UNIVERSITY OF NEW SOUTH WALES

SCHOOL OF MATHEMATICS AND STATISTICS

# MATH1141 Higher Mathematics 1A

# CALCULUS

**Semester 1 2018**

Professor Wolfgang K. Schief

# Chapter 1

# Sets, inequalities and functions

## 1.1 Sets of numbers

A set is a collection of distinct objects. The objects in a set are called the elements or members of the set.

- The set $\mathbb{N}$ of natural numbers is given by

$$\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}.$$

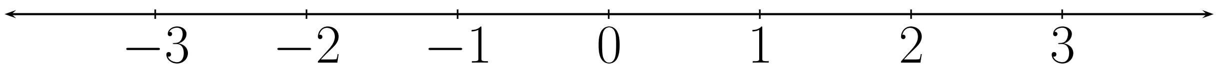
- The set $\mathbb{Z}$ of integers is given by

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

- The set $\mathbb{Q}$ of rational numbers is given by

$$\mathbb{Q} = \{\text{fractions of integers}\} = \left\{ \frac{p}{q} : p, q \in \mathbb{Z},\ q \neq 0 \right\}.$$

- The set $\mathbb{R}$ of real numbers may be represented as the collection of points lying on the number line.




**Notation.** If $x$ is a member of a set $A$, then we write $x \in A$. If $x$ is not a member of $A$ then we write $x \notin A$.

**Example.**

$$\frac{22}{7} \notin \mathbb{N}, \quad \frac{22}{7} \notin \mathbb{Z}, \quad \frac{22}{7} \in \mathbb{Q}, \quad \frac{22}{7} \in \mathbb{R}$$

**Remark.** The set

$$\{x \in \mathbb{R} : x \notin \mathbb{Q}\}$$

is the set of irrational numbers.

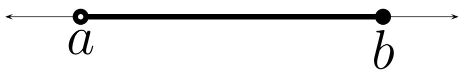
**Notation for intervals.** Suppose that $a$ and $b$ are real numbers and that $a < b$. Then

- $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ (open)


- $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ (closed)


- $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$ (???)

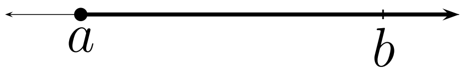

- $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$ (???)



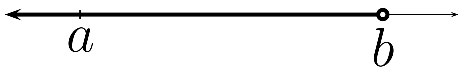

An interval $[a, b]$ that includes its endpoints $a$ and $b$ is called a closed interval, while an interval $(a, b)$ that excludes its endpoints is called an open interval. The intervals $[a, b)$ and $(a, b]$ are neither open nor closed.

**Rays** of the real line using the symbol $\infty$.

- $[a, \infty) = \{x \in \mathbb{R} : a \leq x\}$




- $(-\infty, b) = \{x \in \mathbb{R} : x < b\}$




- $(-\infty, \infty) = \mathbb{R}$

**Definition.** We say that a set $A$ is a subset of a set $B$ if every element of $A$ is an element of $B$. If $A$ is a subset of $B$ then we also say that $B$ contains the set $A$.

**Examples.**

- $\mathbb{N}$ is a subset of $\mathbb{Z}$, and $\mathbb{Z}$ is a subset of $\mathbb{Q}$, and $\mathbb{Q}$ is a subset of $\mathbb{R}$.
- $\{0, 2, 3\}$ is a subset of $\{0, 1, 2, 3, 5\}$.
- $(-1, 2]$ is not a subset of $[0, \infty)$.
- Any set is a subset of itself.
- $(1, 3)$ is a subset of $[1, 3)$.

## 1.2 Solving inequalities

**Inequalities.** Remember the following rules for manipulating inequalities.

- If $x > y$ then $x + z > y + z$.
- If $x > y$ and $z > 0$ then $xz > yz$.
- If $x > y$ and $z < 0$ then $xz < yz$.

Here, $x, y$ and $z$ are real numbers.

Two types of inequalities deserve special attention: quadratic inequalities and rational inequalities.

**Examples.** (a) Solve the quadratic inequality

$$x^2 + 4x \geq 21.$$

(b) Solve the rational inequality

$$\frac{1}{x+1} < \frac{1}{(x-2)(x-3)}.$$

### 1.3 Absolute values

The magnitude, or absolute value, of a real number $x$ is denoted by $|x|$ and defined by

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

**Properties.** Suppose that $x$ and $y$ are real numbers. Then

- $|-x| = |x|$,
- $|xy| = |x||y|$,
- $\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$ provided that $y \neq 0$,
- $|x + y| \leq |x| + |y|$,
- $|x - y| \geq |x| - |y|$.

The fourth property is known as the triangle inequality.

The following facts are useful for solving inequalities.

- For every real number $x$,

$$|x| = \sqrt{x^2}, \quad |x|^2 = x^2.$$

- Geometrically, $|x-a|$ represents the distance between $x$ and $a$ on the real number line.

- For any positive real number $a$,

$$|x| < a \quad \Leftrightarrow \quad x^2 < a^2 \quad \Leftrightarrow \quad -a < x < a.$$

- For any positive real number $a$,

$$|x| > a \quad \Leftrightarrow \quad x^2 > a^2 \quad \Leftrightarrow \quad x < -a \quad \text{or} \quad x > a.$$

**Examples.** Solve the following inequalities.

(a)

$$|3x + 1| \geq 4$$

(b)

$$\frac{|x + 5|}{|x - 11|} < 1$$

### 1.4 Functions

A function
$$f : A \to B$$
is a rule which assigns to every element $x$ belonging to a set $A$ exactly one element $f(x)$ belonging to a set $B$, that is
$$x \mapsto f(x).$$

**Terminology.**

- $A$ is called the domain of the function $f$, that is
$$A = \mathrm{Dom}(f) = \{\text{all allowable inputs}\}.$$

- $B$ is called the codomain of $f$, that is
$$B = \mathrm{Codom}(f) = \{\text{all allowable outputs}\}.$$

- The range of $f$ is
$$\begin{aligned}\mathrm{Range}(f) &= \{f(x) : x \in A\} \\ &= \{\text{all outputs that actually occur}\}.\end{aligned}$$

**Example.**

$$\begin{aligned} f : [0, \infty) &\to \mathbb{R} \\ x &\mapsto 2 + \sqrt{x} \end{aligned}$$

$$\mathrm{Dom}(f) = [0, \infty), \quad \mathrm{Codom}(f) = \mathbb{R}, \quad \mathrm{Range}(f) = [2, \infty)$$

**Remarks.**

- $f$ denotes a function, while $f(x) \in B$ is a number, namely the value of $f$ at the point $x \in A$.

- The codomain of $f$ may be changed but it must contain all the outputs of $f$.

- The statement

  $$\text{‘}f(x) = \sqrt{x} \text{ for all } x \text{ in } [0, \infty)\text{’}$$

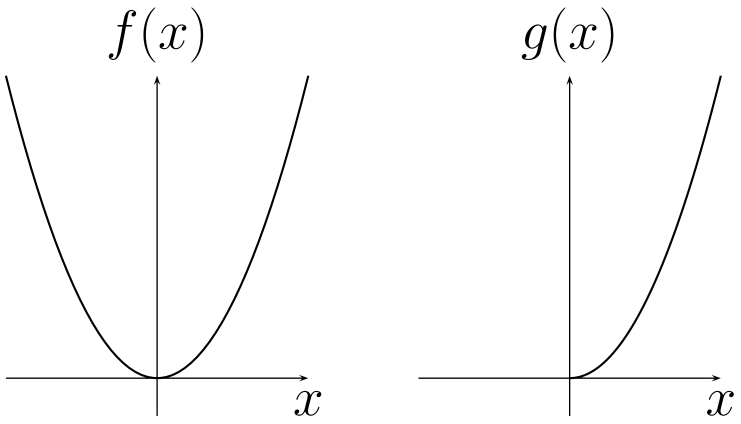
  may be abbreviated as

  $$f(x) = \sqrt{x} \qquad \forall x \in [0, \infty).$$

- Functions which are defined by the same rule but have different domains are not the same.

**Example.**

$$
\begin{aligned}
f : \mathbb{R} \to \mathbb{R}, &\quad f(x) = x^2 \\
g : [0, \infty) \to \mathbb{R}, &\quad g(x) = x^2
\end{aligned}
$$




$g$ is invertible, while $f$ is not!

**Natural domain.** If, for whatever reason, the domain of a function is not defined then we may choose the natural domain or maximal domain, that is the largest possible domain for which the rule makes sense (for real numbers).

**Examples.** Find the maximal domain for

(a)

$$f(x) = \sqrt{x - 1/x}.$$

(b)

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

**Remark.** We distinguish between the range and the codomain of a function since we may not actually be able to determine the range of a function.

**Exercise.** What is the range of the function $f$ defined by

$$f : \mathbb{R} \to [0, \infty), \quad f(x) = x^2 + e^x\,?$$

**Combining functions.** If $f$ and $g$ are two functions with the same domain, then one can combine $f$ and $g$ to form new functions.

**Definition.** Suppose that $f : A \to B$ and $g : A \to B$ are real-valued functions. Then, the functions $f + g$, $f - g$, $f.g$ and $f/g$ are defined by the rules

$$\begin{aligned}
(f+g)(x) &= f(x) + g(x) && \forall x \in A \\
(f-g)(x) &= f(x) - g(x) && \forall x \in A \\
(f.g)(x) &= f(x)g(x) && \forall x \in A \\
(f/g)(x) &= \frac{f(x)}{g(x)} && \forall x \in A^*,
\end{aligned}$$

where $A^* = A \setminus \{x \in A : g(x) = 0\}$.

**Example.** Determine the maximal domains of $f/g$ and $g/f$, where

$$f(x) = 1 + x^2, \quad g(x) = \cos(x).$$

Another way of constructing new functions is given below.

**Definition.** Suppose that

$$f : C \to D \qquad \text{and} \qquad g : A \to B$$

are functions such that $\mathrm{Range}(g)$ is a subset of $\mathrm{Dom}(f)$. Then the composition

$$f \circ g : A \to D$$

is defined by the rule

$$(f \circ g)(x) = f(g(x)) \qquad \forall x \in A.$$

**Example.** Suppose that the functions $f$ and $g$ are given by the rules

$$f(x) = \ln x, \quad g(x) = \sqrt{4 - x^2}$$

What is the maximal domain of the composition $f \circ g$?

## 1.5 Polynomials and rational functions

**Polynomials.** A function $f : \mathbb{R} \to \mathbb{R}$ is called a polynomial if

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0,$$

where $n \in \mathbb{N}$ is the degree and the coefficients $a_0, a_1, \ldots, a_n$ are real numbers with the leading coefficient $a_n \neq 0$.

**Rational functions.** Suppose that $p$ and $q$ are polynomials. The function $f$ defined by the rule

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

with

$$\mathrm{Dom}(f) = \{x \in \mathbb{R} : q(x) \neq 0\}$$

is called a rational function.

**Previous example.** $f$ defined by

$$f(x) = \frac{1}{x^2 + x - 2}, \qquad \mathrm{Dom}(f) = \mathbb{R} \setminus \{-2, 1\}$$

is rational.

**Remark.** The function defined by

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

is also rational and $\mathrm{Dom}(f) = \mathbb{R}$.

