1} - \frac{2}{2x+1}$$

and simplify this to get

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

Solution. The problem is that in order to get the equality

$$\ln\left(\frac{x+1}{2x+1}\right) = \ln(x+1) - \ln(2x+1)$$

we need to assume that both $x+1$ and $2x+1$ are positive. So this solution is valid only for $x > \frac{-1}{2}$. The function is defined also when $x < -1$. We can also calculate, separately, in the interval $(-\infty, -1)$, and write there

$$\ln\left(\frac{x+1}{2x+1}\right) = \ln(-x-1) - \ln(-2x-1).$$

We can then derive this and get the same formula for the derivative for $x < -1$.

Example 2.82. Find the derivative of a^x where $a > 0$ is some number.

Solution. Using the fact that $a = e^{\ln(a)}$ we calculate and we get

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

Example 2.83. Find when is the function $f(x) = \ln x + \frac{1}{x}$ increasing, decreasing, and when does it have local maximum / minimum.

Solution. We derive first and get

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

The function is defined for $x > 0$ because that is the domain of definition of $\ln x$. Since x^2 is always positive in the domain of definition, we get that our function increases when $x > 1$ and decreases when $x < 1$. At $x = 1$ we have a local minimum.

2.13 Inverse functions

Remark 2.84. In this section we will look at inverse functions. Not every function actually has an inverse. The following definition tells us the ones that *do* have an inverse.

Definition 2.85 (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.

Example 2.86. The function f defined by $f(x) = x^3$ is one-to-one. One way to see this is from the definition: if $x_1 \neq x_2$ then $x_1^3 \neq x_2^3$. Another way to see it is from the graph (which we will not draw here), in which each horizontal line crosses the graph exactly once.

Example 2.87. The function g defined by $g(x) = x^2$ is *not* one-to-one. One way to see this is that, if we let $x_1 = 1$ and $x_2 = -1$, then we have $x_1 \neq x_2$ but $g(x_1) = 1 = g(x_2)$, so that the definition does not hold. Another way to see it is that if $c > 0$, then the horizontal line $y = c$ passes through the graph of $y = g(x)$ twice.

Definition 2.88 (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

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

Example 2.89. If f is the function defined by $f(x) = x^3$, then f^{-1} is defined by $f^{-1}(x) = x^{\frac{1}{3}}$. This is because, for this function f^{-1}, the condition

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

is equivalent to

$$y^{\frac{1}{3}} = x$$

which is equivalent to

$$y = x^3,$$

which is equivalent to

$$y = f(x).$$

We will see a perhaps more straightforward way of finding inverse functions shortly.

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

Example 2.90. Find the inverse of the function f defined by $f(x) = x^3 + 4$.

Solution. We follow the steps above.

1. The equation $y = f(x)$ is $y = x^3 + 4$.
2. Rearranging this gives $y - 4 = x^3$ and then $x = \sqrt[3]{y-4}$.
3. Swapping x and y gives $y = \sqrt[3]{x-4}$.
4. So $f^{-1}(x) = \sqrt[3]{x-4}$.

It is easy to understand the graph of an inverse function in terms of the original function. Indeed:

$$\begin{aligned} (x,y) \text{ lies on the graph of } f &\iff y = f(x) \\ &\iff x = f^{-1}(y) \\ &\iff (y,x) \text{ lies on the graph of } f^{-1}. \end{aligned}$$

This means that the graph of f^{-1} is obtained from the graph of f by switching the x and y values, or in other words, by reflecting in the line $y = x$.

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

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

Sketch proof of the formula for $(f^{-1})'(a)$. One of the two cancellation equations tells us that $f^{-1}(f(x)) = x$ for all x. Differentiating both sides of this equation gives

$$\frac{d}{dx} f^{-1}(f(x)) = \frac{d}{dx}(x)$$

or in other words

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

Rearranging gives

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

Now, substituting $f^{-1}(x)$ in place of x, and using the cancellation equation $f(f^{-1}(x)) = x$ gives us

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

□

Example 2.91. Define f by $f(x) = 2x + \cos(x)$. Find $(f^{-1})'(1)$.

Solution. Working out f' gives $f'(x) = 2 - \sin(x)$. so the formula for $(f^{-1})'(a)$ in the case $a = 1$ gives us

$$(f^{-1})'(1) = \frac{1}{f'(f^{-1}(1))} = \frac{1}{2 - \sin(f^{-1}(1))}.$$

It just remains to find $f^{-1}(1)$. However, $f(0) = 2 \cdot 0 + \cos(0) = 0 + 1 = 1$, and this exactly means that $f^{-1}(1) = 0$. So in the end we have

$$(f^{-1})'(1) = \frac{1}{2 - \sin(0)} = \frac{1}{2}.$$

Example 2.92. Prove the formula for the derivative of $\sqrt[n]{x}$ using the formula for the derivative of the inverse function

Solution. We think of $\sqrt[n]{x}$ as the inverse function of $f(x) = x^n$, where we restrict the domain to $(0, \infty)$ if n is even. Write $g(x) = \sqrt[n]{x}$. Using the formula $f'(x) = nx^{n-1}$ we have

$$g'(x) = 1/f'(g(x)) = \frac{1}{ng(x)^{n-1}} = \frac{1}{n}x^{\frac{1-n}{n}} = \frac{1}{n}x^{1-\frac{1}{n}}$$

as required.

Example 2.93. Find the derivatives of the inverse trigonometric functions arcsin and arccos. Write them as explicit formulas in x.

Solution. We show here the formula for arcsin. The formula for arccos is similar. Write $f(x) = \sin(x)$, $g(x) = f^{-1}(x) = \arcsin(x)$. We have

$$g'(x) = \frac{1}{f'(g(x))} = \frac{1}{\cos(\arcsin(x))}.$$

What is $\cos(\arcsin(x))$? If $\arcsin(x) = y$ then $\sin(y) = x$, and then $\cos^2(y) + \sin^2(y) = 1$ so $\cos^2(y) + x^2 = 1$ or $\cos^2(y) = 1 - x^2$. Since the range of arcsin is $[-\pi/2, \pi/2]$, we see that $\cos(y)$ is positive, so we get that

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

We similarly get

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

Chapter 3

Integration

3.1 Antiderivatives and the indefinite integral

Example 3.1. What function has, as a derivative, the function

$$f(x) = x^5 + 3x - 1?$$

Solution. The question is asking us to find $F(x)$ such that $g'(x) = f(x)$.

As a warm-up, let's take the derivative of $f(x)$. It is

$$f'(x) = 5x^4 + 3.$$

This doesn't actually help, but reminds us how to differentiate.

Differentiating a power of x, say x^n, is done by multiplying by the power and lowering the power by 1, i.e.,

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

To go backwards, we must therefore divide by the power plus one and then raise the power by one. So, we can take g to be

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

Note that this isn't the only solution, for example we could instead take

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

because the derivative of a constant (2 in this case) is 0.

Proposition 3.2. *Suppose that $f(x)$ is a function and $F(x)$, $G(x)$ are functions which satisfy*

$$F'(x) = f(x), \quad \textit{and} \quad G'(x) = f(x).$$

Then $F(x) - G(x)$ is a constant.

Proof. Set $H(x) = F(x) - G(x)$. Then

$$H'(x) = F'(x) - G'(x) = f(x) - f(x) = 0.$$

We must show that $H(x)$ is a constant, i.e., that $H(x) = H(y)$ for all x, y.

For this, suppose that $x < y$. By the Mean Value Theorem, there exists $z \in (x, y)$ such that

$$H(y) - H(x) = H'(z)(y - x).$$

Since $H'(z) = 0$, it follows that $H(x) = H(y)$. □

Definition 3.3. If $f(x)$ is a function, an *antiderivative* of $f(x)$ is any function $F(x)$ such that

$$F'(x) = f(x)$$

as functions. In this case, we write

$$\int f(x)\,dx = F(x) + C, \quad C \text{ is any constant.}$$

This is called the *indefinite integral* of $f(x)$; $f(x)$ is the *integrand*, x is the *variable of integration*, and C is the *constant of integration*. (By the previous proposition, this is the general form of an antiderivative of $f(x)$.)

Example 3.4. Evaluate the indefinite integral

$$\int x^3 + 3\,dx.$$

Solution. First we find a function $F(x)$ such that $F'(x) = x^3 + 4$. As in the previous example, we can see that the function

$$F(x) = \frac{1}{4}x^4 + \frac{4}{2}x^2$$

works. Thus, the indefinite integral is

$$\int x^3 + 3\,dx = \frac{1}{4}x^4 + 2x^2 + C, \quad C \text{ any constant.}$$

Note: **it is important to remember to write dx at the end of the integral**, for two reasons:

(i) The dx tells us where the integrand stops. If you don't write it, the meaning of what you've written is ambiguous. For example, if we wrote

$$\int x^3 - 3x^2 + 5,$$

it isn't clear if we mean

$$\begin{aligned}\int x^3 - 3x^2 + 5\,dx &= \frac{1}{4}x^4 - x^3 + 5x + C,\\ \int x^3 - 3x^2\,dx + 5 &= \frac{1}{4}x^4 - x^3 + C + 5, \quad \text{or}\\ \int x^3\,dx - 3x^2 + 5 &= \frac{1}{4}x^4 + C - 3x^2 + 5.\end{aligned}$$

(ii) The dx tells us the variable of integration (in this case, x). We are allowed to do integration with different variables of integration (and this will soon be important). While

$$\int 3x^2\,dx = x^3 + C,$$

we likewise have

$$\int 4t^2\,dt = t^3 + C.$$

In some cases, we might have multiple variables, but only one can be the variable of integration – the others are treated as constants. For example,

$$\int Kx^2\,dx = \frac{K}{3}x^3 + C,$$

whereas if we wrote $\int Kx^2\,dK$, we would have to treat x as constant, so the answer is

$$\int Kx^2\,dK = \frac{x^2}{2}K^2 + C.$$

Similarly,

$$\int x^n\,dx = (n+1)x^{n+1} + C,$$

while

$$\int x^n \, dn = \frac{x^n}{\ln(x)} + C.$$

(We usually try to avoid using n as the variable of integration, because usually n stands for an integer variable.)

Proposition 3.5 (Linearity of the integral).

$$\textit{(i)} \int K f(x)\, dx = K \int f(x)\, dx,$$

where K is any scalar (constant). This includes the case of $K < 0$, for example

$$\int -f(x)\, dx = -\int f(x)\, dx.$$

$$\textit{(ii)} \int f(x) + g(x)\, dx = \int f(x)\, dx + \int g(x)\, dx.$$

Proof. These both follow from linearity of differentiation: say

$$\int f(x)\, dx = F(x) + C,$$

which means that $F'(x) = f(x)$. Then:

(i) we know that $(KF)'(x) = KF'(x) = Kf(x)$, so that

$$\int f(x)\, dx = KF'(x) + D, \quad D \text{ any constant.}$$

(ii) Suppose likewise that $G'(x) = g(x)$, so that

$$\int g(x)\, dx = G(x) + E, \quad E \text{ any constant.}$$

Then $(F + G)'(x) = F'(x) + G'(x) = f(x) + g(x)$. Thus, (avoiding writing the constants of integration), we have

$$\int f(x) + g(x)\, dx = F(x) + G(x) = \int f(x)\, dx + \int g(x)\, dx.$$

□

By contrast, note that

$$\int f(x)g(x)\,dx \neq \int f(x)\,dx \int g(x)\,dx$$

and

$$\int \frac{f(x)}{g(x)}\,dx \neq \int f(x)\,dx \Big/ \int g(x)\,dx.$$

(Because analogous formulae don't hold for differentiation.) This is one of the things that makes integration interesting!

3.2 Computing integrals: the basics

It is very easy to differentiate – there are rules that work every time. Antidifferentiation (finding antiderivatives) is not easy; it is not always possible. For example, $f(x) = e^{-x^2}$ doesn't have a closed form antiderivative (we can't write down a formula for its antiderivative, except by making up new functions). Even when it is possible to find an antiderivative, **there are no rules that always work**.

Instead, we shall learn a number of techniques for finding antiderivatives. For any integration problem, we may need to try a few approaches, and often combine techniques, before discovering the solution.

We start with some integrals we just know by differentiating basic functions. For polynomials:

$$\int x^n\,dx = \frac{1}{n+1}x^{n+1} + C.$$

In fact, this works for any value of n (not just integers) **except** $n = -1$.

Recalling the derivatives of trigonometric functions, we have

$$\begin{aligned}
\int \cos(x)\,dx &= \sin(x) + C,\\
\int \sin(x)\,dx &= \cos(x) + C,\\
\int \sec(x)^2\,dx &= \tan(x) + C,\\
\int \sec(x)\tan(x)\,dx &= \sec(x) + C,\\
\int \csc(x)^2\,dx &= -\cot(x) + C,\\
\int \csc(x)\cot(x)\,dx &= -\csc(x) + C,\\
\int \frac{1}{x^2+1}\,dx &= \tan^{-1}(x) + C,\\
\int \frac{1}{\sqrt{1-x^2}}\,dx &= \sin^{-1}(x) + C.
\end{aligned}$$

Recalling the derivatives of exponential functions, we have

$$\begin{aligned}
\int e^x\,dx &= e^x + C,\\
\int a^x\,dx &= \frac{1}{\ln(a)}a^x + C,\\
\int \frac{1}{x}\,dx &= \int x^{-1}\,dx = \ln|x| + C.
\end{aligned}$$

(Note that, we don't have a formula for the antiderivative of $\ln(x)$ – this will require some later techniques.)

Example 3.6. Evaluate the following:

$$\text{(i)} \quad \int 7t^3 - 2t^{-4}\, dt$$

$$\text{(ii)} \quad \int 5\sqrt{x} - \frac{2}{\sqrt[3]{x^2}}\, dx$$

$$\text{(iii)} \quad \int \frac{4u^3 + 6}{u^4}\, du$$

$$\text{(iv)} \quad \int \sin(2x) + 4e^x\, dx$$

$$\text{(v)} \quad \int \cos(x+1) + \frac{5}{1+x^2} - \sqrt{x^7}, dx$$

Solution. (i):

$$\int 7t^3 - 2t^{-4}\, dt = \frac{7}{4}t^4 - \frac{2}{-3}t^{-3} + C = \frac{7}{4}t^4 + \frac{2}{3t^3} + C.$$

(ii):

$$\begin{aligned}\int 5\sqrt{x} - \frac{2}{\sqrt[3]{x^2}}\, dx &= \int 5x^{1/2} - 2x^{-2/3}\, dx \\ &= \frac{5}{3/2}x^{3/2} - \frac{2}{1/3}x^{1/3} + C \\ &= \frac{10}{3}x^{3/2} - 6x^{1/3} + C.\end{aligned}$$

(iii):

$$\begin{aligned}\int \frac{4u^3 + 6}{u^4}\, du &= \int 4u^{-1} + 6u^{-4}\, du \\ &= 4\ln|u| + \frac{6}{-3}u^{-3} + C \\ &= 4\ln|u| - 2u^{-3} + C.\end{aligned}$$

(iv):

$$\int \sin(2x) + 4e^x\, dx = -\frac{1}{2}\cos(2x) + 4e^x + C.$$

Here we didn't use the integrals derived above, but instead we remember that the derivative of $\cos(2x)$ is $-2\sin(2x)$.

(v):

$$\int \cos(x+1) + \frac{5}{1+x^2} - \sqrt{x^7}, dx = \sin(x+1) + 5\tan^{-1}(x) - \frac{2}{9}x^{9/2} + C.$$

Again, we compute that the $\frac{d}{dx}\sin(x+1) = \cos(x+1)$ in order to deal with the first term.

Example 3.7. Solve

$$\int \sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right)\, dx.$$

Solution. Although our later techniques will give us other options for dealing with this one, we can do this one by simplifying the integrand. Recall the double-angle formula

$$\sin(2t) = 2\sin(t)\cos(t),$$

and thus,

$$\sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right) = \frac{1}{2}\sin(x).$$

Therefore,

$$\begin{aligned}\int \sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right)\, dx &= \frac{1}{2}\int \sin(x)\,dx \\ &= -\frac{1}{2}\cos(x) + C.\end{aligned}$$

Even when we have other tools at our disposal, remember as in the previous example, to **try to simplify the integrand** as one possible way to solve the integral.

3.3 Substitution rule

In this section we learn the first technique for solving more complicated integrals.

Consider the integral

$$\int 12x^2\sqrt[5]{4x^3+7}\,dx.$$

At first this one may look very difficult; but, there is a trick. Notice that if we let

$$u = 4x^3 + 7,$$

then we compute

$$du = 12x^2\, dx$$

(by differentiating with respect to x), and the integral simplifies

$$\int 12x^2 \sqrt[5]{4x^3+7}\, dx = \int \sqrt[5]{4x^3+7}\,(12x^2\, dx) = \int \sqrt[5]{u}\, du.$$

Is this legitimate? Yes, it is justified by the **chain rule** for differentiation.

Proposition 3.8 (Substitution rule)**.** *Suppose that*

$$\int f(u)\, du = F(u) + C.$$

Then

$$\int f(g(x))\, g'(x)\, dx = F(g(x)) + C.$$

Proof. We compute, using the chain rule,

$$\frac{d}{dx} F(g(x)) = F'(g(x))\, g'(x) = f(g(x))\, g'(x).$$

Therefore,

$$\int f(g(x))\, g'(x)\, dx = F(g(x)) + C.$$

□

We write the substitution rule succinctly as

$$\int f(g(x))\, g'(x)\, dx = \int f(u)\, du, \quad \text{where } u = g(x).$$

The skill in using the substitution rule is to identify the right thing to substitute. We want to pick some function $u = g(x)$ so that, after dividing by $g'(x)$, the integrand can be expressed purely in terms of u. If we find that there are some x's left over, we have probably chosen the wrong function $g(x)$.

To learn this skill, **you simply must practice!**

Example 3.9. Evaluate the following integrals.

$$\text{(i)} \quad \int 12x^2 \sqrt[5]{4x^3+7}\, dx.$$

$$\text{(ii)} \quad \int \left(1 - \frac{2}{t}\right) \cos(t - 2\ln(t))\, dt.$$

$$\text{(iii)} \quad \int 7(6x-1)e^{3x^2-x}\, dx.$$

$$\text{(iv)} \quad \int \cos(y)(1-5\sin(y))^7\, dy.$$

$$\text{(v)} \quad \int \frac{t}{\sqrt[4]{1-2t^2}}\, dt.$$

$$\text{(vi)} \quad \int x^5 \frac{(2+\sqrt{1-x^6})^3}{\sqrt{1-x^6}}\, dx.$$

Solution. (i): This is the example we already started. We use

$$u = 4x^3 + 7, \quad du = 12x^2\, dx,$$

so that

$$\begin{aligned} \int 12x^2 \sqrt[5]{4x^3+7}\, dx &= \int \sqrt[5]{4x^3+7}\,(12x^2\, dx) \\ &= \int \sqrt[5]{u}\, du \\ &= \frac{5}{6}u^{6/5} + C \\ &= \frac{5}{6}(4x^3+7)^{6/5} + C. \end{aligned}$$

Always remember to get rid of the substitution variable at the last step. (I.e., express your final solution in terms of the original variable.)

(ii): Use

$$u = t - 2\ln(t), \quad du = 1 - \frac{2}{t}\, dt.$$

Then we have

$$\int \left(1-\frac{2}{t}\right)\cos(t-2\ln(t))\,dt = \int \cos(u)\,du$$
$$= \sin(u)+C$$
$$= \sin(t-2\ln(t))+C.$$

(iii): Use

$$u = 3x^2 - x, \quad du = 6x-1\,dx.$$

This gives

$$\int 7(6x-1)e^{3x^2-x}\,dx = \int 7e^u\,du$$
$$= 7e^u + C$$
$$= 7e^{3x^2-x}+C.$$

(iv): Use

$$u = 1-5\sin(y), \quad du = -5\cos(y)\,dy$$

to get

$$\int \cos(y)(1-5\sin(y))^7\,dy = \int u^7\,\frac{du}{-5}$$
$$= -\frac{1}{5\cdot 8}u^8 + C$$
$$= -\frac{1}{40}(1-5\sin(y))^8 + C.$$

(v): Use

$$u = 1-2t^2, \quad du = -4t\,dt.$$

This gives

$$\int \frac{t}{\sqrt[4]{1-2t^2}}\,dt = \int -\frac{1}{4}u^{-1/4}\,du$$
$$= -\frac{1}{4}\frac{4}{3}u^{3/4}+C$$
$$= -\frac{1}{3}\sqrt[4]{(1-2t^2)^3}+C.$$

(vi): Use

$$u = 1 - x^6, \quad du = -6x^5\, du$$

and we get

$$\int x^5 \frac{(2+\sqrt{1-x^6})^3}{\sqrt{1-x^6}}\, dx = -\int \frac{(2+\sqrt{u})^3}{6\sqrt{u}}\, du.$$

To solve this, we need to do another substitution:

$$v = 2 + \sqrt{u}, \quad dv = \frac{1}{2\sqrt{u}}\, du$$

to obtain

$$\begin{aligned} -\int \frac{(2+\sqrt{u})^3}{6\sqrt{u}}\, du &= -\int \frac{v^3}{3}\, dv \\ &= -\frac{1}{12}v^4 + C. \end{aligned}$$

Putting this together, we get

$$\begin{aligned} \int x^5 \frac{(2+\sqrt{1-x^6})^3}{\sqrt{1-x^6}}\, dx &= -\frac{1}{12}v^4 + C \\ &= -\frac{1}{12}(2+\sqrt{u})^4 + C \\ &= -\frac{1}{12}(2+\sqrt{1-x^6})^4 + C. \end{aligned}$$

After a long problem like this, it is always best to check your solution. Fortunately, this is easy to do, since it's just integration!

We have

$$\begin{aligned} \frac{d}{dx}\frac{1}{12}(2+\sqrt{1-x^6})^4 &= \frac{1}{12}4(2-\sqrt{1-x^6})^3 \frac{1}{2\sqrt{1-x^6}}6x^5 \\ &= x^5 \frac{2-\sqrt{1-x^6})^3}{\sqrt{1-x^6}}, \end{aligned}$$

which confirms that our answer is correct.

Note. If you are a clever clog, you might have seen how to do this problem with just one substitution,

$$w = 2 + \sqrt{1-x^6}, \quad dw = \frac{1}{2\sqrt{1-x^6}}6x^5\, dx = \frac{x^5}{3\sqrt{1-x^6}}\, dx.$$

Example 3.10. Evaluate the following.

$$\text{(i)} \quad \int \frac{7}{2x+3}\,dx.$$
$$\text{(ii)} \quad \int \frac{7x}{2x^2+3}\,dx.$$
$$\text{(iii)} \quad \int \frac{7x}{(2x^2+3)^2}\,dx.$$
$$\text{(iv)} \quad \int \frac{7}{2x^2+3}\,dx.$$

Solution. (i): For this, use

$$u = 2x+3, \quad du = 2\,dx$$

so that

$$\begin{aligned}\int \frac{7}{2x+3}\,dx &= \int \frac{7}{u}\,\frac{du}{2} \\ &= \frac{7}{2}\ln|u| + C \\ &= \frac{7}{2}\ln|2x+3| + C.\end{aligned}$$

(ii): Use

$$u = 2x^2+3, \quad du = 4x\,dx,$$

to get

$$\begin{aligned}\int \frac{7x}{2x^2+3}\,dx &= \int \frac{7}{4u}\,du \\ &= \frac{7}{4}\ln|u| + C \\ &= \frac{7}{4}\ln(2x^2+3) + C.\end{aligned}$$

(Note here, we dropped the absolute value sign because $2x^2+3$ is always positive. We don't have to do this – it would be acceptable to give the final answer as $\frac{7}{4}\ln|2x^2+3| + C$.)

(iii): Using the same substitution as in (ii), we obtain

$$\begin{aligned}\int \frac{7x}{(2x^2+3)^2}\,dx &= \int \frac{7}{4u^2}\,du \\ &= \int \frac{7}{4}u^{-2}\,du \\ &= \frac{7}{4}\cdot\frac{1}{-1}u^{-1}+C \\ &= -\frac{7}{4(2x^2+3)}+C.\end{aligned}$$

(iv): At first, it might look like we want to use the same substitution again, that is,

$$u = 2x^2+3, \quad du = 4x\,dx.$$

However, we don't have a term x, so we would get

$$\int \frac{7}{2x^2+3}\,dx = \int \frac{7}{4xu}\,du,$$

and this doesn't get rid of all occurrences of x, so we cannot proceed further. We **can't treat** x **as a constant**, since u is defined in terms of x, so we don't have

$$\int \frac{7}{4xu}\,du = \frac{7}{4x}\ln|u|.$$

We could solve for x in terms of u:

$$x = \pm\sqrt{\frac{u-3}{2}},$$

but the result is

$$\pm\int \frac{7}{4u\sqrt{(u-3)/2}}\,du,$$

which looks even more intimidating than the original integral. (Also we need to worry about the $\pm$, which is undesirable.)

Whenever we can't get rid of all occurrences of the old variable, it is a dead end; the substitution didn't work and we have to go back and try something else.

Here, we might instead recognise that the integrand

$$\frac{7}{2x^2+3}$$

looks something like the integrand

$$\frac{1}{x^2+1}$$

whose antiderivative is $\tan^{-1}(x)$.

We manipulate things to make it look more like the derivative of $\tan^{-1}(x)$:

$$\int \frac{7}{2x^2+3}\,dx = \frac{7}{3}\int \frac{1}{\frac{2}{3}x^2+1}\,dx = \frac{7}{3}\int \frac{1}{(\sqrt{2/3}x)^2+1}\,dx$$

and now we see that the correct substitution is

$$u = \sqrt{\frac{2}{3}}x, \quad du = \sqrt{\frac{2}{3}}\,dx.$$

Thus,

$$\begin{aligned}\int \frac{7}{2x^2+3}\,dx &= \frac{7}{3}\int \frac{1}{(\sqrt{2/3}x)^2+1}\,dx \\ &= \frac{7}{3}\int \frac{1}{u^2+1}\sqrt{\frac{3}{2}}\,du \\ &= \frac{7\sqrt{3}}{3\sqrt{2}}\tan^{-1}(u)+C \\ &= \frac{7\sqrt{3}}{3\sqrt{2}}\tan^{-1}\left(\sqrt{\frac{2}{3}}x\right)+C.\end{aligned}$$

Example 3.11. Solve:

$$\int \frac{t+1}{\sqrt{1-9t^2}}\,dt.$$

Solution. Whenever the integrand is a sum of two things, it is advisable to break it into the two parts using linearity:

$$\int \frac{t+1}{\sqrt{1-9t^2}}\,dt = \int \frac{1}{\sqrt{1-9t^2}}\,dt + \int \frac{t}{\sqrt{1-9t^2}}\,dt.$$

Now, we need to solve the two integrals separately:

$$\text{(i)} \int \frac{t}{\sqrt{1-9t^2}}\,dt \quad \text{and}$$
$$\text{(ii)} \int \frac{1}{\sqrt{1-9t^2}}\,dt.$$

(i) We do this by the substitution

$$u = 1 - 9t^2, \quad du = -18t\,dt.$$

This gives

$$\begin{aligned}\int \frac{t}{\sqrt{1-18t^2}}\,dt &= \int \frac{1}{-18\sqrt{u}}\,du \\ &= -\frac{2}{18}\sqrt{u} + C' \\ &= -\frac{1}{9}\sqrt{1-t^2} + C',\end{aligned}$$

C' any constant.

(ii): While it may seem like we want to again substitute

$$u = 1 - 9t^2, \quad du = -18t\,dt,$$

there isn't a "t" available for the du part. Instead, we recognise that the integrand looks similar to

$$\frac{1}{\sqrt{1-t^2}},$$

whose antiderivative is

$$\sin^{-1}(t).$$

We thus do a substitution to arrive at this exactly:

$$u = 3t, \quad du = 3\,dt.$$

This leads to

$$\begin{aligned}\int \frac{1}{\sqrt{1-9t^2}}\,dt &= \int \frac{1}{3\sqrt{1-u^2}}\,du \\ &= \frac{1}{3}\sin^{-1}(u) + C'' \\ &= \frac{1}{3}\sin^{-1}(3t) + C''.\end{aligned}$$

C'' any constant.

Putting this together, we get

$$\int \frac{t+1}{\sqrt{1-9t^2}}\,dt = -\frac{1}{9}\sqrt{1-9t^2} + \frac{1}{3}\sin^{-1}(3t) + C.$$

(We combined the two integration constants into one. It is best practice to use different symbols for different integration constants, but it probably won't cause much confusion if this isn't done.)

In the previous two examples, we saw that sometimes similar-looking integrands may require very different methods.

As mentioned before, we simply **cannot** solve every integral problem. One way to say this is that we just don't have names for all the functions we would need. However, we might make up new functions, and then try to solve other integrals in terms of these. For example, it is a fact that there is a differentiable function $G : \mathbb{R} \to \mathbb{R}$ which satisfies

$$G'(t) = \frac{\sin(t)}{t}$$

for all $t \neq 0$; however, this function cannot be expressed in terms of elementary functions (polynomials, trig, exponential, logarithm).

Example 3.12. Using the function G above to express the answer, solve

$$\int \sin(e^t)\, dt.$$

Solution. Use the substitution

$$u = e^t, \quad du = e^t\, dt.$$

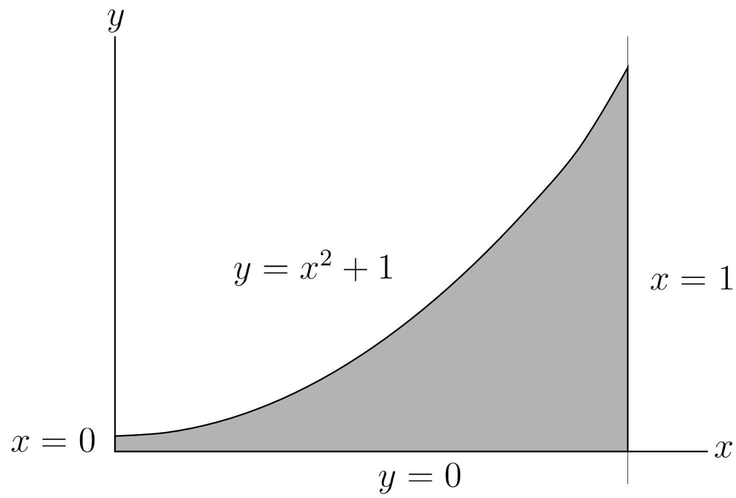
This gives

$$\begin{aligned}\int \sin(e^t)\, dt &= \int \frac{\sin(e^t)}{e^t} \frac{dt}{e^t} \\ &= \int \frac{\sin(u)}{u}\, du \\ &= G(u) + C \\ &= G(e^t).\end{aligned}$$

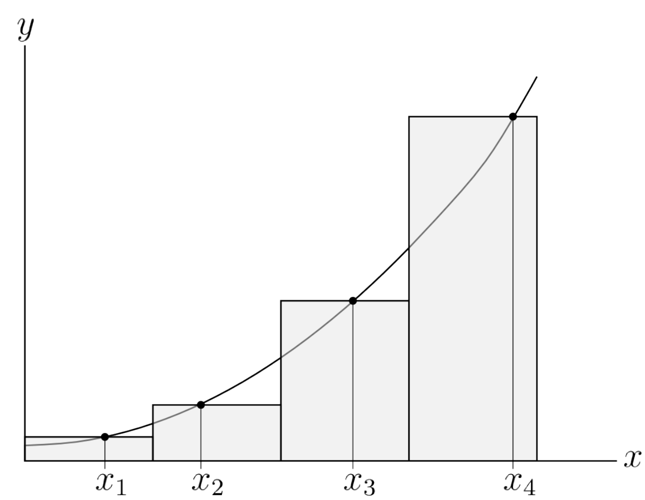
3.4 Area and the definite integral

The problem of computing area will tie into the main application of integrals. Suppose we want to compute the area of some region. To begin, we need to define the area precisely; for now, say it is the area bound by the x-axis, some curve $y = f(x)$, and two vertical lines $x = a$ and $x = b$.