## 1.6 The trigonometric functions

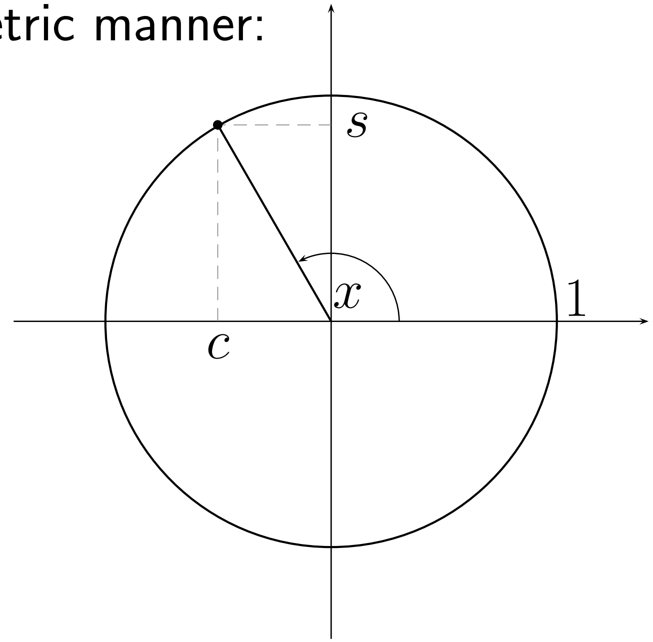
**Definition.** The trigonometric functions

$$\sin : \mathbb{R} \to \mathbb{R} \qquad \text{and} \qquad \cos : \mathbb{R} \to \mathbb{R}$$

are defined by

$$\sin x = s \qquad \text{and} \qquad \cos x = c,$$

where $s$ and $c$ are defined in a geometric manner:




**Question.** Is it obvious that

$$\frac{d}{dx} \sin x = \cos x\,?$$

The following properties are evident.

- $\text{Range}(\sin) = \text{Range}(\cos) = [-1, 1]$.
- $\sin$ and $\cos$ are periodic of period $2\pi$, that is

$$\sin(x + 2\pi) = \sin x, \quad \cos(x + 2\pi) = \cos x.$$

- $\cos$ is an even function, that is

$$\cos(-x) = \cos x.$$

- $\sin$ is an odd function, that is

$$\sin(-x) = -\sin x.$$

- $\sin^2 x + \cos^2 x = 1$.

Other trigonometric functions with suitable domains are defined by

$$\tan x = \frac{\sin x}{\cos x}, \qquad \cot x = \frac{\cos x}{\sin x}$$
$$\sec x = \frac{1}{\cos x}, \quad \operatorname{cosec} x = \frac{1}{\sin x}.$$

The six trigonometric functions are related by various identities and formulae (which you are supposed to know):

- complementary identities

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

- Pythagorean identities

$$\cos^2 x + \sin^2 x = 1$$
$$1 + \tan^2 x = \sec^2 x$$
$$\cot^2 x + 1 = \operatorname{cosec}^2 x$$

- the sum and difference formulae

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$$
$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y$$
$$\tan(x \pm y) = \frac{\tan x \pm \tan y}{1 \mp \tan x \tan y}$$

- double-angle formulae

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

## 1.7 The elementary functions

The elementary functions are all those functions that can be constructed by combining (a finite number of) polynomials, exponentials, logarithms, roots and trigonometric functions (including the inverse trigonometric functions) via function composition, addition, subtraction, multiplication and division. Hence the following expressions give rise to elementary functions:

$$f(x) = e^{\sin x} + x^2,$$

$$g(x) = \frac{\ln x - \tan x}{\sqrt{x}},$$

$$h(x) = \sqrt[3]{x^4 - 2x^2 + 5},$$

$$k(x) = |x| = \sqrt{x^2}.$$

It also follows that every rational function is an elementary function.

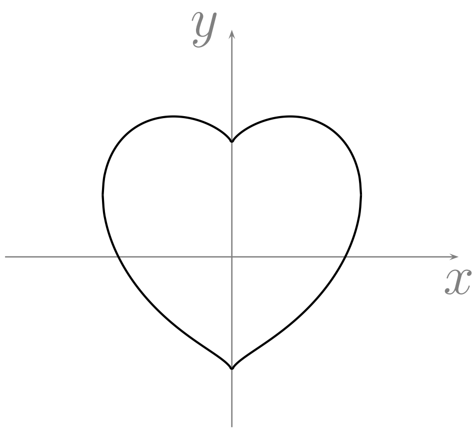
However, there exist important functions which are not elementary!

## 1.8 Implicitly defined functions

Many curves on the plane can be described as all those points $(x, y)$ on the plane that satisfy some equation involving $x$ and $y$. For example, consider the equation

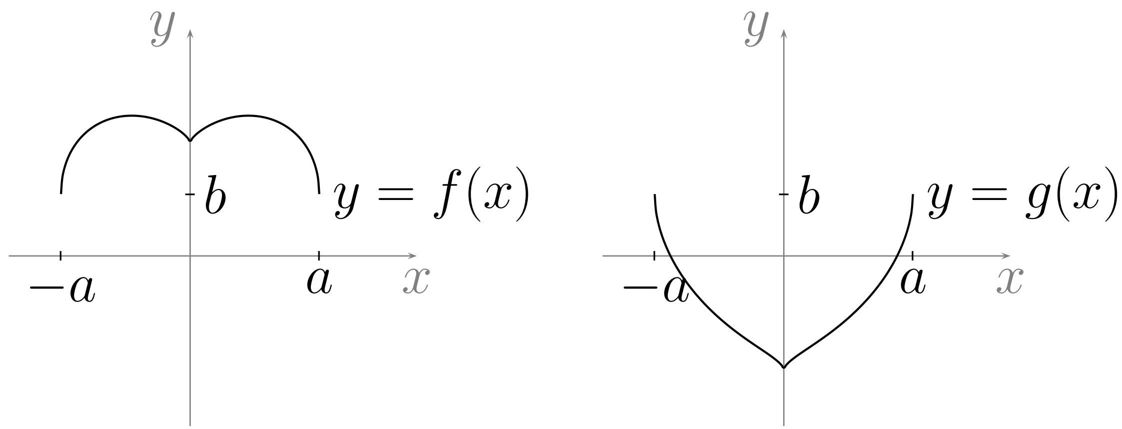
$$(x^2 + y^2 - 1)^3 - x^2y^3 = 0. \qquad (\heartsuit)$$

The set of points $(x, y)$ satisfying this equation are shown on the graph below.




**Properties.**

- There exist several $y$-values for some $x$-values. Hence, the curve cannot be the graph of one function of $x$.
- The curve may be decomposed into two curves which may be regarded as the graphs of two functions, $f$ and $g$, say.




- We say that the functions $f$ and $g$ are implicitly defined by the relation ($\heartsuit$).

  [In this case, can one determine formulae for these functions $f$ and $g$?]

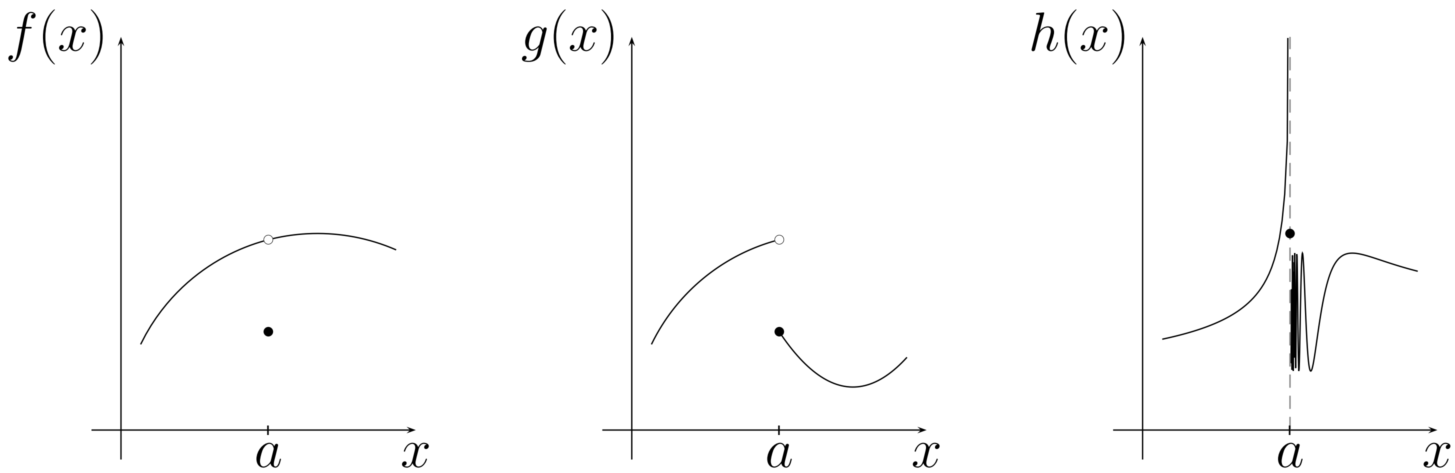
# Chapter 2

# Limits

**Question.** How would you define continuity?

**'Intuitive' (incorrect) answer.** The function is continuous if 'its graph can be drawn without lifting the pencil off the page'.

**Rigorous (correct) answer.** Via limits.



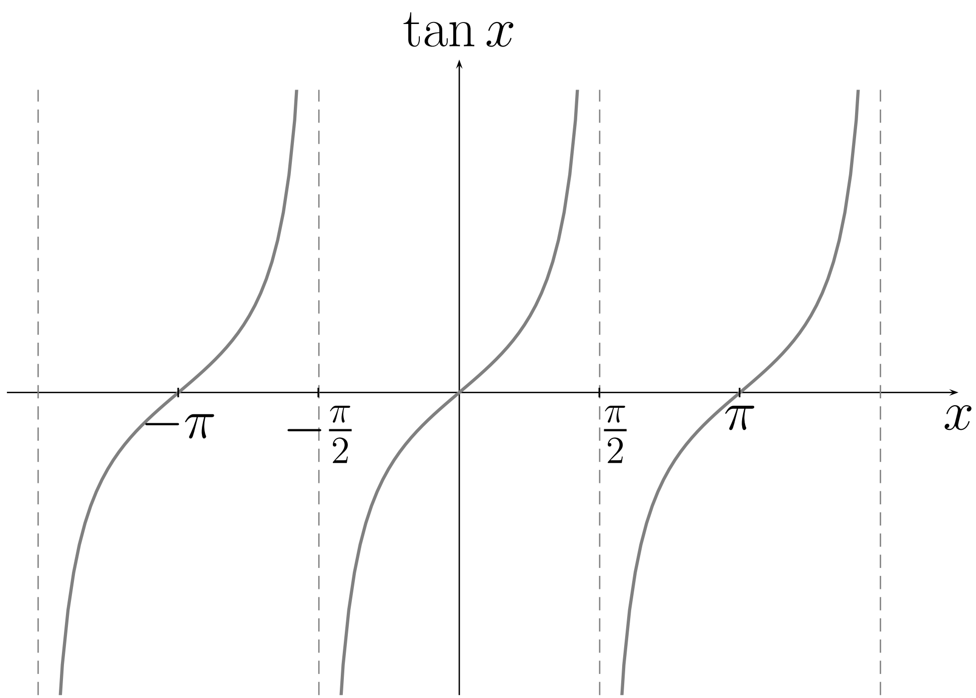

- $f$ has a removable discontinuity at $x = a$.
- $g$ has a jump discontinuity at $x = a$.
- $h$ has an essential discontinuity at $x = a$.

**'Counterexample'.** Consider the function

$$\tan : A \to \mathbb{R}$$

with

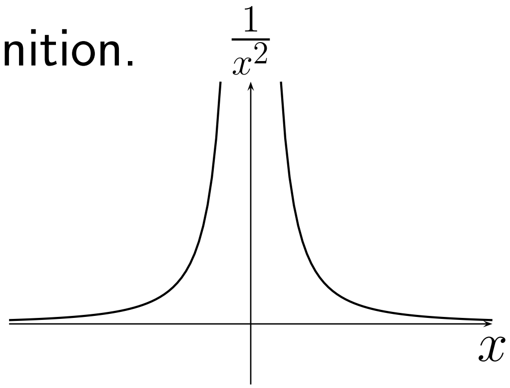
$$A = \mathrm{Dom}(f) = \{x \in \mathbb{R} : x \neq \tfrac{\pi}{2} + n\pi,\ n \in \mathbb{Z}\}.$$

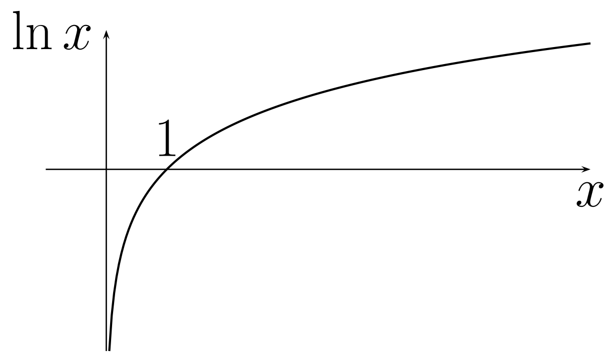



The function $\tan$ is continuous everywhere! The break in the graph is merely due to the 'missing' points in the domain $A$.