Let us use, as an example, the area bound by the x-axis, $y = x^2 + 1$, $x = 0$, and $x = 1$.



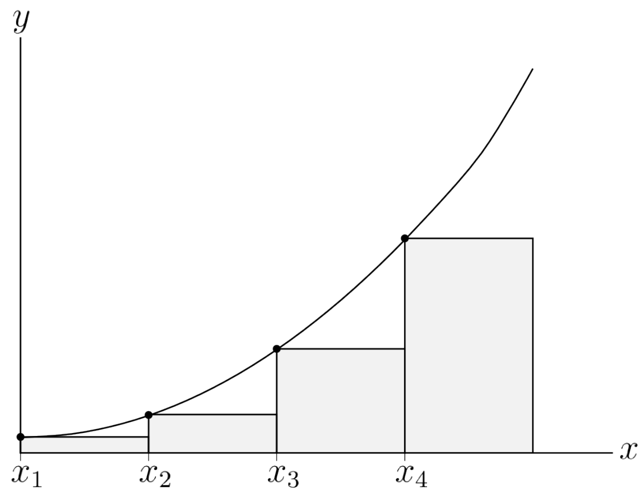
A first step towards (exactly) computing an area is finding a good way to estimate the error. We can partition the interval $[0, 1]$ into a number of subintervals; for now, let's partition it into the subintervals $[0, \frac{1}{4}]$, $[\frac{1}{4}, \frac{1}{2}]$, $[\frac{1}{2}, \frac{3}{4}]$, $[\frac{3}{4}, 1]$. We can then pick points $x_1 \in [0, \frac{1}{4}]$, $x_2 \in [\frac{1}{4}, \frac{1}{2}]$, and so on. Form the following rectangles:



The total area of these rectangles (which is easy to compute!) gives us an approximation of the area we are interested in. Of course, the total area of these rectangles depends on the choice of points $x_1, \ldots, x_5$. If we took left-hand endpoints, i.e., $x_i = \frac{i-1}{4}$, we get the estimate

$$\sum_{i=1}^{4} (\left(\tfrac{(i-1)}{4}\right)^2 + 1) \cdot \tfrac{1}{4} = \tfrac{1}{4}\left((0+1) + (\tfrac{1}{16}+1) + (\tfrac{1}{4}+1) + (\tfrac{9}{16}+1)\right) = \tfrac{39}{32} = 1.21875.$$

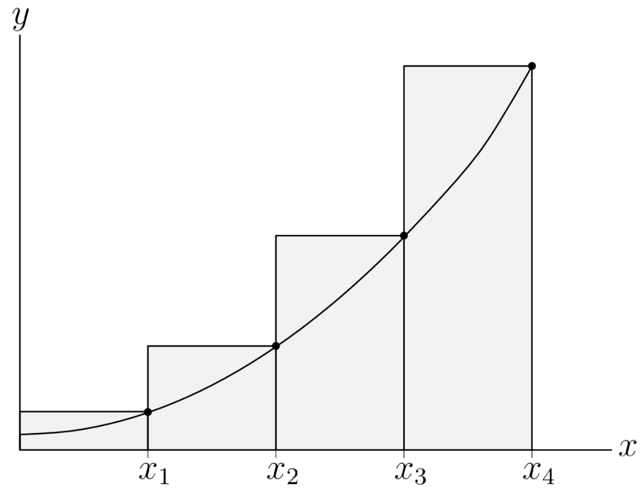
In this case, the rectangles are contained in the region we are interested in, so it is clear that this underestimates the correct area.



If we took right-hand endpoints, i.e., $x_i = \frac{i}{4}$, we get the estimate

$$\sum_{i=1}^{4} (\left(\tfrac{i}{4}\right)^2 + 1) \cdot \tfrac{1}{4} = \tfrac{1}{4}\left((\tfrac{1}{16}+1) + (\tfrac{1}{4}+1) + (\tfrac{9}{16}+1) + (1+1)\right) = \tfrac{47}{32} = 1.46875.$$

Since these rectangles (together) completely contain the region, this overestimates the correct area.



Altogether, this tells us that

$$\text{Area} \in [\tfrac{39}{32}, \tfrac{47}{32}].$$

The average of these two estimates is $\frac{43}{32} = 1.34375$, which is $\frac{1}{8}$ away from each of the over- and the under-estimation. Hence, this average estimate is accurate to within $\frac{1}{8}$.

To get a certain improvement, we need to take a finer partition. If we started with the partition $[0, \frac{1}{10}], [\frac{1}{10}, \frac{2}{10}], \ldots$, then we arrive at the estimates

$$\frac{257}{200} = 1.285, \quad \frac{277}{200} = 1.385.$$

If we started with the partition $[0, \frac{1}{100}], [\frac{1}{100}, \frac{2}{100}], \ldots$, then we arrive at the estimates

$$\frac{26567}{20000} = 1.32835, \quad \frac{26767}{20000} = 1.33835.$$

In this situation, the two estimates (coming from using left- and right-hand endpoints of the subintervals respectively) are always under- and over-estimates. This is because the function in question $y = x^2 + 1$, is increasing on the given interval. If we used a function that is decreasing, then using left-hand endpoints would instead give an overestimate, and right-hand endpoints would give an underestimate. For a general function (which is neither increasing nor decreasing), we **do not know whether the estimates are greater or less than the correct area**.

We have a name for the sums appearing when we compute the areas of the rectangles: these are called Riemann sums. More precisely, let $f : [a, b] \to \mathbb{R}$ be a function, and partition $[a, b]$ into n equally sized subintervals,

$$I_1 = \left[a, a + \frac{b-a}{n}\right],\ I_2 = \left[a + \frac{b-a}{n}, a + 2\frac{b-a}{n}\right], \ldots,$$
$$I_i = \left[\frac{n-i+1}{n}a + \frac{i-1}{n}b, \frac{n-i}{n}a + \frac{i}{n}b\right], \ldots,$$
$$I_n = \left[b - \frac{b-a}{n}, b\right].$$

Pick a point $x_i \in I_i$ for each $i = 1, \ldots, n$. Then the associated *Riemann sum* is

$$\sum_{i=1}^{n} f(x_i)\Delta x = f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x,$$

where $\delta x = \frac{b-a}{n}$ (the length of each interval in the partition).

Theorem 3.13. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. For each n, take a partition of $[a, b]$ into n equally sized subintervals $I_{1,n}, \ldots, I_{n,n}$ as above, and pick points $x_{i,n} \in I_{i,n}$. Then the Riemann sums converge to a limit, i.e.,*

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(x_{i,n})\tfrac{1}{n}$$

exists. This limit does not depend on the choice of points $x_{i,n} \in I_{i,n}$.

We will not prove this theorem in this course (for the proof, take MA2509, "Analysis II").

A word of caution. the last statement of the above theorem tells us that the *limit* doesn't depend on the choices of points. This does not mean that the individual Riemann sums don't depend on the choices of points (and as we've seen in the example, the Riemann sums **do** depend on these choices).

Since the limit in the above theorem exists, and doesn't depend on the choices of points, it is a well-defined value, and we call it the *definite integral* of $f(x)$ on $[a, b]$. In the setting of the above theorem, we use the notation

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_{i,n})\tfrac{1}{n}.$$

Note that although our example of measuring area concerned a positive function, the above theorem and the definition of the definite integral does not require this assumption; we allow $f(x) < 0$ to occur.

We extend the definition of the definite integral to the case that $a > b$. Here we ask that $f(x)$ is a continuous function on $[b, a]$, and we define

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

A note about notation. When we write

$$\int_a^b f(x)\,dx,$$

the variable x is a bound, dummy variable. This expression has the exact same meaning as

$$\int_a^b f(t)\,dt,$$

(or with any other variable in place of t). It does **not** make sense to ask if $\int_a^b f(x)\,dx$ depends on x, or to have x outside the integrand, like

$$\int_a^b f(x)\,dx + x \quad \text{or} \quad \int_a^x f(x)\,dx.$$

(By contrast, the variable x does make sense outside of an **indefinite** integral: it does make sense to write $\int f(x)\,dx + x$.)

Example 3.14. Let $f(x) = 5 - 3x$ on the interval $[0,1]$. For each n, we partition $[0,1]$ into n equally sized subintervals $I_{1,n}, \dots, I_{n,n}$, so that

$$I_{i,n} = [\tfrac{i-1}{n}, \tfrac{i}{n}].$$

Pick $x_{i,n} \in I_{i,n}$ as the right-hand endpoint, $x_i = \frac{i}{n}$. Then the Riemann sum is

$$\sum_{i=1}^{n} \left(5 - 3 \cdot \frac{i}{n}\right) \frac{1}{n} = 5 \cdot \frac{n}{n} - \frac{3}{n^2} \sum_{i=1}^{n} i = 5 - \frac{3}{n^2} \cdot \frac{n(n+1)}{2} = 5 - \frac{3(n+1)}{2n}.$$

Taking the limit, we get

$$\int_0^1 5 - 3x = \lim_{n\to\infty} 5 - \frac{3(n+1)}{2n} = 5 - \frac{3}{2} = \frac{7}{2}.$$

Proposition 3.15 (Basic properties of the definite integral)**.** *Let $f(x), g(x)$ be continuous functions, let $K \in \mathbb{R}$, and let $a, b, c \in \mathbb{R}$. Then:*

(i) $\displaystyle\int_a^a f(x)\,dx = 0.$

(ii) $\displaystyle\int_a^b Kf(x)\,dx = K\int_a^b f(x)\,dx.$

(iii) $\displaystyle\int_a^b f(x) + g(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$

(iv) $\displaystyle\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$

(v) $\displaystyle\int_a^b K\,dx = K(b-a).$

(vi) $\displaystyle\int_a^b f(x)\,dx \geq 0, \quad \text{if } f(x) \geq 0\ \forall x.$

(vii) $\displaystyle\int_a^b f(x)\,dx \leq \int_a^b g(x), \quad \text{if } f(x) \leq g(x)\ \forall x.$

(viii) $\displaystyle\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx.$

We will not prove these statements, although they are not too difficult (apart from (iv)), and make good exercises.

3.5 The Fundamental Theorem of Calculus

In general, it can be very tedious to compute a definite integral using Riemann sums. The following makes it much easier.

Theorem 3.16 (Fundamental Theorem of Calculus)**.** *Let* $f : [a, b] \to \mathbb{R}$ *be a continuous function.*

(i) For each $x \in [a, b]$*, define*

$$F(x) = \int_a^x f(t)\, dt.$$

Then

$$\int f(x)\, dx = F(x) + C$$

(i.e., $F(x)$ *is differentiable and* $F'(x) = f(x)$*).*

(ii) If $G(x)$ *is any antiderivative of* $f(x)$ *then*

$$\int_a^b f(x)\, dx = G(x)\Big|_{x=a}^{b},$$

where $G(x)\big|_{x=a}^{b}$ *means* $G(b) - G(a)$*.*

Again, we won't discuss the proof of this theorem, leaving it for MA2509 "Analysis II".

The purpose of this theorem is two-fold. Part (i) tells us that **every continuous function** has an antiderivative (even though we may not have a closed form expression). Part (ii) (which is more important in this course) tells us how to use indefinite integrals to compute definite integrals. Concisely, it says

$$\int_a^b f(x)\, dx = \left(\int f(x)\, dx\right)\Big|_{x=a}^{b}.$$

One interpretation of the integral is as a way of defining the average value of a continuous function: if $f(x)$ is a continuous function, then its average value on $[a, b]$ is

$$\frac{1}{b-a}\int_a^b f(x)\, dx. \tag{3.1}$$

Consider the case that the function $f(t)$ represents the speed of a car, at time t. If $p(t)$ represents the position of the car, then $p'(t) = f(t)$. Two ways of measuring the average speed between times a and b are: divide the total distance travelled by the total time:

$$\frac{p(b) - p(a)}{b - a},$$

or use the integral formula (3.1)

$$\frac{1}{b-a}\int_a^b f(t)\,dt.$$

Since $p'(t) = f(t)$, FTC tells us that these two ways of measuring the average are the same:

$$\frac{p(b) - p(a)}{b - a} = \frac{1}{b-a}\int_a^b p'(t)\,dt = \frac{1}{b-a}\int_a^b f(t)\,dt.$$

Example 3.17. Compute the area enclosed by the x-axis and the curve $y = 4 - x^2$.

Solution. Note that the curve $y = 4 - x^2$ intersects the x-axis at the points $(-2, 0)$ and $(2, 0)$. The area we need to compute is thereby given by the integral

$$\int_{-2}^{2} 4 - t^2\,dt.$$

We compute

$$\int 4 - t^2\,dt = 4t - \frac{t^3}{3} + C,$$

so that

$$\int_{-2}^{2} 4 - t^2\,dt = 4t - \frac{t^3}{3}\Big|_{t=-2}^{2} = 8 - \frac{8}{3} - (-8 + \frac{8}{3}) = \frac{32}{3}.$$

When computing a definite integral using FTC, we always drop the integration constant (i.e., the "$+\,C$"). (If we left it in, it would cancel with itself.)

Example 3.18. Differentiate the following functions.

$$\text{(i)} \quad f(x) = \int_5^x e^{7t^2} \sqrt{5 + \cos(t)^3}\, dt$$
$$\text{(ii)} \quad g(x) = \int_{x^2}^5 e^{7t^2} \sqrt{5 + \cos(t)^3}\, dt.$$

Solution. (i): It would be a mistake to try to compute this integral first! Rather, we can appeal directly to FTC (Theorem 3.16 (i)), which tells us that

$$f'(x) = e^{7x^2} \sqrt{5 + \cos(x)^3}.$$

(ii): Again, we don't want to try to compute the integral. However, we can't appeal immediately to FTC since our interval of integration is $[x^2, 5]$, rather than something of the form $[a, x]$. We first reverse the endpoints:

$$g(x) = \int_{x^2}^5 e^{7t^2} \sqrt{5 + \cos(t)^3}\, dt = -\int_5^{x^2} e^{7t^2} \sqrt{5 + \cos(t)^3}\, dt.$$

Next, we note that what we get is a function of x^2, namely

$$g(x) = -f(x^2),$$

where f is from part (i). We may therefore use the Chain Rule:

$$g'(x) = -f'(x^2)\frac{d}{dx}x^2 = -e^{7(x^2)^2} \sqrt{5 + \cos(x^2)^3} 2x.$$

Example 3.19. Evaluate the following

$$\text{(i)} \quad \int_0^8 2x - 7\sqrt[3]{x^4}\, dx.$$
$$\text{(ii)} \quad \int_0^{\pi/2} 2\cos(t)\, dt.$$
$$\text{(iii)} \quad \int_{-5}^5 5^x + x^5\, dx.$$
$$\text{(iv)} \quad \int_1^6 x + \frac{1}{x}\, dx.$$

Solution. (i): We have

$$\int 2x - 7\sqrt[3]{x^4}\,dx = \int 2x - 7x^{4/3}\,dx = \frac{2}{2}x^2 - \frac{7}{7/3}x^{7/3} = x^2 - 3x^{7/3}.$$

Hence,

$$\begin{aligned}\int_0^8 2x - 7\sqrt[3]{x^4}\,dx &= x^2 - 3x^{7/3}\Big|_{x=0}^{8} \\ &= 8^2 - 3\cdot 8^{7/3} - (0^2 - 3\cdot 0^{7/3}) \\ &= 64 - 3\cdot 128 = -320.\end{aligned}$$

(ii): From now on, when the indefinite integral isn't too complicated, we won't do it separately.

$$\begin{aligned}\int_0^{\pi/2} 2\cos(t)\,dt &= 2\sin(t)\big|_{t=0}^{\pi/2} \\ &= 2\sin(\pi/2) - 2\sin(0) \\ &= 2\cdot 1 - 2\cdot 0 = 2.\end{aligned}$$

(iii): We have

$$\begin{aligned}\int_{-5}^{5} 5^x + x^5\,dx &= \frac{5^x}{\ln(5)} + \frac{x^6}{6}\Big|_{x=-5}^{5} \\ &= \frac{5^5}{\ln(5)} + \frac{5^6}{6} - \frac{5^{-5}}{\ln(5)} - \frac{(-5)^6}{6} \\ &= \frac{1}{\ln(5)}\left(5^5 - \frac{1}{5^5}\right).\end{aligned}$$

(iv): We have

$$\begin{aligned}\int_1^6 x + \frac{1}{x}\,dx &= \frac{x^2}{2} + \ln(x)\Big|_{x=1}^{6} \\ &= \frac{6^2}{2} + \ln(6) - \frac{1^2}{2} - \ln(1) \\ &= \frac{36}{2} + \ln(6) - \frac{1}{2} - 0 = \frac{35}{2} + \ln(6).\end{aligned}$$

Example 3.20. Compute

$$\text{(i)} \quad \int_1^2 x^2\sqrt{8+x^3}\,dx.$$
$$\text{(ii)} \quad \int_0^{\sqrt{\pi}} t\cos\left(\tfrac{\pi}{3}-t^2\right)\,dt.$$
$$\text{(iii)} \quad \int_2^4 \frac{1+t}{1+t^2}\,dt.$$

Solution. (i): Here the indefinite integral is more complicated, so it is best to solve it first, then plug the answer into the FTC formula for the definite integral. To solve

$$\int x^2\sqrt{8+x^3}\,dx,$$

use the substitution

$$u = 8 + x^3, \quad du = 3x^2\,dx.$$

Thus we have

$$\begin{aligned}\int x^2\sqrt{7+x^3}\,dx &= \int \sqrt{u}\,\frac{du}{3}\\ &= \frac{2}{9}u^{3/2} + C\\ &= \frac{2}{9}(8+x^3)^{3/2} + C.\end{aligned}$$

Therefore,

$$\begin{aligned}\int_1^2 x^2\sqrt{8+x^3}\,dx &= \frac{2}{9}(8+x^3)^{3/2}\Big|_{x=1}^{2}\\ &= \frac{2}{9}16^{3/2} - \frac{2}{9}9^{3/2}\\ &= \frac{2}{9}(64-27) = \frac{74}{9}.\end{aligned}$$

It is crucial that we finished up the indefinite integration with an expression in terms of x and not the substituted variable u. It is not correct that

$$\int_1^2 x^2\sqrt{8+x^3}\,dx = \frac{2}{9}u^{3/2}\Big|_{u=1}^{2}.$$

(ii): Using the substitution

$$u = \tfrac{\pi}{3} - t^2, \quad du = -2t\,dt,$$

we have

$$\begin{aligned} \int t\cos\left(\tfrac{\pi}{3} - t^2\right)\,dt &= \int \cos(u)\,\frac{du}{-2} \\ &= -\frac{1}{2}\sin(u) + C \\ &= -\frac{1}{2}\sin(\tfrac{\pi}{3} - t^2) + C. \end{aligned}$$

Therefore,

$$\begin{aligned} \int_0^{\sqrt{\pi}} t\cos\left(\tfrac{\pi}{3} - t^2\right)\,dt &= -\frac{1}{2}\sin(\tfrac{\pi}{3} - t^2)\Big|_{t=0}^{\sqrt{\pi}} \\ &= -\frac{1}{2}(\sin(\tfrac{\pi}{3} - \pi) - \sin(\tfrac{\pi}{3})) \\ &= -\frac{1}{2}(\sin(-\tfrac{2\pi}{3}) - \sin(\tfrac{\pi}{3})) \\ &= -\frac{1}{2}(-\frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2}) = \frac{\sqrt{3}}{2}. \end{aligned}$$

(iii): We split the indefinite integral into two parts, and use the substititution $u = 1 + t^2$ on the second part:

$$\begin{aligned} \int \frac{1+t}{1+t^2}\,dt &= \int \frac{1}{1+t^2}\,dt + \int \frac{t}{1+t^2}\,dt \\ &= \tan^{-1}(t) + \int \frac{1}{2u}\,du \\ &= \tan^{-1}(t) + \frac{\ln(u)}{2} + C \\ &= \tan^{-1}(t) + \frac{\ln(1+t^2)}{2} + C. \end{aligned}$$

Hence,

$$\begin{aligned} \int_2^4 \frac{1+t}{1+t^2}\,dt &= \tan^{-1}(t) + \frac{\ln(1+t^2)}{2}\Big|_{t=2}^{4} \\ &= \tan^{-1}(4) + \frac{\ln(17)}{2} - \tan^{-1}(2) - \frac{\ln(5)}{2}. \end{aligned}$$

Example 3.21. There is a bounded region enclosed by the curves

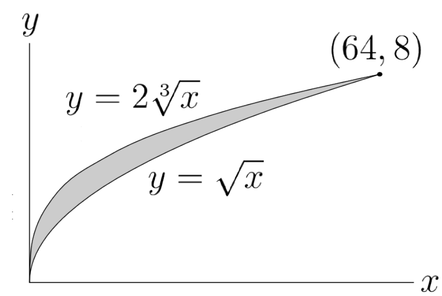
$$y = 2\sqrt[3]{x}$$

and

$$y = \sqrt{x}.$$

Find its area.

Solution. Let's first determine where these curves intersect.



They intersect when

$$2\sqrt[3]{x} = y = \sqrt{x},$$

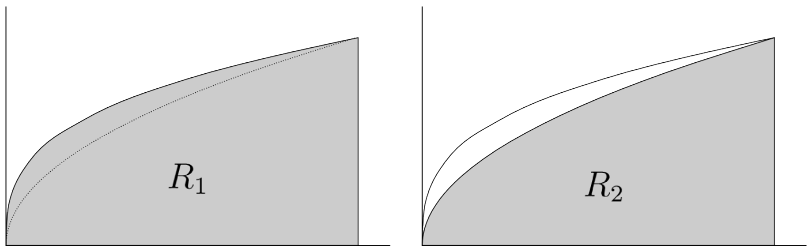
giving

$$\begin{aligned} 64x^2 &= x^3 \\ &= x^2(x - 64), \end{aligned}$$

and the solutions are $(x, y) = (0, 0)$ and $(x, y) = (64, 8)$.

Thus far, we've learned how to find the area between a region bound by the x-axis and a curve. The area we need to find for this problem can be viewed as the difference of two areas that we already know. Namely, if R_1 is the region enclosed by $y = 2\sqrt[3]{x}$, $x = 0$, $y = 0$, and $x = 4$, and R_2 is the region enclosed by $y = \sqrt{x}$, $x = 0$, $y = 0$, and $x = 4$, then the area we need to find is

$$\text{Area}(R_1) - \text{Area}(R_2).$$



We now compute this,

$$\begin{aligned}\text{Area}(R_1) - \text{Area}(R_2) &= \int_0^{64} 2\sqrt[3]{x}\,dx - \int_0^{64} \sqrt{x}\,dx\\ &= 2\tfrac{3}{4}x^{4/3} - \tfrac{2}{3}x^{3/2}\Big|_{x=0}^{64}\\ &= \tfrac{3}{2}(256-0) - \tfrac{2}{3}(512-0)\\ &= \frac{1152-1024}{3}\\ &= \frac{128}{3}.\end{aligned}$$

More generally, for functions $f(x), g(x)$ with $f(x) \leq g(x)$, the area of a region enclosed by $y = f(x)$, $y = g(x)$, $x = a$, and $x = b$ is computed by the integral

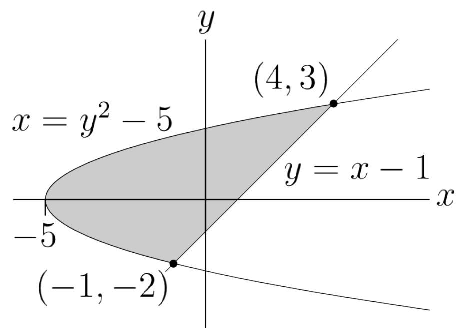
$$\int_a^b g(x) - f(x)\,dx.$$

Example 3.22. Find the area of the region enclosed by the curves $x = y^2 - 5$ and $y = x - 1$.

Solution. This problems differs a bit from previous ones, because the first curve is not in the form $y = f(x)$. Let's first compute where the curves intersect, by solving the two equations as a system. Substituting $y = x - 1$ into the first equation gives

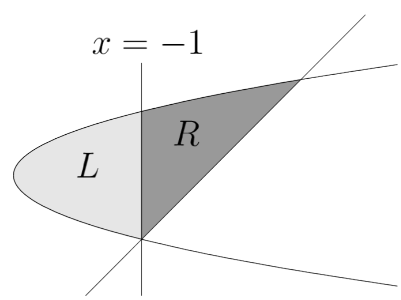
$$\begin{aligned}x &= (x-1)^2 - 5\\ &= x^2 - 2x - 4,\\ 0 &= x^2 - 3x - 4\\ &= (x-4)(x+1).\end{aligned}$$

The intersection points are therefore $(4,3)$ and $(-1,-2)$.



There are two ways that we can attack this problem.

(a): We divide the region by the line $x=-1$. On the left-hand side (L in the diagram below), we want to find the area between $y=-\sqrt{x+5}$ and $y=\sqrt{x+5}$ (from $x=-5$ to $x=-1$). On the right-hand side (R in the diagram below), we want to find the area between $y=x-1$ and $y=\sqrt{x+5}$ (from $x=-1$ to $x=4$).



Altogether, the area can be computed as

$$\begin{aligned}
\text{Area} &= \int_{-5}^{-1} \sqrt{x+5} - (-\sqrt{x+5})\,dx + \int_{-1}^{4} \sqrt{x+5} - (x-1)\,dx \\
&= 2\cdot\frac{2}{3}(x+5)^{3/2}\bigg|_{x=-5}^{-1} + \frac{2}{3}(x+5)^{3/2} - \frac{x^2}{2} + x\bigg|_{x=-1}^{4} \\
&= \frac{4}{3}(8-0) + \frac{2}{3}(27-8) - \frac{1}{2}(16-1) + (4-(-1)) \\
&= \frac{125}{6}
\end{aligned}$$

(b): Looking at this sideways, we see that an easier approach is to swap the roles of the x and y coordinates. Both curves can be put into the form

$x = f(y)$; for the second one, it is $x = y + 1$. The area is thus given by

$$\begin{aligned}
\text{Area} &= \int_{-2}^{3} y + 1 - (y^2 - 5)\, dy \\
&= -\frac{y^3}{3} + \frac{y^2}{2} + 6y\Big|_{y=-2}^{3} \\
&= -\frac{1}{3}(27 - (-8)) + \frac{1}{2}(9 - 4) + 6(3 - (-2)) \\
&= \frac{125}{6}.
\end{aligned}$$

3.6 Integration by parts

Integration by Parts (IBP) is a new technique, which will enable us to solve integrals such as

$$\int xe^{6x}\, dx,$$

which cannot be solved by substitution.

Proposition 3.23 (Integration by parts)**.** *Let $f(x), g(x)$ be differentiable functions, and suppose that*

$$\int f(x)g'(x)\, dx = H(x) + C.$$

Then

$$\int f'(x)g(x)\, dx = f(x)g(x) - H(x) + C'.$$

Proof. This is derived from the product rule for differentiation. By assumption, we know that

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

Define $F(x) = f(x)g(x) - H(x)$, so that

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

i.e.,

$$\int f'(x)g(x)\, dx = F(x) + C',$$

as required. □

Here is a more concise way of remembering Integration by Parts. Use

$$\begin{aligned} u &= f(x), & dv &= g'(x)\,dx \\ du &= f'(x),\,dx & v &= g(x) \end{aligned}$$

and then IBP becomes

$$\int u\,dv = uv - \int v\,du.$$

Since a new integration constant will occur for the integral $\int v\,du$, we have dropped the integration constant ("$+\,C$").

Example 3.24. Solve

$$\int xe^{6x}\,dx.$$

Solution. Use

$$\begin{aligned} u &= x, & dv &= e^{6x}\,dx \\ du &= dx, & v &= \frac{1}{6}e^{6x} \end{aligned}$$

to get

$$\begin{aligned} \int xe^{6x} &= \int u\,dv \\ &= uv - \int v\,du \qquad \text{(IBP)} \\ &= \frac{1}{6}\left(xe^{6x} - \int e^{6x}\,dx\right) \\ &= \frac{1}{6}xe^{6x} - \frac{1}{36}e^{6x} + C \end{aligned}$$

Check:

$$\frac{d}{dx}(\frac{1}{6}xe^{6x} - \frac{1}{36}e^{6x} + C) = \frac{1}{6}(e^{6x} + 6xe^{6x}) - \frac{1}{36}6e^{6x} = xe^{6x}.$$

How did we choose u and dv in the above example?
Generally, we want to factor the integrand as $u\,dv$ where:

1. we can find the antiderivative to dv (i.e., find v), and
2. the integral $\int v\,du$, is **easier** to solve.

(While this is generally what we want to do, there is one trick where 2 doesn't hold – $\int v\,du$ isn't any simpler; see Example 3.31.)

In the above example, it helped to remember that we knew how to integrate e^x (and, by substition, also e^{6x}). Through **practice** with computing integrals, we get to know which things we can integrate more easily. It might be that we need to use substitution, or even IBP again, to find v or solve $\int v\,du$.

Example 3.25. Evaluate the integral

$$\int x\sqrt{x+1}\,dx$$

Solution. There are, in fact, two ways of doing this one: (a) Integration by Parts, and (b) Substitution.

(a) Integration by Parts.

Notice that there are no trigonometric or exponential functions here. Although often the IBP integrals we will be solving contain trig or exponential functions, **don't assume** that you can't use IBP if you don't see them. We'll use IBP with

$$\begin{aligned} u &= x, \quad & dv &= \sqrt{x+1}\,dx \\ du &= dx, \quad & v &= \frac{2}{3}(x+1)^{3/2}. \end{aligned}$$

Then

$$\begin{aligned} \int x\sqrt{x+1}\,dx &= \frac{2}{3}x(x+1)^{3/2} - \frac{2}{3}\int (x+1)^{3/2}\,dx \qquad \text{(IBP)} \\ &= \frac{2}{3}x(x+1)^{3/2} - \frac{2}{3}\cdot\frac{2}{5}(x+1)^{5/2} + C \\ &= \frac{2}{3}x(x+1)^{3/2} - \frac{4}{15}(x+1)^{5/2} + C \end{aligned}$$

(b) Substitution.