## 2.1 Limits of functions at infinity

In this section, we examine some techniques for calculating limits of the form $\lim_{x\to\infty} f(x)$ and postpone their precise definition.




**Notation.**

- If $f(x)$ gets 'arbitrarily close' to some real number $L$ as $x$ tends to infinity, then we write

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

  or

$$f(x) \to L \quad \text{as} \quad x \to \infty.$$

- If $f(x)$ gets 'arbitrarily large' (that is, 'approaches' $\infty$) as $x$ tends to $\infty$ then we write

$$f(x) \to \infty \quad \text{as} \quad x \to \infty.$$

- We never write $\lim_{x\to\infty} f(x) = \infty$ since $\infty$ is not a real number.

### 2.1.1 Basic rules for limits

**'Elementary' rules.**

- If $f$ is a constant function, that is $f(x) = c$ for all $x$, then

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

- If $f(x) \to \infty$ as $x \to \infty$ then

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

These are intuitively obvious and give limits such as

$$\lim_{x\to\infty} \frac{1}{x} = 0, \qquad \lim_{x\to\infty} e^{-x} = \lim_{x\to\infty} \frac{1}{e^x} = 0.$$

**Theorem.** Suppose that

$$\lim_{x\to\infty} f(x) = a, \qquad \lim_{x\to\infty} g(x) = b$$

for some functions $f$ and $g$. Then

- $\lim\limits_{x\to\infty} [f(x) + g(x)] = a + b$
- $\lim\limits_{x\to\infty} [f(x) - g(x)] = a - b$
- $\lim\limits_{x\to\infty} [f(x)g(x)] = ab$
- $\lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)} = \dfrac{a}{b}$ provided that $b \neq 0$.

**Proof.** Later.

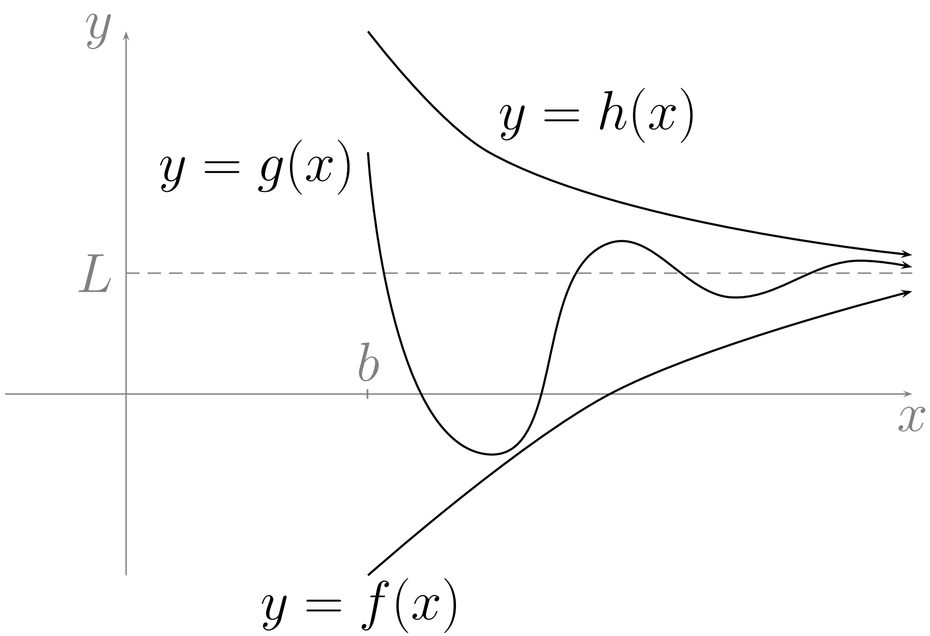
**Example.** Determine the limit of

$$f(x) = \frac{1 + \dfrac{2}{3x+4}}{5 - 6e^{-x}}$$

as $x \to \infty$.

### 2.1.2 The pinching theorem

**Idea.** Assume that two functions $f$ and $h$ have the same limit as $x \to \infty$ and the graph of a function $g$ lies between the graphs of $f$ and $h$ (if $x$ is large enough). Then, $g$ has the same limit as $f$ and $h$.



**The pinching theorem.** Suppose that $f$, $g$ and $h$ are three functions such that

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

on an interval $(b, \infty)$ for some $b \in \mathbb{R}$ and

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

Then

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

**Example.** Determine the limit of

$$g(x) = e^{-x^2} 2^{\sin 3x}$$

as $x \to \infty$.

### 2.1.3 Limits of the form $f(x)/g(x)$

Suppose that we want to calculate a limit of the form

$$\lim_{x \to \infty} \frac{f(x)}{g(x)},$$

where both $f(x)$ and $g(x)$ tend to infinity as $x \to \infty$.

**Problem.** We cannot apply the preceding rules since $f$ and $g$ do not have limits.

**Idea.** Divide both $f$ and $g$ by the leading term, that is the fastest growing term appearing in the denominator $g$ (if it exists).

**Examples.** Find the following limits (if they exist).

(a)

$$\lim_{x\to\infty} \frac{5x^3 + 6x^2 - 4\sin x}{\cos 3x + 5x - 2x^3}$$

(b)

$$\lim_{x\to\infty} \frac{x^2 + 3x}{\sqrt{2x^4 + 3} - 4x}$$

### 2.1.4 Limits of the form $\sqrt{f(x)} - \sqrt{g(x)}$

In this case, one needs to be very careful since a similar 'leading term argument' may be erroneous!

In fact, it is the 'lower order' terms that may determine the limit!

**Example.** Determine the limit of

$$f(x) = \sqrt{2x^2 + 3x} - \sqrt{2x^2 - x}$$

as $x \to \infty$.

**Exercise.** Does

$$\lim_{x\to\infty} \sqrt{x^4 - x^3} - \sqrt{x^4 + 1}$$

exist?

### 2.1.5 Indeterminate forms

The following limits have the form "$\frac{\infty}{\infty}$" but each displays a very different limiting behaviour as $x \to \infty$:

- $\dfrac{x^2}{x} \to \infty$
- $\dfrac{x}{x^2} \to 0$
- $\dfrac{2x^2}{x^2} \to 2$

Since we cannot determine in advance what kind of limiting behaviour something of the form "$\frac{\infty}{\infty}$" has, we say that "$\frac{\infty}{\infty}$" is an indeterminate form.

Other types of indeterminate forms are

- "$\frac{0}{0}$"
- "$\infty - \infty$"
- "$0 \times \infty$"

and appear in various applications.

## 2.2 The definition of $\lim_{x \to \infty} f(x)$

**Example.** Why do we believe that

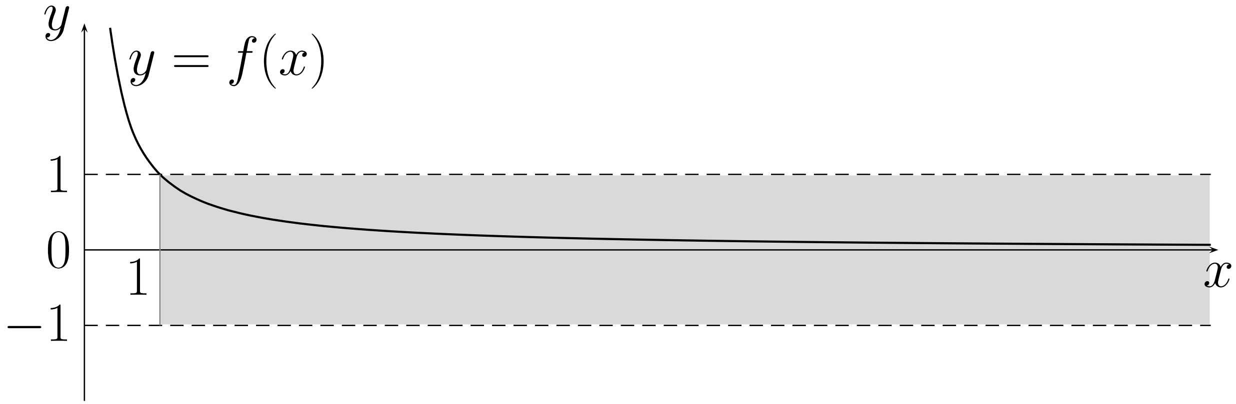
$$\lim_{x\to\infty} f(x) = 0 \quad \text{for} \quad f(x) = \frac{1}{x}\,?$$

Consider the distance between $f(x)$ and $0$ denoted by

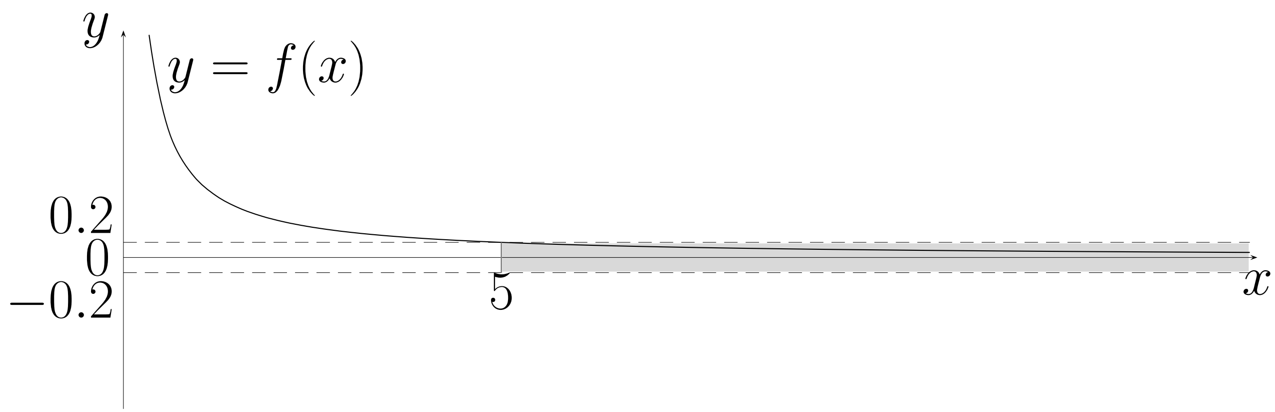
$$d(x) = |f(x) - 0|.$$

**Facts.**

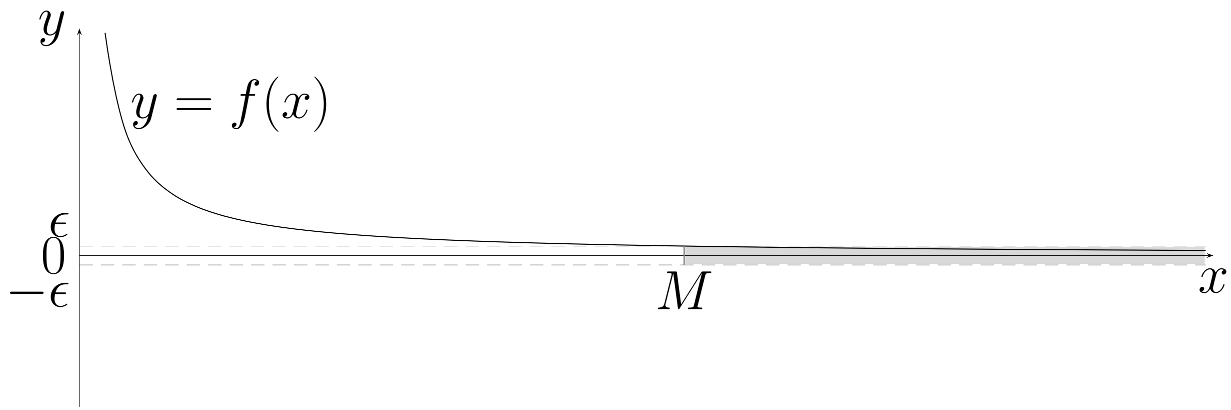
- $d(x) < 1$ whenever $x > 1$.



- $d(x) < 0.2$ whenever $x > 5$.