We use the following substitution

$$u = x + 1, \quad x = u - 1, \quad dx = du.$$

Then

$$\begin{aligned}
\int x\sqrt{x+1}\,dx &= \int (u-1)u^{1/2}\,du \\
&= \int u^{3/2} - u^{1/2}\,du \\
&= \frac{5}{2}u^{5/2} - \frac{3}{2}u^{3/2} + C' \\
&= \frac{5}{2}(x+1)^{5/2} - \frac{3}{2}(x+1)^{3/2} + C'
\end{aligned}$$

We got different answers! Did we do something wrong? No, in fact, with some algebraic manipulation, we can show that

$$\frac{5}{2}(x+1)^{5/2} - \frac{3}{2}(x+1)^{3/2} = \frac{2}{3}x(x+1)^{3/2} - \frac{4}{15}(x+1)^{5/2}.$$

(Note, in some cases, there might be two correct solutions that aren't exactly the same, but instead differ by a constant. This is also okay — it is precisely the purpose of the integration constant.)

Example 3.26. Solve

$$\int x^2 \sin(10x)\,dx$$

Solution. We use IBP with

$$\begin{aligned}
u = x^2, &\quad dv = \sin(10x)\,dx \\
du = 2x\,dx, &\quad v = -\frac{1}{10}\cos(10x).
\end{aligned}$$

Then

$$\int x^2 \sin(10x)\,dx = -\frac{1}{10}\left(x^2\cos(10x) - 2\int x\cos(10x)\,dx\right) \qquad \text{(IBP)}$$

At first, this might not seem helpful, since the new integral, $\int x\cos(10x)\,dx$, is still not something we recognise. We need to do IBP again, this time with

$$u = x, \quad dv = \cos(10x)\,dx$$
$$du = dx, \quad v = \frac{1}{10}\sin(10x).$$

Then

$$\begin{aligned}\int x\cos(10x)\,dx &= \frac{1}{10}\left(x\sin(10x) - \int \sin(10x)\,dx\right) \qquad \text{(IBP)}\\ &= \frac{1}{10}x\sin(10x) + \frac{1}{100}\cos(10x) + C\end{aligned}$$

and thus,

$$\int x^2\sin(10x)\,dx = -\frac{1}{10}x^2\cos(10x) + \frac{1}{50}x\sin(10x) + \frac{1}{500}\cos(10x) + C'$$

Example 3.27. Evaluate

$$\int \ln(x)\,dx.$$

Solution. Unlike in previous examples, the integrand doesn't look like it can factor at all! One choice that you might be tempted to try is

$$u = 1, \quad dv = \ln(x)\,dx.$$

However, this begs the question, because we don't know how to find v (i.e., antidifferentiate $\ln(x)$) in the first place!

Instead, we use

$$u = \ln(x), \quad dv = dx$$
$$du = \frac{1}{x}\,dx, \quad v = x.$$

Then

$$\begin{aligned}\int \ln(x)\,dx &= x\ln(x) - \int x\frac{1}{x}\,dx \qquad \text{(IBP)}\\ &= x\ln(x) - \int 1\,dx\\ &= x\ln(x) - x + C.\end{aligned}$$

Example 3.28. Solve
$$\int \ln(t)^2 \, dt.$$

Solution. From the previous example, we know an antiderivative to $\ln(t)$. We can use

$$u = \ln(t), \quad dv = \ln(t)\,dt$$
$$du = \frac{dt}{t}, \quad v = t\ln(t) - t,$$

to get

$$\begin{aligned}\int \ln(t)^2 \, dt &= \ln(t)(t\ln(t) - t) - \int \frac{t\ln(t) - t}{t} \, dt \quad \text{(IBP)} \\ &= t\ln(t)^2 - t\ln(t) - \int \ln(t) - 1 \, dt \\ &= t\ln(t)^2 - t\ln(t) - (t\ln(t) - t - t) + C \\ &= t\ln(t)^2 - 2t\ln(t) + 2t + C.\end{aligned}$$

This one could also have been done with $u = \ln(t)^2, dv = 1\,dt$.

Example 3.29. Solve the integral

$$\int x^5\sqrt{x^3+1}\,dx$$

Solution. The most obvious way to factorise the integrand is

$$u = x^5, \quad dv = \sqrt{x^3+1}\,dx.$$

To find v, however, requires solving

$$\int \sqrt{x^3+1}\,dx$$

which is not ~~easy~~ possible!

But this is not the only choice! There are many others, e.g., $x \cdot x^4\sqrt{x^3+1}$, $x^2 \cdot x^3\sqrt{x^3+1}$, etc. The best way to do this one is by choosing $dv = x^2\sqrt{x^3+1}$. This is because we can do a substitution to solve

$$\int x^2\sqrt{x^3+1}\,dx,$$

namely $w = x^3+1$, $dw = 3x^2\,dx$, which gives

$$\int x^2\sqrt{x^3+1}\,dx = \frac{1}{3}\int \sqrt{w}\,dw = \frac{2}{9}(x^3+1)^{3/2} + C.$$

Thus, we use

$$\begin{aligned} u &= x^3, \quad dv = x^2\sqrt{x^3+1}, \\ du &= 3x^2, \quad v = \frac{2}{9}(x^3+1)^{3/2} \end{aligned}$$

and get

$$\begin{aligned} \int x^5\sqrt{x^3+1}\,dx &= \frac{2}{9}x^3(x^3+1)^{3/2} - \frac{2}{3}\int x^2(x^3+1)^{3/2}\,dx \qquad \text{(IBP)} \\ &= \frac{2}{9}x^3(x^3+1)^{3/2} - \frac{2}{9}\int w^{3/2}\,dw \qquad \text{(subst. } w = x^3+1\text{)} \\ &= \frac{2}{9}x^3(x^3+1)^{3/2} - \frac{4}{45}w^{5/2} + C \\ &= \frac{2}{9}x^3(x^3+1)^{3/2} - \frac{4}{45}(x^3+1)^{5/2} + C. \end{aligned}$$

Example 3.30. Evaluate

$$\text{(i)} \quad \int x^3 e^{x/2}\, dx.$$
$$\text{(ii)} \quad \int \frac{\ln(t)^2}{t^2}\, dt.$$

Solution. (i): Use IBP with

$$u = x^3, \quad dv = e^{x/2}\, dx$$
$$du = 3x^2\, dx, \quad v = 2e^{x/2}$$

to get

$$\int x^3 e^{x/2}\, dx = 2x^3 e^{x/2} - 6 \int x^2 e^{x/2}\, dx.$$

Does this help? Well the new integral isn't something we can solve directly; but it looks less complicated (because it contains x^2 instead of x^3).

We simply need to use IBP again (and again). This time use

$$u = x^2, \quad dv = e^{x/2}\, dx$$
$$du = 2x\, dx, \quad v = 2e^{x/2},$$

which gives

$$\int x^2 e^{x/2}\, dx = 2x^2 e^{x/2} - 4 \int x e^{x/2}\, dx \quad \text{(IBP)}.$$

We use IBP again; this time,

$$u = x, \quad dv = e^{x/2}\, dx$$
$$du = 1\, dx, \quad v = 2e^{x/2},$$

which gives

$$\begin{aligned} \int x e^{x/2}\, dx &= 2x e^{x/2} - 2 \int e^{x/2}\, dx \quad \text{(IBP)} \\ &= 2x e^{x/2} - 4e^{x/2} + C. \end{aligned}$$

Altogether, we get

$$\begin{aligned}\int x^3 e^{x/2}\,dx &= 2x^3e^{x/2} - 6[2x^2e^{x/2} - 4\left(2xe^{x/2} - 4e^{x/2} + C\right)]\\ &= 2e^{x/2}(x^3 - 6x^2 + 24x - 48) + C'.\end{aligned}$$

(ii): Use

$$\begin{aligned}u = \ln(t)^2, &\quad dv = \frac{1}{t^2}\,dt\\ du = \frac{2\ln(t)}{t}\,dt, &\quad v = -\frac{1}{t}.\end{aligned}$$

Then IBP gives

$$\int \frac{\ln(t)^2}{t^2}\,dt = -\frac{\ln(t)^2}{t} + 2\int \frac{\ln(t)}{t^2}\,dt \quad \text{(IBP)}.$$

The situation is much like in (i): we arrive at a new integral that we can't solve directly, but it seems that we are getting closer to something that we can solve directly, and so we push on.

We now use

$$\begin{aligned}u = \ln(t), &\quad dv = \frac{1}{t^2}\,dt\\ du = \frac{1}{t}\,dt, &\quad v = -\frac{1}{t}\end{aligned}$$

to get

$$\begin{aligned}\int \frac{\ln(t)}{t^2}\,dt &= -\frac{\ln(t)}{t} + \int \frac{1}{t^2}\,dt \quad \text{(IBP)}\\ &= -\frac{\ln(t)}{t} - \frac{1}{t} + C.\end{aligned}$$

Putting these together we obtain

$$\begin{aligned}\int \frac{\ln(t)^2}{t^2}\,dt &= -\frac{\ln(t)^2}{t} + 2\int \frac{\ln(t)}{t^2}\,dt\\ &= -\frac{\ln(t)^2}{t} - \frac{2\ln(t)}{t} - \frac{2}{t} + C'.\end{aligned}$$

In the previous example, it was clear that we should continue solving the integral using IBP, since each time the new integral appearing became easier. In the next example, this will not be the case; it involves a trick.

Example 3.31. Evaluate
$$\int e^t \cos(t)\, dt.$$

Solution. Try
$$u = e^t, \quad dv = \cos(t)\, dt,$$
$$du = e^t\, dt, \quad v = \sin(t),$$
and we find
$$\int e^t \cos(t)\, dt = e^t \sin(t) - \int e^t \sin(t)\, dt \qquad \text{(IBP)}. \tag{3.2}$$
The new integral, $\int e^t \sin(t)\, dt$, looks no easier than the original!

We might have instead tried
$$u = \cos(t), \quad dv = e^t\, dt,$$
$$du = -\sin(t)\, dt, \quad v = e^t,$$
which gives
$$\int e^t \cos(t)\, dt = e^t \cos(t) + \int e^t \sin(t)\, dt \qquad \text{(IBP)}. \tag{3.3}$$
This still leads to something involving the integral $\int e^t \sin(t)\, dt$. It seems that there is no way to get around something involving this other integral.

So let's perservere and try
$$u = \sin(t), \quad dv = e^t\, dt,$$
$$du = \cos(t)\, dt, \quad v = e^t,$$
which gives
$$\int e^t \sin(t)\, dt = e^t \sin(t) - \int e^t \cos(t)\, dt \qquad \text{(IBP)},$$

which expresses the new integral $\int e^t \sin(t)\,dt$ in terms of the original integral. Does this help? Well, combining this with either (3.2) or (3.3) leads to an equation involving only the original integral. With (3.2), we get

$$\begin{aligned}\int e^t \cos(t)\,dt &= e^t \sin(t) - \int e^t \sin(t)\,dt \\ &= e^t \sin(t) - (e^t \sin(t) - \int e^t \cos(t)\,dt) \\ &= \int e^t \cos(t)\,dt.\end{aligned}$$

This is not helpful at all — what has happened was that second IBP undid the first IBP.

However, with (3.3), we get

$$\begin{aligned}\int e^t \cos(t)\,dt &= e^t \cos(t) + \int e^t \sin(t)\,dt \\ &= e^t \cos(t) + e^t \sin(t) - \int e^t \cos(t)\,dt\end{aligned}$$

and rearranging, this becomes

$$2\int e^t \cos(t) = e^t \cos(t) + e^t \sin(t),$$

or

$$\int e^t \cos(t) = \frac{1}{2} e^t (\cos(t) + \sin(t)).$$

Note that there should be a constant of integration, so we add it, to get

$$\int e^t \cos(t) = \frac{1}{2} e^t (\cos(t) + \sin(t)) + C.$$

(There should **always** be a constant of integration when solving an indefinite integral; the only reason it wasn't there in our original solution is that we've established a habit of not adding one every time we do IBP, because usually it appears later when we get to an integral we can solve.)

Example 3.32. Solve the following.

$$\text{(i)} \quad \int \cos(\sqrt{1-y})\,dy.$$
$$\text{(ii)} \quad \int e^x \sin^{-1}(e^x)\,dx.$$

Solution. (i): Initially it looks like there isn't much that can be done here. In fact, we want to start with the substitution

$$w = \sqrt{1-y}, \quad dw = -\frac{1}{2\sqrt{1-y}}\,dy = -\frac{1}{2w}\,dy$$

(Note that we could have equally done this as

$$w^2 = 1-y, \quad 2w\,dw = -1\,dy.)$$

This gives

$$\int \cos(\sqrt{1-y})\,dy = \int -2w\cos(w)\,dw,$$

and this looks like something we can handle using IBP (we've done something very similar in Example 3.26).

We set

$$u = w, \quad dv = \cos(w)\,dw$$
$$du = dw, \quad v = \sin(w)$$

and use this with IBP, to get

$$\begin{aligned}\int \cos(\sqrt{1-y})\,dy &= -2\int w\cos(w)\,dw \\ &= -2\left(w\sin(w) - \int \sin(w)\,dw\right) \quad \text{(IBP)} \\ &= -2\,(w\sin(w) + \cos(w)) \\ &= -2(\sqrt{1-y}\sin(\sqrt{1-y}) + \cos(\sqrt{1-y})).\end{aligned}$$

(ii): Again, we start with a substitution:

$$w = e^x, \quad dw = e^x\,dx,$$

which gives

$$\int e^x \sin^{-1}(e^x)\,dx = \int \sin^{-1}(w)\,dw.$$

Now we do IBP with

$$\begin{aligned} u &= \sin^{-1}(w), \quad dv = 1\,dw \\ du &= \frac{1}{\sqrt{1-w^2}}\,dw, \quad v = w. \end{aligned}$$

In the following, we will find we need to do another substitution:

$$\begin{aligned} \int e^x \sin^{-1}(e^x)\,dx &= \int \sin^{-1}(w)\,dw \\ &= w\sin^{-1}(w) - \int \frac{w}{\sqrt{1-w^2}}\,dw \quad \text{(IBP)} \\ &= w\sin^{-1}(w) - \int -\frac{1}{2\sqrt{z}}\,dz \quad (\text{subst. } z = 1-w^2) \\ &= w\sin^{-1}(w) + 2\cdot\frac{1}{2}\sqrt{z} + C \\ &= w\sin^{-1}(w) + \sqrt{1-w^2} + C \\ &= e^x\sin^{-1}(e^x) + \sqrt{1-e^{2x}} + C. \end{aligned}$$

Summary of IBP techniques

(a) **IBP to reduce a power of the integration variable**. If the integrand factors as $x^k g(x)$, and we know how to integrate $g(x)$, then we set $u = x^k$ and $v = g(x)$ (Examples 3.24, 3.25). When $k > 1$, we will probably need to do this again (and again ...) (Examples 3.26, 3.30 (i)).

(b) **IBP with $u \neq x^k$**. Sometimes we don't have a factor of x^k, or we do but we can't integrate the other factor. Then we need to try something else. Don't forget to try $dv = 1\,dx$. (Examples 3.27, 3.28, 3.29, 3.30 (ii).)

(c) **IBP twice, returning something involving the original integral**. Provided the second IBP didn't undo the first one, we get an equation which we can solve, yielding a solution to the original integral (Example 3.31).

(d) **Combining substitution and IBP**. Example 3.32 (i), which starts with substitution then uses (a). Example 3.32 (ii), which starts with substitution and then uses (b), and finishes with another substitution.

3.7 Integrals containing trigonometric functions

Example 3.33. Solve the following integrals.

$$\text{(i)} \quad \int \sin(x)\cos(x)^7\,dx.$$
$$\text{(ii)} \quad \int \cos(x)^7\,dx.$$
$$\text{(iii)} \quad \int \cos(x)^4\sin(x)^5\,dx.$$

Solution. (i): We have already seen problems like this in Section 3.3 (Substitution Rule). We substitute

$$u = \cos(x), \quad du = -\sin(x)\,dx$$

to get

$$\begin{aligned} \int \sin(x)\cos(x)^7\,dx &= -\int u^7\,du \\ &= -\frac{u^8}{8} + C \\ &= -\frac{\cos(x)^8}{8} + C. \end{aligned}$$

(ii): We want to make a substitution as in (i), but at first this doesn't seem to have the right form. Using the identity

$$\cos(x)^2 + \sin(x)^2 = 1,$$

we get

$$\cos(x)^7 = (1 - \sin(x)^2)^3\cos(x),$$

and thus the substitution

$$u = \sin(x), \quad du = \cos(x)\,dx$$

will work. We compute

$$\begin{aligned}
\int \cos(x)^7\,dx &= \int (1-\sin(x)^2)^3\cos(x)\,dx \\
&= \int (1-u^2)^3\,du \\
&= \int 1-3u^2+3u^4-u^6\,du \\
&= u-u^3+\frac{3u^5}{5}-\frac{u^7}{7}+C \\
&= \sin(x)-\sin(x)^3+\frac{3\sin(x)^5}{5}-\frac{\sin(x)^7}{7}+C.
\end{aligned}$$

(iii): Again we use the identity $\sin(x)^2+\cos(x)^2=1$, and then substitute

$$u=\cos(x),\quad du=-\sin(x)\,dx,$$

to get

$$\begin{aligned}
\int \cos(x)^4\sin(x)^5\,dx &= \int \cos(x)^4(1-\cos(x)^2)^2\sin(x)\,dx \\
&= -\int u^4(1-u^2)^2\,du \\
&= \int -u^4+2u^6-u^8\,du \\
&= -\frac{u^5}{5}+\frac{2u^7}{7}-\frac{u^9}{9}+C \\
&= -\frac{\cos(x)^5}{5}+\frac{2\cos(x)^7}{7}-\frac{\cos(x)^9}{9}+C.
\end{aligned}$$

More generally, using the identity

$$\sin(x)^2+\cos(x)^2=1$$

and substitution, we can solve

$$\int \sin(x)^m\cos(x)^n\,dx$$

provided that m and n are positive integers, and one of m or n is odd.

What if both m and n are even? To handle these, we will need to make use of angle-sum formulas from trigonometry; we recall these are

$$\begin{aligned} \sin(A)\sin(B) &= \frac{1}{2}(\cos(A-B) - \cos(A+B)), \\ \cos(A)\cos(B) &= \frac{1}{2}(\cos(A+B) + \cos(A-B)), \\ \sin(A)\cos(B) &= \frac{1}{2}(\sin(A+B) + \sin(A-B)). \end{aligned}$$

It is helpful to remember the special cases where $A = B$:

$$\begin{aligned} \sin(A)^2 &= \frac{1}{2}(1 - \cos(2A)), \\ \cos(A)^2 &= \frac{1}{2}(1 + \cos(2A))), \\ \sin(A)\cos(A) &= \frac{1}{2}\sin(2A). \end{aligned}$$

With these, we can reduce any expression of the form

$$\sin(x)^m \cos(x)^n$$

into something involving only single powers of $\sin(kx)$ and $\cos(kx)$ (for various values of k). In fact, we can even handle *any* product of terms of the form $\sin(kx)$ and $\cos(kx)$.

Example 3.34. Solve the following integrals.

(i) $\displaystyle\int \sin(x)^2 \cos(x)^2 \, dx.$

(ii) $\displaystyle\int \sin(x)\sin(\sqrt{2}x)\sin(\sqrt{3}x) \, dx.$

(iii) $\displaystyle\int \sin(x)^2 \cos(5x)^2 \, dx.$

Solution. (i): Start by reducing the integrand,

$$\begin{aligned}
\sin(x)^2\cos(x)^2 &= \sin(x)^2(1-\sin(x)^2)\\
&= \sin(x)^2 - \sin(x)^4\\
&= \frac{1}{2}(1-\cos(2x)) - \left(\frac{1}{2}(1-\cos(2x))\right)^2\\
&= \frac{1}{2} - \frac{\cos(2x)}{2} - \frac{1}{4} + \frac{\cos(2x)}{2} - \frac{\cos(2x)^2}{4}\\
&= \frac{1}{4} - \frac{1}{4}\cdot\frac{1}{2}(\cos(4x)+1)\\
&= \frac{1}{8} - \frac{\cos(4x)}{8}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\int \sin(x)^2\cos(x)^2\,dx &= \int \frac{1}{8} - \frac{\cos(4x)}{8}\,dx\\
&= \frac{x}{8} - \frac{\sin(4x)}{4\cdot 8} + C.
\end{aligned}$$

(ii): Simplify the integrand:

$$\begin{aligned}
\sin(x)\sin(\sqrt{2}x)\sin(\sqrt{3}x) &= \frac{1}{2}(\cos(x-\sqrt{2}x) - \cos(x+\sqrt{2}x))\sin(\sqrt{3}x)\\
&= \frac{1}{4}(\sin(x-\sqrt{2}x+\sqrt{3}x) - \sin(x-\sqrt{2}x-\sqrt{3}x)\\
&\quad - \sin(x+\sqrt{2}x+\sqrt{3}x) + \sin(x+\sqrt{2}x-\sqrt{3}x)),
\end{aligned}$$

Hence,

$$\begin{aligned}
\int \sin(x)\sin(\sqrt{2}x)\sin(\sqrt{3}x)\,dx &= \frac{1}{4}\int \sin(x-\sqrt{2}x+\sqrt{3}x) - \sin(x-\sqrt{2}x-\sqrt{3}x)\\
&\qquad - \sin(x+\sqrt{2}x+\sqrt{3}x) + \sin(x+\sqrt{2}x-\sqrt{3}x)\,dx\\
&= \frac{1}{4}\left(-\frac{\cos((1-\sqrt{2}+\sqrt{3})x)}{1-\sqrt{2}+\sqrt{3}} + \frac{\cos((1-\sqrt{2}-\sqrt{3})x)}{1-\sqrt{2}-\sqrt{3}}\right.\\
&\qquad \left.+\frac{\cos((1+\sqrt{2}+\sqrt{3})x)}{1+\sqrt{2}+\sqrt{3}} - \frac{\cos((1+\sqrt{2}-\sqrt{3})x)}{1+\sqrt{2}-\sqrt{3}}\right)
\end{aligned}$$

(iii): Reduce the integrand,

$$\begin{aligned}\sin(x)^2\cos(5x)^2 &= \frac{1}{4}(1-\cos(2x))(1+\cos(10x))\\ &= \frac{1}{4}(1-\cos(2x)+\cos(10x)-\cos(2x)\cos(10x))\\ &= \frac{1}{4}(1-\cos(2x)+\cos(10x))-\frac{1}{8}(\cos(12x)+\cos(8x)).\end{aligned}$$

Thus,

$$\begin{aligned}\int \sin(x)^2\cos(5x)^2\,dx &= \frac{1}{4}\int 1-\cos(2x)+\cos(10x)\,dx-\frac{1}{8}\int \cos(12x)+\cos(8x)\,dx\\ &= \frac{1}{4}\left(x-\frac{\sin(2x)}{2}+\frac{\sin(10x)}{10}\right)-\frac{1}{8}\left(\frac{\sin(12x)}{12}+\frac{\sin(8x)}{8}\right)+C.\end{aligned}$$

What if we have something of the form $\int \sin(x)^m\cos(x)^n\,dx$, where one (or both) of m, n are negative integers?

It is often helpful as a first step to rewrite the integrand in terms of

$$\begin{aligned}\sec(x) &= \cos(x)^{-1}, \quad \csc(x) = \sin(x)^{-1},\\ \tan(x) &= \frac{\sin(x)}{\cos(x)}, \quad \cot(x) = \frac{\cos(x)}{\sin(x)}.\end{aligned}$$

(Often the question will already be in terms of these functions.) The identities

$$\begin{aligned}\tan(x)^2+1 &= \sec(x)^2,\\ \cot(x)^2+1 &= \csc(x)^2\end{aligned}$$

are useful. Also, common substitutions in these cases are

$$\begin{aligned}u &= \tan(x), \quad du = \sec(x)^2\,dx,\\ u &= \sec(x), \quad du = \sec(x)\tan(x)\,dx,\\ u &= \cot(x), \quad du = -\csc(x)^2\,dx, \quad \text{or}\\ u &= \csc(x), \quad du = -\csc(x)\cot(x)\,dx.\end{aligned}$$

Example 3.35. Solve the following integrals:

$$\text{(i)} \quad \int \frac{\sin(x)^5}{\cos(x)^4}\,dx.$$
$$\text{(ii)} \quad \int \frac{\sin(x)^4}{\cos(x)^6}\,dx.$$
$$\text{(iii)} \quad \int \tan(x)\,dx.$$
$$\text{(iv)} \quad \int \tan(x)^2\,dx.$$
$$\text{(v)} \quad \int \sec(x)\,dx.$$
$$\text{(vi)} \quad \int \sec(x)^3\,dx.$$

Solution. (i): We have $\sin(x)$ to an odd power, so we can use the substitution

$$u = \cos(x), \quad du = -\sin(x)\,dx.$$

With this we get

$$\begin{aligned}
\int \frac{\sin(x)^5}{\cos(x)^4}\,dx &= \int \sin(x) \cdot \frac{(1-\cos(x)^2)^2}{\cos(x)^4}\,dx \\
&= -\int \frac{(1-u^2)^2}{u^4}\,du \\
&= -\int \frac{1}{u^4} - \frac{2}{u^2} + 1\,du \\
&= \frac{1}{3u^3} - \frac{2}{u} - u + C \\
&= \frac{1}{3\cos(x)^3} - \frac{2}{\cos(x)} - \cos(x) + C.
\end{aligned}$$

(ii): Here we don't have an odd power on either $\sin(x)$ or $\cos(x)$. However, we may rewrite the integrand as

$$\tan(x)^4 \sec(x)^2,$$

which suggests we could try to substitute either $u = \tan(x)$ or $u = \sec(x)$. Since $\tan(x)$ has an even exponent, we can't substitute $u = \sec(x)$, so we use

$$u = \tan(x), \quad du = \sec(x)^2\, dx.$$

This yields

$$\begin{aligned}\int \frac{\sin(x)^4}{\cos(x)^6}\, dx &= \int \tan(x)^4 \sec(x)^2\, dx \\ &= \int u^4\, du \\ &= \frac{u^5}{5} + C \\ &= \frac{\tan(x)^5}{5} + C.\end{aligned}$$

(iii): We rewrite $\tan(x) = \cos(x)/\sin(x)$ and then notice that the best substitution is

$$u = \sin(x), \quad du = \cos(x)\, dx.$$

Thus,

$$\begin{aligned}\int \tan(x)\, dx &= \int \frac{\cos(x)}{\sin(x)}\, dx \\ &= \int \frac{1}{u}\, du \\ &= \ln|u| + C \\ &= \ln|\sin(x)| + C.\end{aligned}$$

(iv): This one is just a little trick. Using the identity $\tan(x)^2 + 1 = \sec(x)^2$, one gets

$$\int \tan(x)^2\, dx = \int \sec(x)^2 - 1\, dx = \tan(x) - x + C.$$

(v): This one requires a manipulation of the integrand that isn't obvious.

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

The numerator is the derivative of the denominator, so this is now amenable to the substitution

$$u = \sec(x) + \tan(x), \quad du = (\sec(x)\tan(x) + \sec(x)^2)\, dx.$$

We get

$$\begin{aligned}\int \sec(x)\, dx &= \int \frac{\sec(x)^2 + \sec(x)\tan(x)}{\sec(x)+\tan(x)}\, dx \\ &= \int \frac{1}{u}\, du \\ &= \ln|u| + C \\ &= \ln|\sec(x)+\tan(x)| + C.\end{aligned}$$

(vi): Whether we write the integrand as $\sec(x)^3$ or $1/\cos(x)^3$, it seems that there is no substitution that will help. Instead we do integration by parts, with

$$\begin{aligned}u = \sec(x), &\quad dv = \sec(x)^2\, dx \\ du = \sec(x)\tan(x)\, dx, &\quad v = \tan(x).\end{aligned}$$

With this we get

$$\int \sec(x)^3\, dx = \sec(x)\tan(x) - \int \tan(x)^2 \sec(x)\, dx \quad \text{(IBP)}.$$

At first this might not look like it helps, because we equally can't solve

$$\int \tan(x)^2 \sec(x)\, dx = \int \frac{\sin(x)^2}{\cos(x)^3}\, dx$$

using substitutions as before.

The correct next step is to rewrite $\tan(x)^2$ in terms of $\sec(x)$:

$$\begin{aligned}\int \sec(x)^3\, dx &= \sec(x)\tan(x) - \int \tan(x)^2\sec(x)\, dx \\ &= \sec(x)\tan(x) - \int (\sec(x)^2 - 1)\sec(x)\, dx \\ &= \sec(x)\tan(x) - \int \sec(x)^3\, dx + \int \sec(x)\, dx \\ &= \sec(x)\tan(x) - \int \sec(x)^3\, dx + \ln|\sec(x)+\tan(x)| + C,\end{aligned}$$

where we used (v) to get $\int \sec(x)\,dx$. Now we rearrange to solve for $\int \sec(x)^3\,dx$:

$$2\int \sec(x)^3\,dx = \sec(x)\tan(x) + \ln|\sec(x) + \tan(x)| + C,$$
$$\int \sec(x)^3\,dx = \frac{1}{2}(\sec(x)\tan(x) + \ln|\sec(x) + \tan(x)|) + C'.$$

3.8 Trigonometric substitutions

Trigonometric substitutions is one approach which is sometimes needed to solve integrals. For these, we define the substitution variable (u) implicitly:

$$x = \sin(u), \quad dx = \cos(u)\,du$$

or

$$x = \tan(u), \quad dx = \sec(u)^2\,du,$$

or

$$x = \sec(u), \quad dx = \tan(u)\sec(u)\,du.$$

Since we are defining x in terms of u (instead of the other way around as we did before with substitution), there will be an additional complication when getting the final answer in terms the original variable of integration.