- $d(x) < 0.1$ whenever $x > 10$.
- $d(x) < 0.01$ whenever $x > 100$.
- $d(x) < 0.0001$ whenever $x > 10000$.
- Set $\epsilon = 1/M$. Then, $d(x) < \epsilon$ whenever $x > M$.



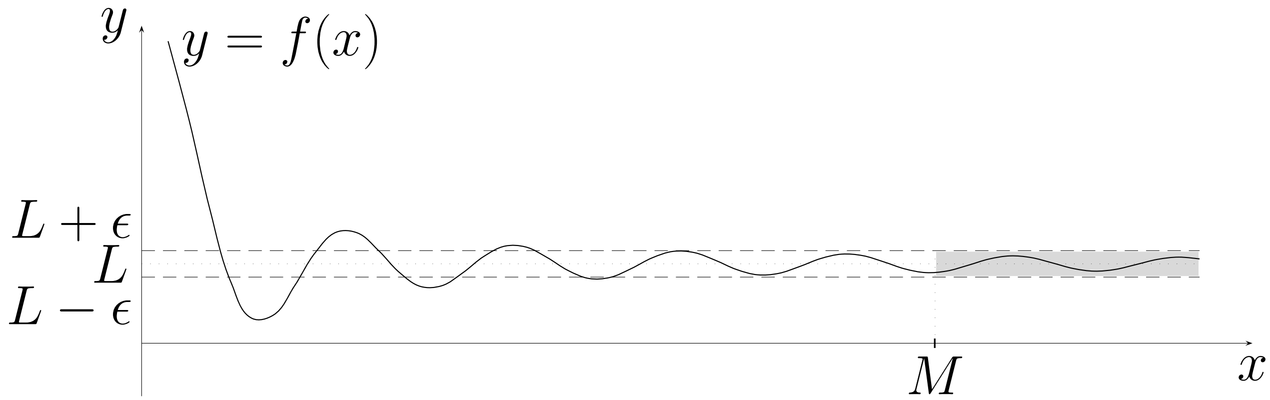
**Definition.** Let $f$ be a function defined on some interval $(b, \infty)$ and let $L$ be a real number. We say that

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

if

for every $\epsilon > 0$, there exists a real number $M$ such that

$$\text{if } x > M \quad \text{then} \quad |f(x) - L| < \epsilon.$$



## 2.3 Proving that $\lim_{x\to\infty} f(x) = L$ using the limit definition

**Example.** Prove that

$$\lim_{x\to\infty} \frac{5x}{x+3} = 5.$$

**Proof.** We consider the distance

$$\begin{aligned}
|f(x) - L| &= \left|\frac{5x}{x+3} - 5\right| \\
&= \left|\frac{5x - 5(x+3)}{x+3}\right| \\
&= \left|\frac{-15}{x+3}\right| \\
&= \frac{15}{x+3} && \text{for } x > -3 \\
&< \frac{15}{x} && \text{for } x > 0.
\end{aligned}$$

In summary,

$$|f(x) - L| < \frac{15}{x}.$$

This inequality gives an upper bound for the distance between $f(x)$ and $L$! Accordingly,

$$|f(x) - L| < \epsilon \quad \text{whenever} \quad \frac{15}{x} < \epsilon.$$

The latter condition is equivalent to

$$x > \frac{15}{\epsilon}$$

and hence if we set

$$M = \frac{15}{\epsilon}$$

then

$$|f(x) - L| < \epsilon \quad \text{whenever} \quad x > M,$$

as required.

**General strategy.** Given $\epsilon$, we need to find a number $M$ such that

$$|f(x) - L| < \epsilon \quad \text{whenever} \quad x > M.$$

The number $M$ can be found by following the procedure below.

1. Find a good upper bound for $|f(x) - L|$.
2. Find a simple condition on $x$ such that this upper bound is less than $\epsilon$.
3. Use this condition to state an appropriate value for $M$ (in terms of $\epsilon$).

**Example.** Show that

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

**Remarks.**

- In general, $M$ depends on $\epsilon$ but it is not uniquely defined.
- The definition of the limit does not tell you what the limit is. The definition may be used to prove theorems which allow you to justify methods of finding limits.
- Applying the definition to verify an educated guess for a limit is usuallly the last resort. Make use of the theorems unless you are specifically asked to apply the definition.

## 2.4 Proofs of basic limit results

**Proof of the limit of a sum of two functions.** Suppose that

$$\lim_{x\to\infty} f(x) = L_1, \qquad \lim_{x\to\infty} g(x) = L_2$$

and set

$$\epsilon_1 = \epsilon_2 = \frac{\epsilon}{2}$$

for any given $\epsilon > 0$.

By definition of the limit, there exist numbers $M_1$ and $M_2$ such that

$$|f(x) - L_1| < \epsilon_1 \quad \text{whenever} \quad x > M_1$$

and

$$|g(x) - L_2| < \epsilon_2 \quad \text{whenever} \quad x > M_2.$$

Hence, by the triangle inequality,

$$\begin{aligned}\Big|[f(x)+g(x)]-(L_1+L_2)\Big| &\le |f(x)-L_1|+|g(x)-L_2| \\ &< \epsilon_1+\epsilon_2 \\ &= \epsilon\end{aligned}$$

whenever both $x > M_1$ and $x > M_2$.

If we set $M = \max\{M_1, M_2\}$ then

$$\Big|[f(x)+g(x)]-(L_1+L_2)\Big| < \epsilon \quad \text{whenever} \quad x > M$$

so that

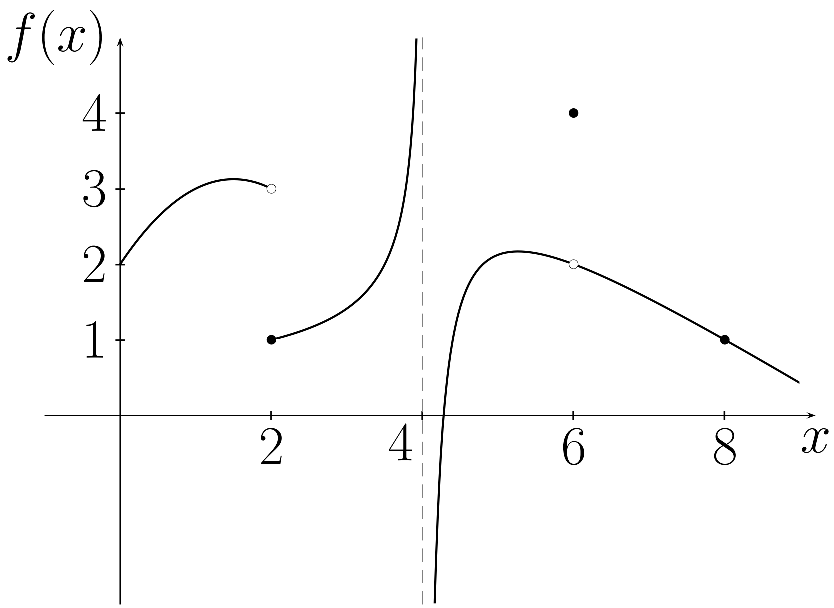
$$\lim_{x\to\infty}[f(x)+g(x)] = L_1+L_2.$$

The other rules may be proven in a similar manner.

**Proof of the pinching Theorem.** See page 37 of the Calculus Notes.

### 2.4.1 Left-hand, right-hand and two-sided limits

**Example.** Consider the function $f$ whose graph is shown below.




With reference to this graph, we will discuss the behaviour of $f(x)$ when $x$ is near the points $2$, $4$, $6$ and $8$.

**Notation.**

- Let $f$ be a function defined on an interval $(c, a)$. We say that $L$ is the left-hand limit of $f(x)$ as $x$ approaches $a$ if $f(x)$ gets 'closer and closer' to $L$ when $x$ gets 'closer and closer' to $a$ from the left. We write

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

- Let $f$ be a function defined on an interval $(a, b)$. We say that $L$ is the right-hand limit of $f(x)$ as $x$ approaches $a$ if $f(x)$ gets 'closer and closer' to $L$ when $x$ gets 'closer and closer' to $a$ from the right. We write

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

- Let $f$ be a function with the property

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

for some point $a \in \mathrm{Dom}(f)$. Then, we say that $L$ is the (two-sided) limit of $f(x)$ as $x$ approaches $a$ and we write

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

**Remark.** In all cases, the value of $f$ at $x = a$ (if defined) is irrelevant!

- For $a = 2$:

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

  The two-sided limit does not exist.

- For $a = 4$:

$$f(x) \to \pm\infty \quad \text{as} \quad x \to 4^{\mp}$$

  No limit exists.

- For $a = 6$:

$$\lim_{x \to 6^-} f(x) = \lim_{x \to 6^+} f(x) = 2, \quad f(6) = 4$$

  The two-sided limit exists but does not coincide with the value of $f$ at $a = 6$.

- For $a = 8$:

$$\lim_{x \to 8^-} f(x) = \lim_{x \to 8^+} f(x) = f(8) = 1$$

  The two-sided limit exists and coincides with the value of $f$ at $a = 8$.

**Question.** What is so special about the above function at $x = 8$?

### 2.4.2 Limits and continuous functions

**Definition.** Let $f$ be defined on some open interval containing the point $a$. We say that $f$ is continuous at $a$ if

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

otherwise we say that $f$ is discontinuous at $a$.

If $f$ is continuous at every point of its domain, we simply say that $f$ is continuous.

**Previous example.** The function $f$ is continuous everywhere except at $x = 2$ and $x = 6$.

Note that $x = 4$ is not part of the domain of $f$ and hence asking whether or not $f$ is continuous at $x = 4$ does not make any sense.

**Remark.** A rigorous definition of the preceding limits exists but, here, we will focus on rules for calculating limits.

### 2.4.3 Rules for limits at a point

Polynomials, rational functions, the trigonometric functions, exponentials and logarithms are continuous at every point in the respective domains.

Limits at a point also have nice arithmetic properties ...

**Theorem.** Suppose that

$$\lim_{x\to a} f(x) = L_1, \qquad \lim_{x\to a} g(x) = L_2$$

for some functions $f$ and $g$. Then

- $\lim_{x\to a}[f(x) + g(x)] = L_1 + L_2$
- $\lim_{x\to a}[f(x) - g(x)] = L_1 - L_2$
- $\lim_{x\to a}[f(x)g(x)] = L_1 L_2$
- $\lim_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{L_1}{L_2}$ provided that $L_2 \neq 0$.

... and interact nicely with function composition.

**Theorem.** If $\lim_{x \to a} f(x) = L$ and $g$ is continuous at $L$ then

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

**Example.** Find the limit

$$\lim_{x \to 5} \sqrt{2x + \cos(\pi x^2)}\,.$$

**The pinching theorem.** Suppose that $f$, $g$ and $h$ are all defined on an open interval $I$ containing the point $a$. If

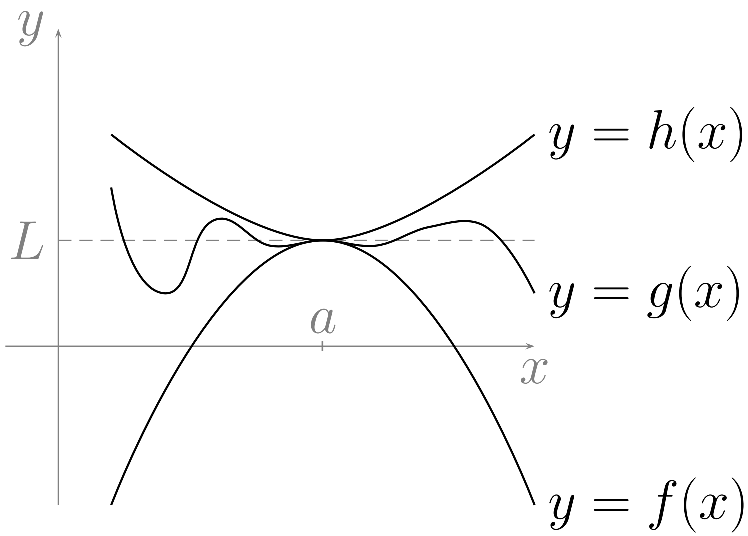
$$f(x) \leq g(x) \leq h(x) \qquad \forall x \in I$$

and

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

then

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



**Example.** Is the function $f$ defined by

$$f(x) = \begin{cases} \dfrac{\sin x}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$$

continuous?