Example 3.36. Solve the following:

(i) $$\int \frac{x^3}{\sqrt{1-x^2}}\,dx.$$

(ii) $$\int \frac{x^3}{\sqrt{x^2-1}}\,dx.$$

(iii) $$\int \frac{y^2}{9y^2+1}\,dy.$$

(iv) $$\int \frac{1}{y\sqrt{2-y^2}}\,dy.$$

Solution. (i): We recognise that if $x = \sin(u)$ then $\sqrt{1-x^2} = |\cos(u)|$ (and we assume that $\cos(u) \geq 0$), so it makes sense to use

$$x = \sin(u), \quad du = \cos(u)\, du.$$

to get

$$\begin{aligned}
\int \frac{x^3}{\sqrt{1-x^2}}\, dx &= \int \frac{\sin(u)^3}{\cos(u)} \cdot \cos(u)\, du \\
&= \int \sin(u)(1-\cos(u)^2)\, du \\
&= -\int 1 - v^2\, dv \quad \text{(subst. } v = \cos(u)\text{)} \\
&= -v + \frac{v^3}{3} + C \\
&= -\cos(u) + \frac{\cos(u)^3}{3} + C \\
&= -\sqrt{1-x^2} + \frac{(1-x^2)^{3/2}}{3} + C.
\end{aligned}$$

(ii): Whereas in (i) we had $\sqrt{1-x^2}$ (which suggests substituting $\sin(u)$ because of the identity $\sin(u)^2 + \cos(u)^2 = 1$), here we have $\sqrt{x^2-1}$. This suggests substituting $\sec(u)$ because of the identity $\tan(u)^2 + 1 = \sec(x)^2$. So, we use

$$x = \sec(u), \quad dx = \sec(u)\tan(u)\, du$$

to get

$$\begin{aligned}
\int \frac{x^3}{\sqrt{x^2-1}}\,dx &= \int \frac{\sec(u)^3}{\tan(u)} \cdot \sec(u)\tan(u)\,du \\
&= \int \sec(u)^4\,du \\
&= \int (\tan(u)^2+1)\sec(u)^2\,du \\
&= \int (v^2+1)\,dv \quad (\text{subst. } v = \tan(u)) \\
&= \frac{v^3}{3} + v + C \\
&= \frac{\tan(u)^3}{3} + \tan(u) + C \\
&= \frac{(x^2-1)^{3/2}}{3} + \sqrt{x^2-1} + C.
\end{aligned}$$

(iii): Observe that the numerator can be written $(3y)^2+1$, which suggests that we want to substitute

$$3y = \tan(u), \quad dy = \frac{\sec(u)^2}{3}\,du.$$

This leads to

$$\begin{aligned}
\int \frac{y^2}{9y^2+1}\,dy &= \int \frac{(\tan(u)/3)^2}{\tan(u)^2+1} \cdot \frac{\sec(u)^2}{3}\,du \\
&= \frac{1}{27}\int \tan(u)^2\,du \\
&= \frac{1}{27}\int \sec(u)^2 - 1\,du \\
&= \frac{1}{27}(\tan(u) - u) + C \\
&= \frac{y}{9} - \frac{1}{27}\tan^{-1}\left(\frac{y}{3}\right) + C.
\end{aligned}$$

(In the last step, we needed to invert $3y = \tan(u)$ to write u in terms of y.)

(iv): We rewrite the integrand as

$$\frac{1}{y\sqrt{2-y^2}} = \frac{1}{y\sqrt{2}\sqrt{1-(y/\sqrt{2})^2}},$$

suggesting that we want to substitute

$$\frac{y}{\sqrt{2}} = \sin(u), \quad dy = \sqrt{2}\cos(u)\,du.$$

This gives

$$\begin{aligned}\int \frac{1}{y\sqrt{2-y^2}}\,dy &= \frac{1}{\sqrt{2}}\int \frac{1}{\sqrt{2}\sin(u)\cos(u)} \cdot \sqrt{2}\cos(u)\,du \\ &= \frac{1}{\sqrt{2}}\int \frac{1}{\sin(u)}\,du \\ &= \frac{1}{\sqrt{2}}\int \csc(u)\,du\end{aligned}$$

The rest of this part can be solved much like Example 3.35 (v), and is left as an exercise (in fact it is Question 1(f) on Exercise Sheet 3).

Example 3.37. Solve the following:

$$\text{(i)} \quad \int \frac{x}{\sqrt{4x^2 + 8x - 5}}\, dx.$$
$$\text{(ii)} \quad \int x\sqrt{x^2 - 6x + 10}\, dx.$$
$$\text{(iii)} \quad \int \frac{1}{(x^2 + 4)^2}, dx.$$

Solution. (i): Whenever we have a square root of something involving x^2, it is very likely that a trigonometric substitution is needed. However, in cases like this, we need to rearrange to get something like $\sqrt{a(y^2 \pm 1)}$ (where $a \in \mathbb{R}$). This is done by "completing the square":

$$\begin{aligned} 4x^2 + 8x - 5 &= 4(x^2 + 2x + 1) - 9 \\ &= 4(x + 1)^2 - 9. \end{aligned}$$

Thus,

$$\sqrt{4(x + 1)^2 - 9} = 3\sqrt{\left(\frac{2(x + 1)}{3}\right)^2 - 1},$$

and, as in (ii), we want to substitute

$$\frac{2(x + 1)}{3} = \sec(u), \quad \frac{2}{3}\, dx = \sec(u)\tan(u)\, du.$$

It is useful to write x in terms of u before performing this substitution:

$$x = \frac{3\sec(u)}{2} - 1.$$

We get

$$
\begin{aligned}
\int \frac{x}{\sqrt{4x^2+8x-5}}\,dx &= \int \frac{x}{3\sqrt{(2(x+1)/3)^2-1}}\,dx \\
&= \int \frac{3\sec(u)/2-1}{3\tan(u)} \cdot \frac{3\sec(u)\tan(u)}{2}\,du \\
&= \frac{1}{2}\int \frac{3}{2}\sec(u)^2 - \sec(u)\,du \\
&= \frac{3}{4}\tan(u) - \frac{1}{2}\ln|\sec(u)+\tan(u)| + C \\
&= \frac{3}{4}\sqrt{\left(\frac{2(x+1)}{3}\right)^2 - 1} \\
&\quad - \frac{1}{2}\ln\left|\frac{2(x+1)}{3} + \sqrt{\left(\frac{2(x+1)}{3}\right)^2-1}\right| + C \\
&= \frac{1}{4}\sqrt{4(x+1)^2-9} \\
&\quad - \frac{1}{2}\left(\ln\left|2(x+1)+\sqrt{4(x+1)^2-9}\right| - \ln(3)\right) + C \\
&= \frac{1}{4}\sqrt{4(x+1)^2-9} \\
&\quad - \frac{1}{2}\ln\left|2(x+1)+\sqrt{4(x+1)^2-9}\right| + C'.
\end{aligned}
$$

(We used Example 3.35 (v) for $\int \sec(u)\,du$.)

(ii): Completing the square, we have

$$\sqrt{x^2-6x+10} = \sqrt{(x-3)^2+1},$$

so we use the substitution

$$x-3=\tan(u), \quad dx = \sec(u)^2\,du.$$

This yields

$$
\begin{aligned}
\int x\sqrt{x^2-6x+10}\,dx &= \int (\tan(u)+3)^2\sec(u)\cdot\sec(u)^2\,du \\
&= \int \tan(u)\sec(u)^3 + 3\sec(u)^3\,du.
\end{aligned}
$$

We break this into two integrals and solve them separately. For the first one, substitute

$$v = \sec(u), \quad dv = \tan(u)\sec(u)\,du$$

to get

$$\begin{aligned}\int \tan(u)\sec(u)^3\,du &= \int v^2\,dv \\ &= \frac{v^3}{3} + C \\ &= \frac{\sec(u)^3}{3} + C.\end{aligned}$$

For the second one, Example 3.35 (vi) tells us that

$$\int \sec(u)^3\,du = \frac{1}{2}(\sec(u)\tan(u) + \ln|\sec(u) + \tan(u)|) + C'.$$

We therefore have

$$\begin{aligned}\int x\sqrt{x^2 - 6x + 10}\,dx &= \int \tan(u)\sec(u)^3 + 3\sec(u)^3\,du \\ &= \frac{\sec(u)^3}{3} + \frac{\sec(u)\tan(u) + \ln|\sec(u) + \tan(u)|}{2} + C'' \\ &= \frac{((x-3)^2+1)^{3/2}}{3} + \frac{\sqrt{(x-3)^2+1}(x-3)}{2} \\ &\quad + \frac{\ln\left|\sqrt{(x-3)^2+1} + (x-3)\right|}{2} + C''.\end{aligned}$$

(iii): Write the integrand as

$$\frac{1}{(x^2+4)^2} = \frac{1}{16((x/2)^2+1)^2},$$

and then use the substitution

$$x/2 = \tan(u), \quad dx = 2\sec(u)^2\,dx.$$

This leads to

$$\begin{aligned}
\int \frac{1}{(x^2+4)^2}, dx &= \int \frac{1}{16((x/2)^2+1)^2}\, dx \\
&= \int \frac{2\sec(u)^2}{16(\tan(u)^2+1)^2}\, du \\
&= \frac{1}{8}\int \frac{\sec(u)^2}{\sec(u)^4}\, du \\
&= \frac{1}{8}\int \cos(u)^2\, du \\
&= \frac{1}{16}\int 1+\cos(2u)\, du \\
&= \frac{1}{16}\left(u+\frac{\sin(2u)}{2}\right)+C \\
&= \frac{1}{16}\left(u+\sin(u)\cos(u)\right)+C \\
&= \frac{1}{16}\left(u+\frac{\tan(u)}{\sec(u)^2}\right)+C \\
&= \frac{1}{16}\left(\tan^{-1}\left(\frac{x}{2}\right)+\frac{x/2}{x^2/4+1}\right)+C. \\
&= \frac{1}{16}\tan^{-1}\left(\frac{x}{2}\right)+\frac{x}{8(x^2+4)}+C.
\end{aligned}$$

As we saw in the previous examples, the general rules for trig substitutions are:

If the integrand contains ...	**Substitute ...**
$1-x^2$ (especially $\sqrt{1-x^2}$)	$x=\sin(u)$
x^2-1 (especially $\sqrt{x^2-1}$)	$x=\sec(u)$
x^2+1 (especially $\sqrt{x^2+1}$)	$x=\tan(u)$.

When the integrand contains a quadratic form (especially under a square root), first complete the square to get it in one of the above forms.

3.9 Integrating rational functions: partial fraction decomposition

We already know of a procedure to integrate polynomials. Here we shall see how to integrate a "rational function", i.e., a quotient of one polynomial by another.

Occasionally, there might be a substitution available to solve a complicated rational function, e.g.,

$$\begin{aligned}\int \frac{3x^2+x}{2x^3+x^2-7}\,dx &= \int \frac{1}{u}\,\frac{du}{2} \quad (\text{subst. } u = 2x^3+x^2-7)\\ &= \ln|u| + C\\ &= \ln\left|2x^3+x^2-7\right| + C.\end{aligned}$$

Usually, this approach won't work, e.g., with

$$\int \frac{5x-15}{x^2+3x-4}\,dx.$$

This integral can be solved easily, however, by noticing

$$\frac{7}{x+4} - \frac{2}{x-1} = \frac{7(x-1)-2(x+4)}{x^2+3x-4} = \frac{5x-15}{x^2+3x-4} \tag{3.4}$$

so that

$$\begin{aligned}\int \frac{5x-15}{x^2+3x-4}\,dx &= \int \frac{7}{x+4} - \frac{2}{x-1}\,dx\\ &= 7\ln|x+4| - 2\ln|x-1| + C.\end{aligned}$$

Finding the decomposition (3.4) might look, at first, like a stroke of luck. In fact, there is a systematic way to do this. We first note that the denominator of the integrand, x^2+3x-4, factors as

$$x^2+3x-4 = (x+4)(x-1).$$

Then we try to solve

$$\frac{5x-15}{x^2+3x-4} = \frac{A}{x+4} + \frac{B}{x-1} \tag{3.5}$$

for real numbers A and B. Multiplying the equation by (x^2+3x-4) produces

$$5x - 15 = A(x-1) + B(x+4) = (A+B)x + (-A+4B).$$

We view this as a system of two equations in two unknowns,

$$\begin{array}{lrcl} (1): & 5 & = & A+B, \\ (2): & -15 & = & -A+4B \end{array}$$

Now we solve this system:

$$\begin{array}{lrcl} (1)+(2): & -10 & = & 5B, \\ & B & = & -2, \\ (1): & 5 & = & A+(-2), \\ & A & = & 7. \end{array}$$

Plugging this solution into (3.5) yields

$$\frac{5x-15}{x^2+3x-4} = \frac{7}{x+4} - \frac{2}{x-1}.$$

This sort of thing can be done very generally, which is the content of the following theorem:

Theorem 3.38 (Partial Fraction Decomposition)**.** *Let $f(x)$ be a rational function,*

$$f(x) = \frac{p(x)}{q_1(x)^{n_1} \cdots q_k(x)^{n_k} r_1^{m_1} \cdots r_l^{m_l}},$$

where $p(x)$ is a polynomial and q_i are distinct monic linear polynomials,

$$q_i(x) = x + a_i, \quad i = 1, \dots, k$$

and the r_i are distinct monic quadratic polynomials,

$$q_i(x) = x^2 + b_i x + c_i \text{ where } b_i^2 - 4c_i < 0, \quad i = 1, \dots, l.$$

Then $f(x)$ decomposes as

$$f(x) = g(x) + \sum_{i=1}^{k} \sum_{j=1}^{n_i} \frac{A_{i,j}}{(x+a_i)^j} + \sum_{i=1}^{l} \sum_{j=1}^{m_i} \frac{B_{i,j}x + C_{i,j}}{(x^2 + b_i^x + c_i)^j},$$

for some polynomial $g(x)$ and some real numbers $A_{i,j}, B_{i,j}, C_{i,j}$.

We do not prove this theorem in this course. The statement of this theorem is fairly involved, so let us break it down a bit, and explain how to use it.

We start with a general rational function

$$f(x) = \frac{p(x)}{q(x)}.$$

We may assume that $q(x)$ is a monic polynomial (i.e., the coefficient on its first term is 1). By the Fundamental Theorem of Algebra, $q(x)$ factors into irreducible linear and quadratic terms. (In practice, given a "random" polynomial $q(x)$, it may be extremely difficult to factor it; in some cases, it is not even possible to factor it exactly. However, in all examples in this course, this factorisation will be possible and not too difficult.) There may be some factors repeated in the factorisation of $q(x)$; these need to be collected together. Altogether, this yields a factorisation

$$q(x) = q_1(x)^{n_1} \cdots q_k(x)^{n_k} r_1(x)^{m_1} \cdots r_l(x)^{m_k}$$

where $q_i(x)$ and $r_i(x)$ are as described in the above theorem.

We can then expect to express $f(x)$ as a sum of the following terms:

1. Some polynomial, $g(x)$.

2. For every linear term $q_i(x) = x + a_i$ in the factorisation of $q(x)$ (occurring n_i times) terms of the form

$$\frac{A_{i,j}}{(x + a_i)^j},$$

   where $j = 1, \ldots, n_i$. (If $n_i = 1$, i.e., if $(x + a_i)$ only appears once in the factorisation of $q(x)$, then we just have one term here.)

3. For every quadratic term $r_i(x) = x^2 + b_i x + c_i$ in the factorisation of $q(x)$ (occurring m_i times) terms of the form

$$\frac{B_{i,j}x + C_{i,j}}{(x^2 + b_i^x + c_i)^j}$$

   where $j = 1, \ldots, m_i$. (Again, if $m_i = 1$, i.e., if $(x^2 + b_i x + c_i)$ only appears once in the factorisation of $q(x)$, then we just have one term here.)

Let us do some examples of how to express rational functions using this theorem. After this, we will go on to integrating rational functions.

Example 3.39. Find the partial fraction decomposition of the following rational functions:

$$\text{(i)} \quad \frac{x+7}{x^2+5x+6}.$$

$$\text{(ii)} \quad \frac{x+7}{x^2+4x+4}.$$

$$\text{(iii)} \quad \frac{31x+51}{x^3+4x^2-15x-18}.$$

$$\text{(iv)} \quad \frac{x^4+x^3+5x^2-10x-1}{(x+3)(x-1)^2}.$$

$$\text{(v)} \quad \frac{x^3+7x^2+6x+4}{x^3-1}.$$

Solution. (i): First factor the denominator:

$$x^2+5x+6 = (x+2)(x+3).$$

Hence we want to solve

$$\frac{x+7}{x^2+5x+6} = g(x) + \frac{A}{x+2} + \frac{B}{x+3}$$

for a polynomial $g(x)$ and scalars $A, B \in \mathbb{R}$.