**Concluding example.** Suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined by the formula

$$f(x) = \begin{cases} \dfrac{|x^2 - 9|}{x - 3} & \text{if } x \neq 3 \\ -6 & \text{if } x = 3. \end{cases}$$

Discuss the limiting behaviour of $f(x)$ as $x$ approaches $3$.

# Chapter 3

# Properties of continuous functions

## 3.1 Combining continuous functions

**Question.** If two continuous functions are combined via function addition, subtraction, multiplication, division or composition, when is the resulting function also continuous?

**Theorem.** Suppose that the functions $f$ and $g$ are continuous at a point $a$. Then

$$f+g, \quad f-g, \quad fg$$

are continuous at $a$. If $g(a) \neq 0$ then

$$f/g$$

is also continuous at $a$.

**Proof.** Suppose that $f$ and $g$ are continuous at $a$. Then,

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

by the definition of continuity at a point. Therefore,

$$\begin{aligned}
\lim_{x\to a}(f+g)(x) &= \lim_{x\to a}(f(x)+g(x)) && \text{(def. of } f+g\text{)}\\
&= \lim_{x\to a} f(x) + \lim_{x\to a} g(x) && \text{(limit rule)}\\
&= f(a)+g(a) && (f, g \text{ cont.})\\
&= (f+g)(a) && \text{(def. of } f+g\text{)}.
\end{aligned}$$

Hence $f+g$ is continuous at $a$.

The proofs that the functions $f-g$, $fg$ and $f/g$ are continuous at $a$ are similar.

**Example.** Based on the fact that constant functions and the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x$ are continuous at every point, show that polynomials and rational functions are continuous at every point of their respective domains.

**Solution.** Any polynomial can be obtained from $f$ and constant functions via addition and multiplication, e.g.

$$x^3 - 4x^2 + 5 = [(x \times x \times x)] + [(-4) \times x \times x] + 5,$$

and hence is continuous everywhere.

Any rational function is of the form

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

where $p$ and $q$ are two (continuous) polynomials, and is therefore continuous at every point $a$ for which $q(a) \neq 0$.

If we can prove that trigonometric functions are continuous at every point of their domain then functions defined by rules such as

$$f(x) = \frac{1 + \tan x}{1 + x^2 + x^6}$$

are likewise continuous on the appropriate domain.

Thus, it suffices to verify the following Lemma:

**Lemma.** Show that the functions $\sin$ and $\cos$ are continuous everywhere.

**Interlude.** Why is the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = x$$

continuous everywhere?

**Answer.** Prescribe $a \in \mathbb{R}$ and $\epsilon > 0$. Define $\delta = \epsilon$ and assume that

$$|x - a| < \delta.$$

Then,

$$|f(x) - f(a)| < \epsilon.$$

Thus, we conclude that

$$\text{for every } \epsilon > 0\text{, there exists a number } \delta > 0 \text{ such that}$$
$$|f(x) - f(a)| < \epsilon \quad \text{whenever} \quad |x - a| < \delta.$$

The latter is the formal definition of continuity of any function $f$ at a point $a$.

Accordingly, the particular function considered above is continuous everywhere.

Even larger classes of continuous functions may be obtained in the following manner:

**Theorem.** Suppose that $f$ is continuous at $a$ and that $g$ is continuous at $f(a)$. Then $g \circ f$ is continuous at $a$.

**Proof.**

$$\begin{aligned} \lim_{x \to a} (g \circ f)(x) &= \lim_{x \to a} (g(f(x)) && \text{(def. of } g \circ f\text{)} \\ &= g(\lim_{x \to a} f(x)) && \text{(Th. of Ch. 2)} \\ &= g(f(a)) && (f \text{ cont.}) \\ &= (g \circ f)(a). && \text{(def. of } g \circ f\text{)} \end{aligned}$$

Hence $g \circ f$ is continuous at $a$.

**Example.** Suppose that the function $g : \mathbb{R} \to \mathbb{R}$ is continuous at the point $a$ and let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by

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

Show that $f$ is continuous at the point $a$.

## 3.2 Continuity on intervals

Strictly speaking, the previous example requires a natural extension of continuity to endpoints. We first recall the standard definition ...

**Definition.** Suppose that $f$ is a real-valued function defined on an open interval $(a, b)$. We say that $f$ is continuous on $(a, b)$ if $f$ is continuous at every point in the interval $(a, b)$.

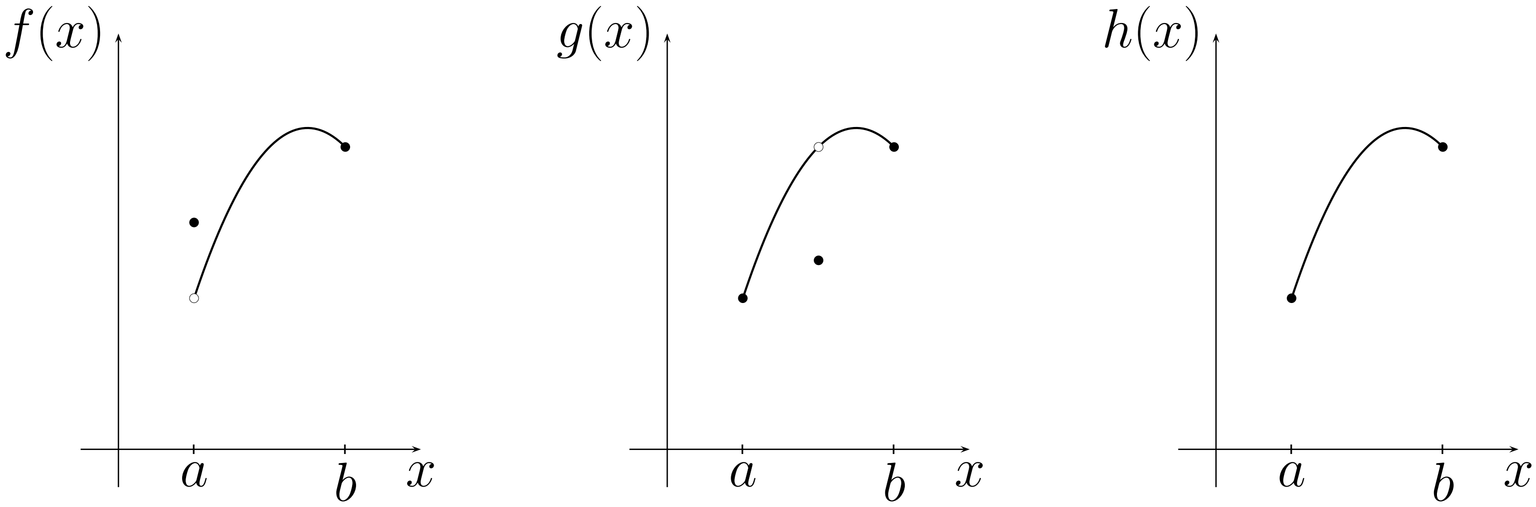
... and make the following generalisation.

**Definition.** Suppose that $f$ is a real-valued function defined on a closed interval $[a, b]$. We say that

- $f$ is continuous at the endpoint $a$ if
$$\lim_{x \to a^+} f(x) = f(a),$$
- $f$ is continuous at the endpoint $b$ if
$$\lim_{x \to b^-} f(x) = f(b),$$
- $f$ is continuous on the closed interval $[a, b]$ if $f$ is continuous on the open interval $(a, b)$ and at each of the endpoints $a$ and $b$.

**Further generalisation.** The function $f : [0, \infty) \to \mathbb{R}$ defined by $f(x) = \sqrt{x}$ is continuous on $[0, \infty)$.

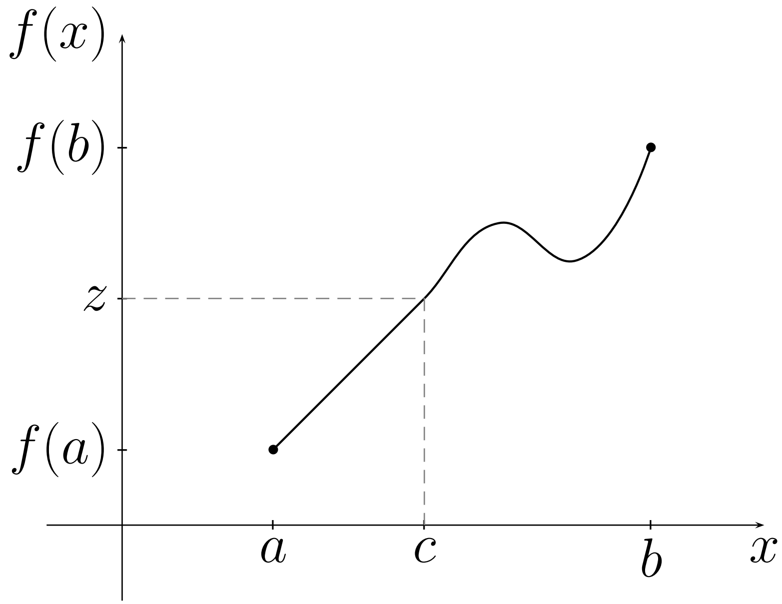
**Example.** Consider the functions $f$, $g$ and $h$, whose graphs are shown below.




All three functions are defined on the interval $[a, b]$.

- $f$ is continuous on the open interval $(a, b)$ and at the endpoint $b$.
- $g$ is continuous at the endpoints $a$ and $b$ but not continuous on the open interval $(a, b)$.
- $h$ is continuous on the closed interval $[a, b]$ (and, by implication, on the open interval $(a, b)$ and at both endpoints $a$ and $b$).

## 3.3 The intermediate value theorem




**The intermediate value theorem.** Suppose that $f$ is continuous on the closed interval $[a, b]$. If $z$ lies between $f(a)$ and $f(b)$ then there exists at least one real number $c$ in $[a, b]$ such that $f(c) = z$.

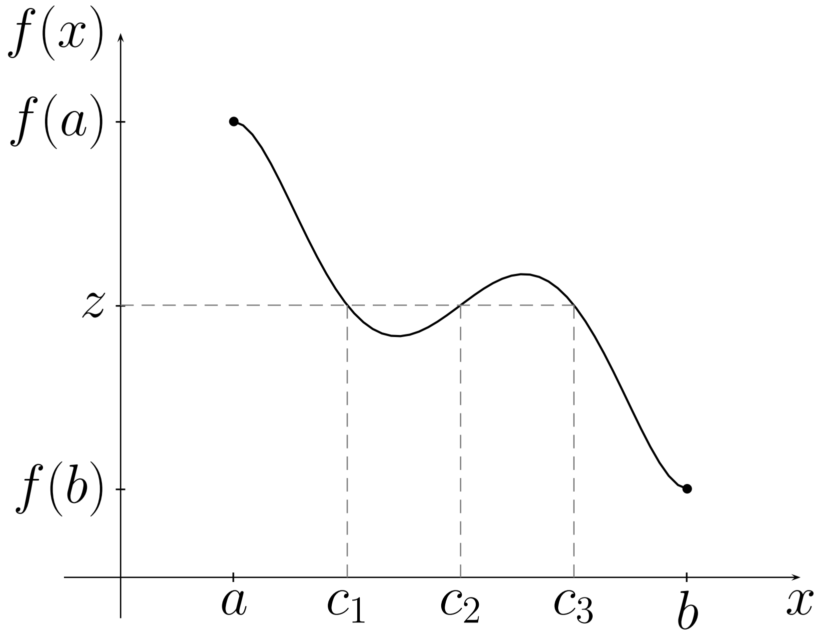
**Proof.** Omitted but based on the the least upper bound property of the real numbers:

If $a_1 \leq a_2 \leq a_3 \leq \cdots \leq M$ is a non-decreasing bounded sequence of real numbers then $a_k$ converges to some real number $L \leq M$ as $k \to \infty$.

(It is this property that distinguishes the real numbers from the rational numbers!)

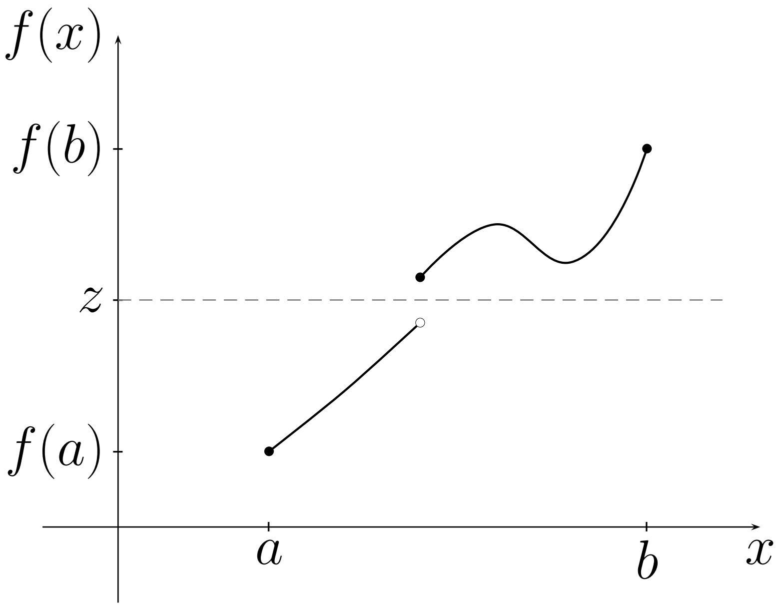
## Remarks.

- The number $c$ in $[a, b]$ may not be unique.




There exists three numbers $c_i$ with $f(c_i) = z$.

- Continuity of $f$ is crucial.




For all $c \in [a, b]$, we see that $f(c) \neq z$.

- The intermediate value theorem is a theorem about real numbers. Consider, for instance, the function

$$f : [0, 2] \to \mathbb{R}, \qquad f(x) = x^2 - 2$$

so that, in particular, $f(0) = -2$ and $f(2) = 2$.

The intermediate value theorem says that there exists a real number $c \in [0, 2]$ such that $f(c) = 0$. In fact,

$$c = \sqrt{2} \notin \mathbb{Q}$$

and, hence, an analogous theorem for rational numbers cannot exist.

**Application.** Show that there exists a solution $c \in [1, 2]$ of the equation

$$\sqrt{c} = c^2 - 1$$

and approximate its value.

**Example.** Suppose that $f : \mathbb{R} \to \mathbb{R}$ is continuous and

$$\lim_{x \to -\infty} f(x) = -1, \quad \lim_{x \to \infty} f(x) = 1.$$

Show that there exists a point $c \in \mathbb{R}$ such that $f(c) = 0$.

## 3.4 The maximum-minimum theorem

**Definition.** Suppose that $f$ is defined on a closed interval $[a, b]$.

- We say that a point $c$ in $[a, b]$ is an absolute minimum point for $f$ on $[a, b]$ if

$$f(c) \leq f(x) \quad \text{for all} \quad x \in [a, b].$$

  The corresponding value $f(c)$ is called the absolute minimum value of $f$ on $[a, b]$. If $f$ has an absolute minimum point on $[a, b]$ then we say that $f$ attains a minimum on $[a, b]$.

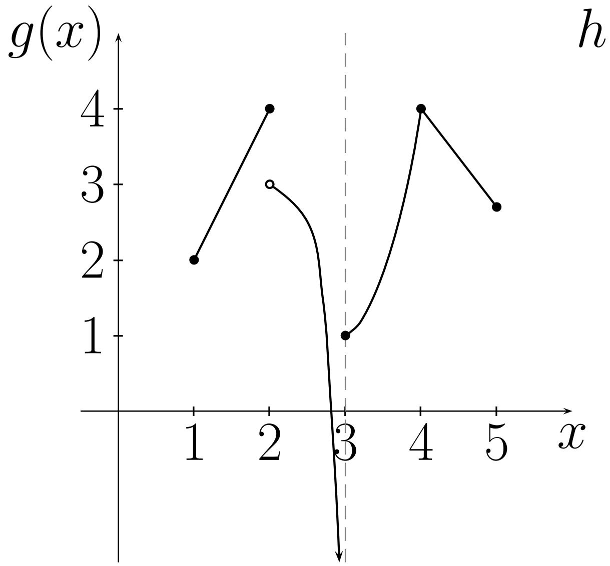
- We say that a point $d$ in $[a, b]$ is an absolute maximum point for $f$ on $[a, b]$ if

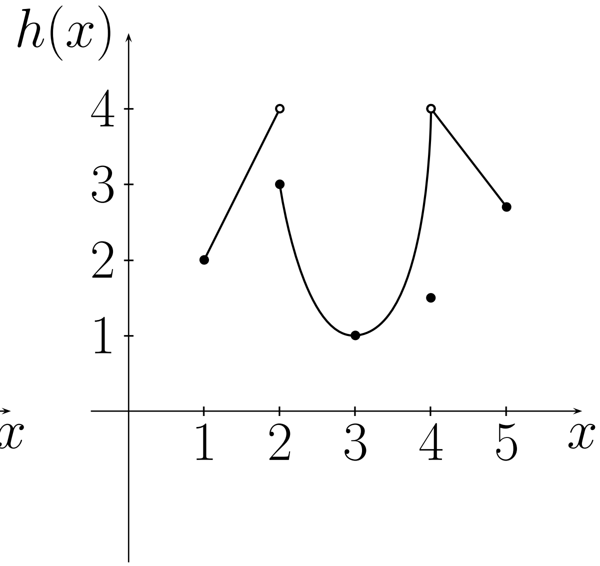
$$f(x) \leq f(d) \quad \text{for all} \quad x \in [a, b].$$

  The corresponding value $f(d)$ is called the absolute maximum value of $f$ on $[a, b]$. If $f$ has an absolute maximum point on $[a, b]$ then we say that $f$ attains a maximum on $[a, b]$.

An absolute maximum point and an absolute minimum point are sometimes referred to as a global maximum point and a global minimum point.

**Example.** Consider the functions $g$ and $h$, whose graphs are illustrated below.




The absolute minimum and maximum points of $g$ and $h$ on $[1, 5]$ are recorded in the following table.

| | $g$ | $h$ |
|---|---:|---:|
| Absolute minimum points | none | $3$ |
| Absolute minimum value | n.a. | $1$ |
| Absolute maximum points | $2$, $4$ | none |
| Absolute maximum value | $4$ | n.a. |

The above example shows that a function $f : [a, b] \to \mathbb{R}$ need not have an absolute maximum point (or an absolute minimum point) on a closed interval $[a, b]$.

**The maximum-minimum theorem.** If $f$ is continuous on a closed interval $[a, b]$ then $f$ attains a minimum and maximum on $[a, b]$. That is, there exist points $c$ and $d$ in $[a, b]$ such that

$$f(c) \leq f(x) \leq f(d)$$

for all $x$ in $[a, b]$.

**Proof.** Omitted but, once again, the least upper bound property of the real numbers is used. The proof relies on the crucial fact that $f$ is continuous on the closed interval $[a, b]$!

**Example.** Suppose that $f : \mathbb{R} \to \mathbb{R}$ is continuous and

$$\lim_{x\to-\infty} f(x) = \lim_{x\to\infty} f(x) = 0$$

Show that $f$ attains a maximum or a minimum (or both) on $\mathbb{R}$.

**Question.** Is there a link between continuity and boundedness?

**Definition.** A function $f$ is said to be bounded on an interval $I$ if there exists some positive number $M$ such that

$$|f(x)| \leq M \quad \text{for all} \quad x \in I.$$

In other words, $f$ is bounded if the $y$-values of its graph lie between $-M$ and $M$ for some positive number $M$.

**Theorem.** If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ then $f$ is bounded on $[a, b]$.

**Proof.** Suppose that $f$ is continuous on $[a, b]$. By the maximum-minimum theorem, $f$ attains a maximum and minimum on $[a, b]$, that is

$$f(c) \leq f(x) \leq f(d)$$

for some $c$ and $d$ and all $x$ in $[a, b]$. If we set

$$M = \max\{|f(c)|, |f(d)|\}$$

then

$$|f(x)| \leq M$$

and hence $f$ is bounded.

**Remark.** A function may be bounded without having a maximum or minimum value.

**Example.** Is the function $f$ defined by

$$f(x) = \begin{cases} \dfrac{1}{x^2+1}\sin\left(\dfrac{1}{x}\right) & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$$

bounded on $[-2, 2]$? Does it attain a minimum or maximum on $[-2, 2]$?

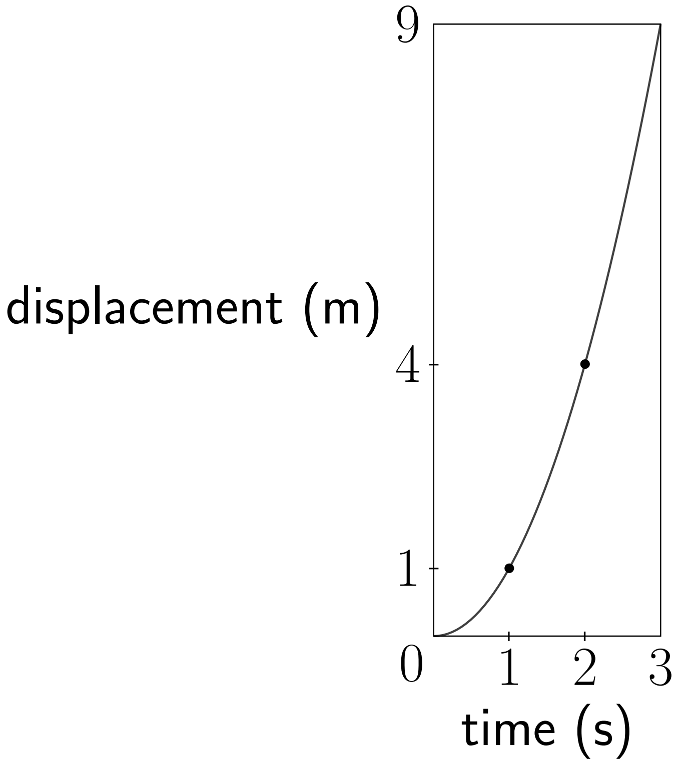
# Chapter 4

# Differentiable functions

**Typical question.** A snowboarder accelerates from rest in such a way that her displacement $s$ (in metres) from her starting position after $t$ seconds is given by

$$s = t^2 \qquad \forall t \in [0, 3].$$

What is the snowboarder's 'instantaneous' speed when $t = 1$?



**'Intuitive solution.'** Calculate the average velocity $\overline{v}$ of the snowboarder over any time interval according to

$$\overline{v} = \frac{\Delta s}{\Delta t},$$

where $\Delta s$ denotes the change in displacement corresponding to a change $\Delta t$ in time ...

| Time interval | $\Delta t$ | Average velocity |
|---|---|---|
| $[1, 1.5]$ | $0.5$ | $2.5$ |
| $[1, 1.4]$ | $0.4$ | $2.4$ |
| $[1, 1.3]$ | $0.3$ | $2.3$ |
| $[1, 1.2]$ | $0.2$ | $2.2$ |
| $[1, 1.1]$ | $0.1$ | $2.1$ |
| $[1, 1.01]$ | $0.01$ | $2.01$ |
| $[1, 1.001]$ | $0.001$ | $2.001$ |

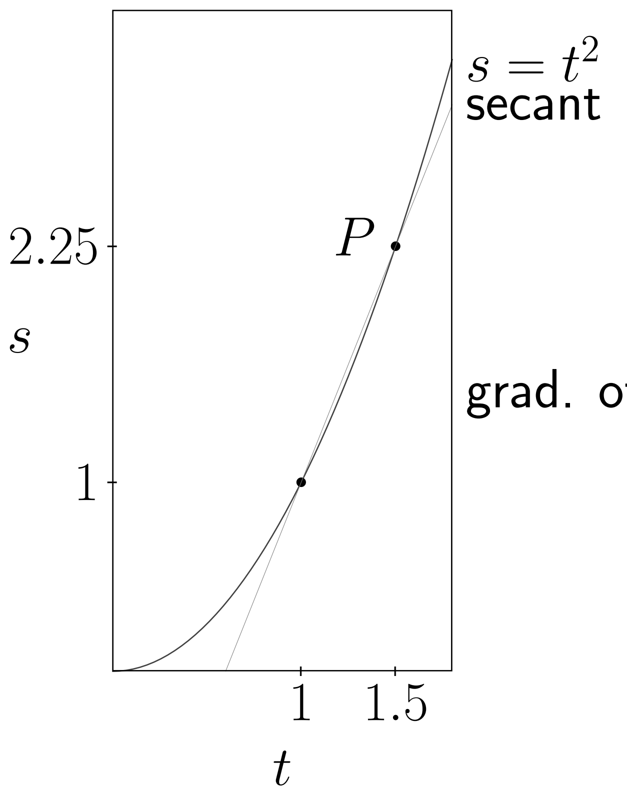
... and let $\Delta t$ 'approach' 0. It appears that

$$\bar{v} \to 2 \quad \text{as} \quad \Delta t \to 0$$

and hence define 2 (meters/second) to be the instantaneous speed at $t = 1$.