Multiplying both sides by (x^2+5x+6) produces

$$\begin{aligned} x+7 &= g(x)(x^2+5x+6) + A(x+3) + B(x+2) \\ &= g(x)(x^2+5x+6) + (A+B)x + (3A+2B). \end{aligned}$$

If the polynomial $g(x)$ were nonzero, then the right-hand side would have some term involving x^k for some $k \geq 2$. Since the left-hand side has degree one, the polynomial $g(x)$ must be zero. This gives the system of equations

$$\begin{array}{lrcl} (1): & 1 & = & A+B, \\ (2): & 7 & = & 3A+2B. \end{array}$$

We solve this as follows:

$$\begin{aligned}(2)-3(1): \quad 4 &= -B \\ B &= -4 \\ (1): \quad 1 &= A-4, \\ A &= 5.\end{aligned}$$

We therefore have

$$\frac{x+7}{x^2+5x+6} = \frac{5}{x+2} - \frac{4}{x+3}.$$

(ii): The denominator factors as

$$x^2+4x+4 = (x+2)^2.$$

Hence we want to solve

$$\frac{x+7}{x^2+4x+4} = g(x) + \frac{A}{x+2} + \frac{B}{(x+2)^2}$$

for a polynomial $g(x)$ and scalars $A, B \in \mathbb{R}$.

Multiplying this equation by $(x+2)^2$ produces

$$x+7 = g(x)(x+2)^2 + A(x+2) + B = g(x)(x+2)^2 + Ax + (2A+B).$$

Again, $g(x)$ must be zero since the left-hand side has degree < 2. This system solves easily:

$$A = 1, \quad B = 7-2 = 5.$$

Therefore,

$$\frac{x+7}{x^2+4x+4} = \frac{1}{x+2} + \frac{5}{(x+2)^2}.$$

(iii): First factor the denominator:

$$x^3+4x^2-15x-18 = (x-3)(x+1)(x+6)$$

Hence we want to solve

$$\frac{31x+51}{x^3+4x^2-15x-18} = g(x) + \frac{A}{x-3} + \frac{B}{x+1} + \frac{C}{x+6}$$

for a polynomial $g(x)$ and scalars $A, B, C \in \mathbb{R}$.

Multiplying this equation by $(x^3+4x^2-15x-18)$ produces

$$\begin{aligned}31x+51 &= g(x)(x^3+4x^2-15x-18)+A(x+1)(x+6)+B(x-3)(x+6)\\&\quad +C(x-3)(x+1)\\&= g(x)(x^3+4x^2-15x-18)+(A+B+C)x^2+(7A+3B-2C)x\\&\quad +(6A-18B-3C).\end{aligned}$$

Since the left-hand side has degree < 3, the polynomial $g(x)$ must be zero. This gives a system of three equations,

$$\begin{aligned}(1):\quad 0 &= A+B+C,\\(2):\quad 31 &= 7A+3B-2C,\\(3):\quad 51 &= 6A-18B-3C.\end{aligned}$$

We now solve this system:

$$\begin{aligned}(4):\ (2)+2(1):\quad 31 &= 9A+5B,\\(5):\ (3)+3(1):\quad 51 &= 9A-15B,\\(5)-(4):\quad 20 &= -20B,\\B &= -1,\\(4):\quad 31 &= 9A-5,\\A &= 36/9=4,\\(1):\quad 0 &= 4-1+C,\\-3 &= C.\end{aligned}$$

Therefore,

$$\frac{31x+51}{x^3+4x^2-15x-18}=\frac{4}{x-3}-\frac{1}{x+1}-\frac{3}{x+6}.$$

(iv): The denominator has already been factored, so we go straight to the partial fraction decomposition with variables to be solved. We want to solve

$$\frac{x^4+x^3+5x^2-10x-1}{(x+3)(x-1)^2}=g(x)+\frac{A}{x+3}+\frac{B}{x-1}+\frac{C}{(x-1)^2}.$$

Multiplying by $(x+3)(x-1)^2$ yields

$$x^4+x^3+5x^2-10x-1=g(x)(x+3)(x-1)^2+A(x-1)^2+B(x+3)(x-1)+C(x+3).$$

Here, the left-hand side has degree 4, while the component of the right-hand side that doesn't involve $g(x)$ can only have degree at most 2. Hence, we need the polynomial $g(x)$ to have degree 1, so we let $g(x) = Dx + E$, giving

$$\begin{aligned}
&x^4 + x^3 + 5x^2 - 10x - 1 \\
&\quad = (Dx + E)(x + 3)(x - 1)^2 + A(x - 1)^2 + B(x + 3)(x - 1) + C(x + 3) \\
&\quad = Dx^4 + (D + E)x^3 + (A + B - 5D + E)x^2 \\
&\qquad + (-2A + 2B + C + 3D - 5E)x + (A - 3B + 3C + 3E),
\end{aligned}$$

translating into the system

$$\begin{array}{lrcl}
(1): & 1 & = & D, \\
(2): & 1 & = & D + E, \\
(3): & 5 & = & A + B - 5D + E, \\
(4): & -10 & = & -2A + 2B + C + 3D - 5E, \\
(5): & -1 & = & A - 3B + 3C + 3E.
\end{array}$$

We now solve this system:

$$\begin{array}{rrcl}
(2): & 1 & = & 1 + E, \\
 & E & = & 0, \\
(6):\ (5) - 3(4): & 29 & = & 7A - 9B - 9(1), \\
 & 38 & = & 7A - 9B, \\
(6) + 9(3): & 83 & = & 16A - 45 \\
 & A & = & 128/16 = 8, \\
(3): & 5 & = & 8 + B - 5, \\
 & B & = & 2, \\
(5): & -1 & = & 8 - 3(2) + 3C \\
 & C & = & -3/3 = -1.
\end{array}$$

Therefore,

$$\frac{x^4 + x^3 + 5x^2 - 10x - 1}{(x + 3)(x - 1)^2} = x + \frac{8}{x + 3} + \frac{2}{x - 1} - \frac{1}{(x - 1)^2}.$$

Generally, when the degree of the numerator is below the degree of the denominator, the polynomial term ("$g(x)$") will be zero in the Partial Fraction Decomposition. Otherwise, the degree of $g(x)$ is the difference between the degree of the numerator and the degree of the numerator.

(v): The denominator factors as

$$(x-1)(x^2+x+1),$$

and we know that we can't factor x^2+x+1, since the discriminant is

$$1^2-4(1)=-3<0.$$

We want to solve

$$\frac{x^3+7x^2+6x+4}{x^3-1}=g(x)+\frac{A}{x-1}+\frac{Bx+C}{x^2+x+1}.$$

Multiplying by x^3-1 yields

$$x^3+7x^2+6x+4=g(x)(x^3-1)+A(x^2+x+1)+(Bx+C)(x-1).$$

Since the right-hand side has degree 3, we see that we need $g(x)$ to be a constant, $g(x)=D$. The above equation becomes

$$\begin{aligned}x^3+7x^2+6x+4&=D(x^3-1)+A(x^2+x+1)+(Bx+C)(x-1)\\&=Dx^3+(A+B)x^2+(A-B+C)x+(-D+A-C).\end{aligned}$$

Our system is thus

$$\begin{array}{rrcl}(1): & 1 & = & D,\\(2): & 7 & = & A+B,\\(3): & 6 & = & A-B+C,\\(4): & 4 & = & A-C-D\\ & & = & A-C-1.\end{array}$$

We solve this system:

$$\begin{array}{rrcl}(2)+(3)+(4): & 17 & = & 3A-1,\\ & A & = & 18/3=6,\\(2): & 7 & = & 6+B,\\ & B & = & 1,\\(3): & 6 & = & 6-1+C,\\ & C & = & 1.\end{array}$$

Therefore,

$$\frac{x^3+7x^2+6x+4}{x^3-1}=1+\frac{6}{x-1}+\frac{x+1}{x^2+x+1}.$$

Once we've decomposed a rational function using Partial Fraction Decomposition, we should be able to integrate the various terms that appear. Here are some examples.

Example 3.40. Solve the following integrals.

$$\text{(i)} \quad \int \frac{x+7}{x^2+5x+6}\,dx.$$

$$\text{(ii)} \quad \int \frac{x+7}{x^2+4x+4}\,dx.$$

$$\text{(iii)} \quad \int \frac{x^4+x^3+5x^2-10x-1}{(x+3)(x-1)^2}\,dx.$$

$$\text{(iv)} \quad \int \frac{7x^2-x+4}{x^3+x}\,dx.$$

$$\text{(v)} \quad \int \frac{x^3-7x^2}{x^2-6x+10}\,dx.$$

$$\text{(vi)} \quad \int \frac{4x^3-8x^2-x+1}{16x^4+8x^2+1}\,dx.$$

Solution. (i): Using Example 3.39 (i), we get

$$\begin{aligned}\int \frac{x+7}{x^2+5x+6}\,dx &= \int \frac{5}{x+2} - \frac{4}{x+3}\,dx \\ &= 5\ln|x+2| - 4\ln|x+3| + C.\end{aligned}$$

(ii): Using Example 3.39 (ii), we get

$$\begin{aligned}\int \frac{x+7}{x^2+4x+4}\,dx &= \int \frac{1}{x+2} + \frac{5}{(x+2)^2}\,dx \\ &= \ln|x+2| - \frac{5}{x+2} + C.\end{aligned}$$

(iii): Using Example 3.39 (ii), we get

$$\begin{aligned}\int \frac{x^4+x^3+5x^2-10x-1}{(x+3)(x-1)^2}\,dx &= \int x + \frac{8}{x+3} + \frac{2}{x-1} - \frac{1}{(x-1)^2}\,dx \\ &= x^2 + 8\ln|x+3| + 2\ln|x-1| + \frac{1}{x-1} + C.\end{aligned}$$

(iv): We need to do Partial Fraction Decomposition on the integrand. The denominator factors as

$$x^3 + x = x(x^2 + 1),$$

so we want to solve

$$\frac{7x^2 - x + 4}{x^3 + x} = \frac{A}{x} + \frac{Bx + C}{x^2 + 1}.$$

(We see that we don't need to add a polynomial, because the degree of the numerator is less than the degree of the denominator.)

We get

$$7x^2 - x + 4 = A(x^2 + 1) + (Bx + C)x = (A + B)x^2 + Cx + A,$$

which easily solves as

$$A = 4, \quad B = 3, \quad C = -1.$$

Therefore,

$$\int \frac{7x^2 - x + 4}{x^3 + x}\,dx = \int \frac{4}{x} + \frac{3x - 1}{x^2 + 1}\,dx.$$

to integrate $\frac{3x-1}{x^2+1}$, we separate it into two sums, using the substitution $u = x^2 + 1$ for the first.

$$\begin{aligned}\int \frac{7x^2 - x + 4}{x^3 + x}\,dx &= \int \frac{4}{x} + \frac{3x - 1}{x^2 + 1}\,dx \\ &= 4\ln|x| + \frac{3}{2}\ln(x^2 + 1) - \tan^{-1}(x) + C.\end{aligned}$$

(v): The denominator in this case cannot be further factored, since the discriminant is

$$(-6)^2 - 4(10) = -4 < 0.$$

So we just need to write the integrand as

$$\begin{aligned}\frac{x^3 - 7x^2}{x^2 - 6x + 10} &= \frac{Ax + B}{x^2 - 6x + 10} + Cx + D, \\ x^3 - 7x^2 &= Ax + B + (Cx + D)(x^2 - 6x + 10) \\ &= Cx^3 + (-6C + D)x^2 + (A + 10C - 6D)x + (B + 10D).\end{aligned}$$

This gives the system

$$\begin{aligned} 1 &= C, \\ -7 &= -6C + D, \\ 0 &= A + 10C - 6D, \\ 0 &= B + 10D, \end{aligned}$$

which we can solve,

$$C = 1, \quad , D = -1, \quad A = -16, \quad B = 10.$$

Therefore,

$$\frac{x^3 - 7x^2}{x^2 - 6x + 10} = \frac{-16x + 10}{x^2 - 6x + 10} + x - 1.$$

When we integrate this, we'll want to use the substitution

$$u = x^2 - 6x + 10, \quad du = (2x - 6)\,dx$$

to deal with part of the fractional term. We therefore break apart this fractional term,

$$\frac{-16x + 10}{x^2 - 6x + 10} = -8 \cdot \frac{2x - 6}{x^2 - 6x + 1} - \frac{38}{x^2 - 6x + 10}.$$

The second term here we handle by completing the square in the denominator:

$$x^2 - 6x + 10 = (x - 3)^2 + 1.$$

Putting this together, we get

$$\begin{aligned} &\int \frac{x^3 - 7x^2}{x^2 - 6x + 10}\,dx \\ &\quad = -8 \int \frac{2x - 3}{x^2 - 6x + 1}\,dx - 38 \int \frac{1}{(x-3)^2 + 1}\,dx + \int x - 1\,dx \\ &\quad = -8 \int \frac{1}{u}\,du - 38 \int \frac{1}{(x-3)^2 + 1}\,dx + \int x - 1\,dx \\ &\quad = -8 \ln|u| - 38 \tan^{-1}(x - 3) + \frac{x^2}{2} - x + C \\ &\quad = -8 \ln(x^2 - 6x + 10) - 38 \tan^{-1}(x - 3) + \frac{x^2}{2} - x + C. \end{aligned}$$

(vi): Factor the denominator,

$$16x^4 + 8x^2 + 1 = (4x^2 + 1)^2,$$

so our Partial Fraction Decomposition looks like

$$\begin{aligned}\frac{4x^3 - 8x^2 - x + 1}{16x^4 + 8x^2 + 1} &= \frac{Ax + B}{4x^2 + 1} + \frac{Cx + D}{(4x^2 + 1)^2},\\ 4x^3 - 8x^2 - x + 1 &= (Ax + B)(4x^2 + 1) + Cx + D\\ &= 4Ax^3 + 4Bx^2 + (A + C)x + (B + D)\end{aligned}$$

This solves as

$$A = 1, \quad B = -2, \quad C = -2, \quad D = 3,$$

so that

$$\frac{4x^2 + 2x - 2}{16x^4 + 8x^2 + 1} = \frac{x - 2}{4x^2 + 1} - \frac{2x - 3}{(4x^2 + 1)^2}.$$

To integrate this, we note that

$$\begin{aligned}\int \frac{x}{4x^2 + 1}\, dx &= \frac{\ln(4x^2 + 1)}{8} + C_1, \quad (\text{subst. } u = 4x^2 + 1)\\ \int \frac{1}{4x^2 + 1}\, dx &= \frac{\tan^{-1}(2x)}{2} + C_2,\\ \int \frac{x}{(4x^2 + 1)^2}\, dx &= -\frac{1}{8(4x^2 + 1)} + C_3, \quad (\text{subst. } u = 4x^2 + 1),\\ \int \frac{1}{(4x^2 + 1)^2}\, dx &= \frac{\tan^{-1}(2x)}{4} + \frac{x}{2(1 + x^2)} + C_4,\end{aligned}$$

while the last one is done with the trig substitution $2x = \tan(u)$, similarly to Example 3.37 (iii). Hence,

$$\begin{aligned}\int \frac{4x^2 + 2x - 2}{16x^4 + 8x^2 + 1}\, dx &= \int \frac{x - 2}{4x^2 + 1} - \frac{2x - 3}{(4x^2 + 1)^2}\, dx\\ &= \frac{\ln(4x^2 + 1)}{8} - 2 \cdot \frac{\tan^{-1}(2x)}{2} - 2\left(-\frac{1}{8(4x^2 + 1)}\right)\\ &\quad + 3\left(\frac{\tan^{-1}(2x)}{4} + \frac{x}{2(1 + x^2)}\right) + C\\ &= \frac{\ln(4x^2 + 1)}{8} - \frac{\tan^{-1}(2x)}{4} - \frac{1 + 6x}{4(4x^2 + 1)} + C.\end{aligned}$$