**Geometric interpretation.** Find the slope (gradient) of the (tangent line to the) graph at $t = 1$.

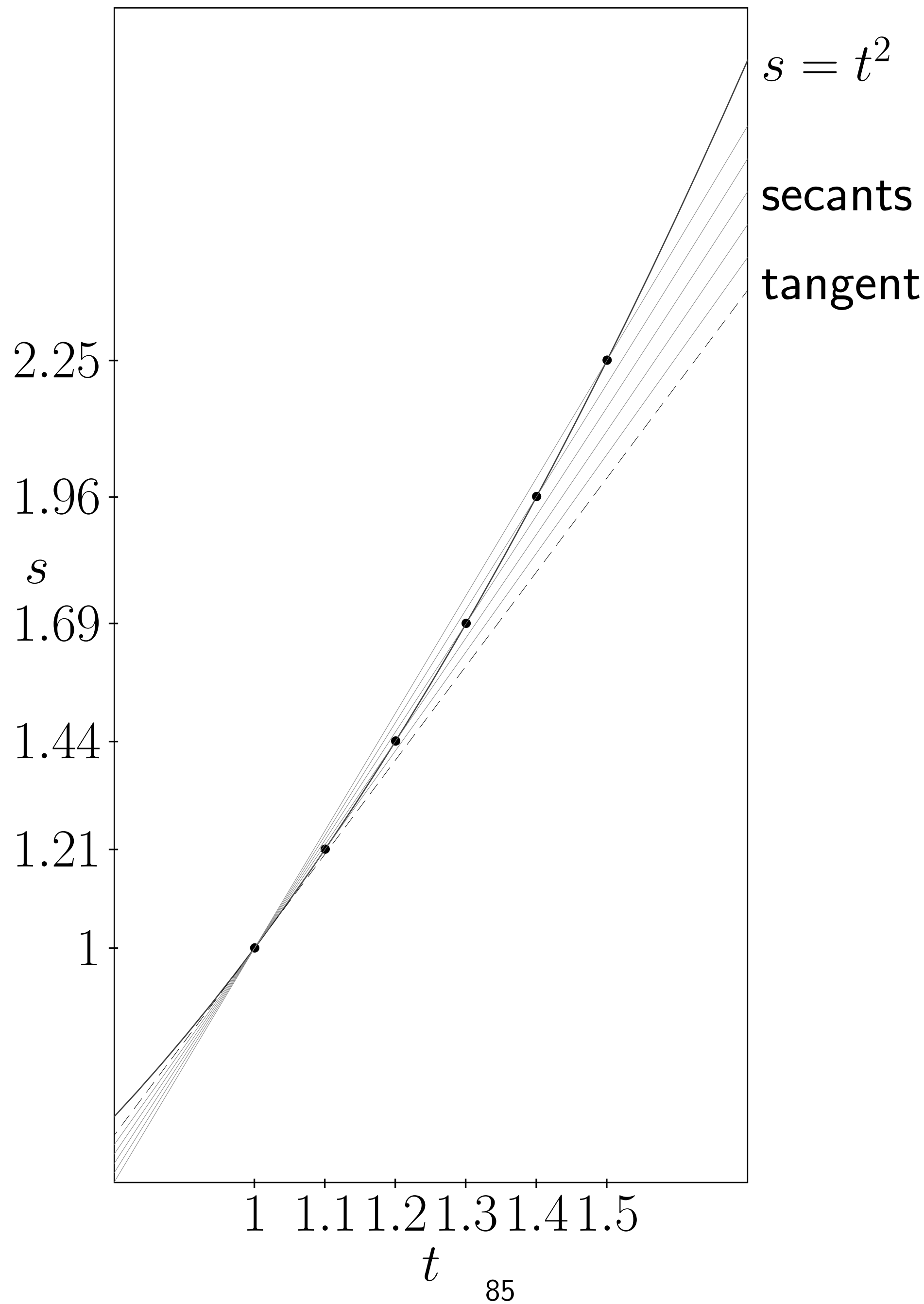
Thus, calculate the gradient of the secant intersecting the graph at $(1, 1)$ and some other point $P$ ...




grad. of secant $= \frac{\Delta s}{\Delta t} = \frac{(1.5)^2 - (1)^2}{1.5 - 1} = 2.5$

| Intersection points | $\Delta t$ | Gradient of secant |
| --- | --- | --- |
| $t = 1, t = 1.5$ | $0.5$ | $2.5$ |
| $t = 1, t = 1.4$ | $0.4$ | $2.4$ |
| $t = 1, t = 1.3$ | $0.3$ | $2.3$ |
| $t = 1, t = 1.2$ | $0.2$ | $2.2$ |
| $t = 1, t = 1.1$ | $0.1$ | $2.1$ |

... and let $P$ 'approach' the point $(1, 1)$.

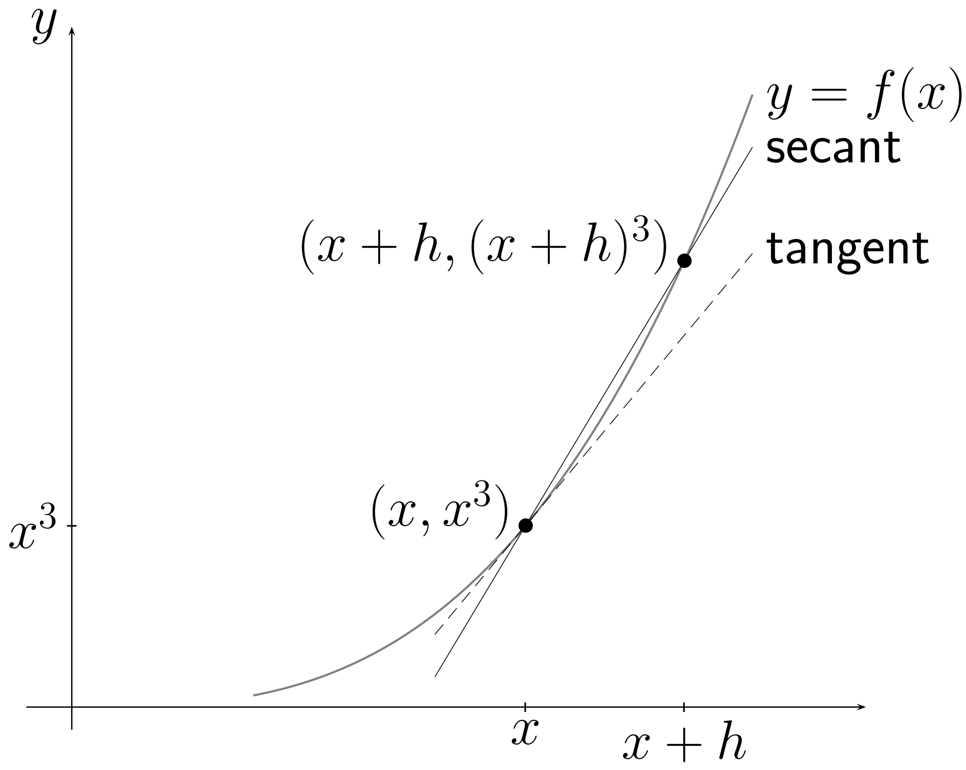


$s = t^2$
secants
tangent
2.25
1.96
$s$
1.69
1.44
1.21
1
1
1.1
1.2
1.3
1.4
1.5
$t$

## 4.1 Gradients of tangents and derivatives

**Example.** Find the gradient of the tangent to the graph of the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$.

$$\begin{aligned}
(\text{grad. at } x) &= \lim_{h\to 0} \frac{(x+h)^3 - x^3}{h} \\
&= \lim_{h\to 0} \frac{3x^2h + 3xh^2 + h^3}{h} \\
&= \lim_{h\to 0} 3x^2 + 3xh \\
&= 3x^2.
\end{aligned}$$



**Definition.** Suppose that $f$ is defined on some open interval containing the point $x$. If

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

exists then it is called the derivative of $f$ at $x$ and we say that $f$ is differentiable at $x$.

The derivative of $f$ at $x$ is denoted by

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

**Remark.** The ratio

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

is called the difference quotient for $f$ at the point $x$.

**Example.** Show that

$$\frac{d}{dx}\cos x = -\sin x.$$

**Exercise.** Write

$$\cos h - 1 = \frac{\cos^2 h - 1}{\cos h + 1} = -\frac{\sin^2 h}{\cos h + 1}$$

and conclude that

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

**Example.** For which $n \in \mathbb{N}$ is the function $f$ defined by

$$f(x) = \begin{cases} x^n \sin\left(\dfrac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

continuous or differentiable at $x = 0$?

**Exercise.** Show that

- $\dfrac{d}{dx}c = 0$, where $c$ is a constant.
- $\dfrac{d}{dx}\big(x^m\big) = mx^{m-1}$, where $m$ is an integer

  [without using the product rule!]
- $\dfrac{d}{dx}\sin x = \cos x$.

## 4.2 Rules for differentiation

Many differentiable functions may be constructed via function addition, subtraction, multiplication and division ...

**Theorem.** Suppose that $f$ and $g$ are differentiable at $x$. Then,

- $(f+g)'(x) = f'(x) + g'(x)$
- $(cf)'(x) = cf'(x)$, where $c$ is a constant
- $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$ (product rule)
- $(f/g)'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$ (quotient rule)

  provided that $g(x) \neq 0$.

... and function composition:

**Theorem.** Suppose that $g$ is differentiable at the point $x$ and $f$ is differentiable at the point $g(x)$. Then,

$$(f \circ g)'(x) = f'(g(x))g'(x) \quad \text{(chain rule)}.$$

**Exercise.** Use the product rule and induction on $n$ to prove that

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

for $n \in \mathbb{N}$.

**Example.** Find the derivative of

$$f(x) = \left[\sin\left(\frac{x}{x^2+1}\right)\right]^2.$$

### 4.3 Proofs of results in Section 4.2

**Proof of product rule.** Suppose that $f$ and $g$ are differentiable at the point $x$. The difference quotient of $fg$ at $x$ gives

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

We will show later that if a function is differentiable at a point then it must be continuous at that point! Hence,

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

Accordingly,

$$\frac{(fg)(x+h)-(fg)(x)}{h} \quad\to\quad g(x)f'(x)+f(x)g'(x)$$

as $h\to 0$.

Proofs of the other differentiation rules are found in most undergraduate calculus textbooks.

## 4.4 Implicit differentiation

**Idea.** On use of the chain rule, determine the derivative of a function which is implicitly defined.

**Example.** Determine the tangent line to the curve defined by

$$x^4 - x^2y^2 + y^4 = 13$$

at the point $(2, 1)$.

Write

$$x^4 - x^2y(x)^2 + y(x)^4 = 13$$

and differentiate both sides:

$$4x^3 - 2xy(x)^2 - 2x^2y(x)y'(x) + 4y(x)^3y'(x) = 0.$$

Evaluate at $(x, y) = (2, 1)$ and solve for $y'(2)$:

$$32 - 4 - 8y'(2) + 4y'(2) = 0 \quad \Rightarrow \quad y'(2) = 7.$$

Equation of tangent line at $(2, 1)$:

$$y - 1 = 7(x - 2).$$

**Theorem.** Suppose that $q$ is a rational number. Then

$$\frac{d}{dx}x^q = qx^{q-1}.$$

**Proof.** Since $q$ is a rational number, there exist integers $m$ and $n$ such that

$$q = \frac{m}{n}.$$

Write

$$y = x^q = x^{m/n}$$

and take the $n$th power of both sides, leading to

$$y^n = x^m.$$

Differentiation of both sides with respect to $x$ yields

$$ny^{n-1}\frac{dy}{dx} = mx^{m-1}.$$

Hence,

$$\begin{aligned}\frac{dy}{dx} &= \frac{mx^{m-1}}{ny^{n-1}} \\ &= \frac{m}{n}\frac{x^{m-1}}{x^{q(n-1)}} \\ &= qx^{(m-1)-qn+q} \\ &= qx^{q-1}\end{aligned}$$

as required.

## 4.5 Differentiation, continuity and split functions

**Theorem.** If $f$ is differentiable at $a$ then $f$ is continuous at $a$.

**Proof.** It is seen that

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

since $f$ is differentiable at $a$. Accordingly,

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

**Conversely:** If $f$ is not continuous at $a$ then it is not differentiable at $a$!

**Previous example.** At $x = 0$, the function $f$ defined by

$$f(x) = \begin{cases} x^n \sin\left(\dfrac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is continuous for $n \geq 1$ and differentiable for $n \geq 2$.

Extra care is required if one deals with 'split functions'. Whether or not the function is differentiable at the 'split point' can be determined by calculating left- and right-hand limits of the difference quotient. ...

**Example.** Suppose that $f : (0, \infty) \to \mathbb{R}$ is defined by

$$f(x) = \begin{cases} 4\sqrt{x} & \text{if} \quad 0 < x \leq 1 \\ bx^2 + c & \text{if} \quad x > 1, \end{cases}$$

where $b$ and $c$ are real numbers. Find all possible values of $b$ and $c$ such that $f$ is (i) continuous at $x = 1$ and (ii) differentiable at $x = 1$.

... However, in many cases, this may be avoided by using the following theorem:

**Theorem.** Suppose that $a$ is a fixed real number and that a function $f$ is the rule

$$f(x) = \begin{cases} p(x) & \text{if} \quad x \geq a \\ q(x) & \text{if} \quad x < a, \end{cases}$$

where $p$ and $q$ are defined on some open interval containing $a$. If $f$ is continuous at $a$ and $p'(a) = q'(a)$ then $f$ is differentiable at $x = a$.

**Remark.** Note that the requirement of $f$ being continuous at $a$ is equivalent to demanding that $p(a) = q(a)$ since $p$ and $q$ are continuous at $a$.

**Previous example.** Consider the function $f : (0, \infty) \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 4\sqrt{x} & \text{if} \quad 0 < x \leq 1 \\ bx^2 + c & \text{if} \quad x > 1, \end{cases}$$

where $b$ and $c$ are real numbers.

## 4.6 Derivatives and function approximation

By definition, if a function $f$ is differentiable at $a$, we have

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

so that

$$f'(a) \approx \frac{f(x) - f(a)}{x - a}$$

if $x$ is 'sufficiently close' to $a$. Thus,

$$f(x) \approx f(a) + f'(a)(x - a)$$

and the right-hand side may be regarded as an 'approximation' of $f(x)$ in a neighbourhood of $x = a$.

**Example.** Suppose that the horizontal distance $L$ covered by a tennis ball is given by

$$L = \frac{v^2}{g} \sin 2\varphi,$$

where $v$ and $\varphi$ are the initial speed and angle respectively and $g$ is the constant vertical acceleration due to gravity. How does $L$ change if either $v$ or $\varphi$ are changed by a 'small' amount?

## 4.7 Derivatives and rates of change

Many physical processes involve quantities (such as temperature, volume, concentration, velocity) that change with time but may not be independent of each other. Their rates of change may then be obtained by careful application of the chain rule.

**Example.** A spherical balloon is being inflated and its radius is increasing at a constant rate of $6$ mm/sec. At what rate is its volume increasing when the radius of the balloon is $20$ mm?

Let $V(t)$ be the volume of the balloon and $r(t)$ be its radius at time $t$. Alternatively, let $\tilde{V}(r)$ be the volume of the balloon as a function of its radius $r$ given by

$$\tilde{V} = \frac{4}{3}\pi r^3$$

so that

$$\frac{d\tilde{V}}{dr} = 4\pi r^2.$$

Then, the chain rule implies that

$$\frac{dV}{dt} = \frac{d\tilde{V}}{dr}\,\frac{dr}{dt} = 4\pi r^2 \frac{dr}{dt}.$$

We are told that

$$\frac{dr}{dt} = 6$$

so that

$$\frac{dV}{dt} = 4\pi(20)^2 \times 6 = 9600\pi$$

at $r = 20$.

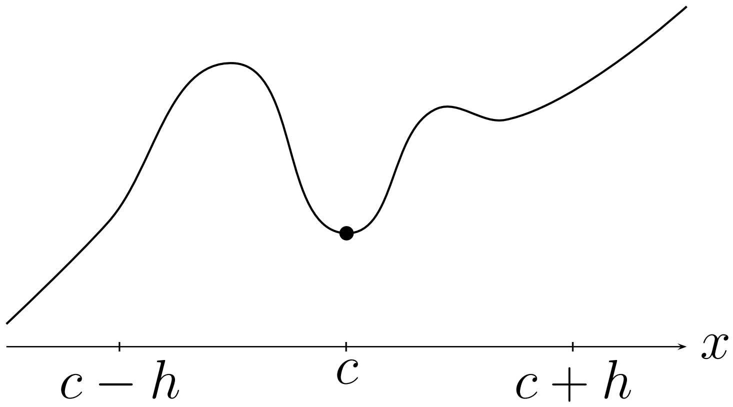
Hence, the volume is increasing at a rate of $9600\pi$ mm$^3$/sec when the radius is $20$ mm.

The above example illustrates an approach to solving such problems.

1. Define variables for the quantities involved.
2. Write down what is known in terms of these variables and their derivatives.
3. Write down what you need to find in terms of these variables and their derivatives.
4. Write down anything else you know that relates the variables (for example, a volume or area formula).
5. Use the chain rule (or implicit differentiation) to find the relevant derivative.

## 4.8 Local maximum, local minimum and stationary points

In this section, we begin to develop a systematic approach to locating maxima and minima. A complete approach will be presented in the next chapter.




Local minimum point c

**Definition.** Let $f$ be defined on some interval $I$.

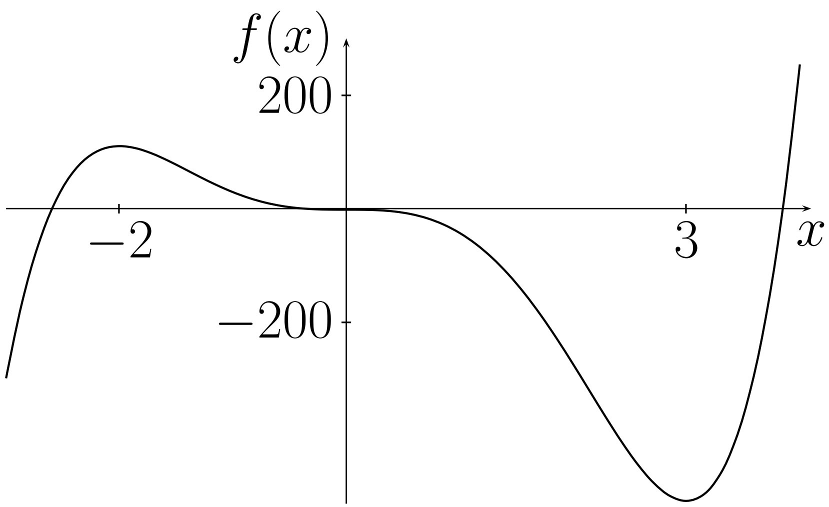
- We say that a point $c$ in $I$ is a local minimum point if there exists an $h > 0$ such that
$$f(c) \leq f(x) \quad \text{for all} \quad x \in (c-h, c+h) \cap I.$$
- We say that a point $d$ in $I$ is a local maximum point if there exists an $h > 0$ such that
$$f(x) \leq f(d) \quad \text{for all} \quad x \in (d-h, d+h) \cap I.$$

**Theorem.** Suppose that $f$ is defined on $(a, b)$ and has a local maximum or minimum point at $c$ for some $c$ in $(a, b)$. If $f$ is differentiable at $c$ then $f'(c) = 0$.

**Definition.** If a function $f$ is differentiable at a point $c$ and $f'(c) = 0$ then $c$ is called a stationary point of $f$.

**Example.** Find all the stationary, maximum and minimum points of the function $f : [-3, 4] \to \mathbb{R}$ defined by

$$f(x) = 4x^5 - 5x^4 - 40x^3 - 2.$$



Differentiation yields

$$\begin{aligned} f'(x) &= 20x^4 - 20x^3 - 120x^2 \\ &= 20x^2(x^2 - x - 6) \\ &= 20x^2(x+2)(x-3). \end{aligned}$$

Result:

- $x = -3$: local minimum point
- $x = -2$: stationary point: local maximum point
- $x = 0$: stationary point: point of inflection
- $x = 3$: stationary point: global minimum point
- $x = 4$: global maximum point

The proof of the above theorem may be found in the Calculus Notes.

# Chapter 5

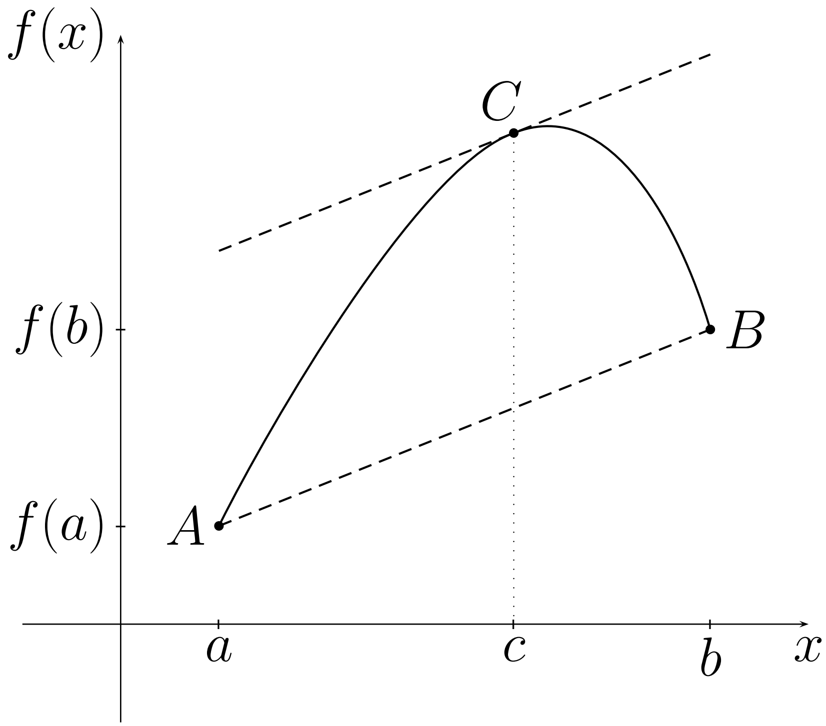
# The mean value theorem and its applications

The mean value theorem is one of the most important results for establishing the theoretical structure of calculus.

Applications of the mean value theorem include

- identifying where a function is increasing or decreasing,
- identifying different types of stationary points,
- determining how many zeros a polynomial has,
- evaluating limits which are indeterminate forms of type $\frac{\infty}{\infty}$ and $\frac{0}{0}$,
- proving useful inequalities and
- estimating errors in approximations.

## 5.1 The mean value theorem




**The mean value theorem.** Suppose that a function $f$ is continuous on $[a, b]$ and differentiable on $(a, b)$. Then, there exists at least one real number $c$ in $(a, b)$ such that

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

**Remark.** In the above theorem, it is required that $f$ is continuous on the closed interval but differentiable only on the open interval!

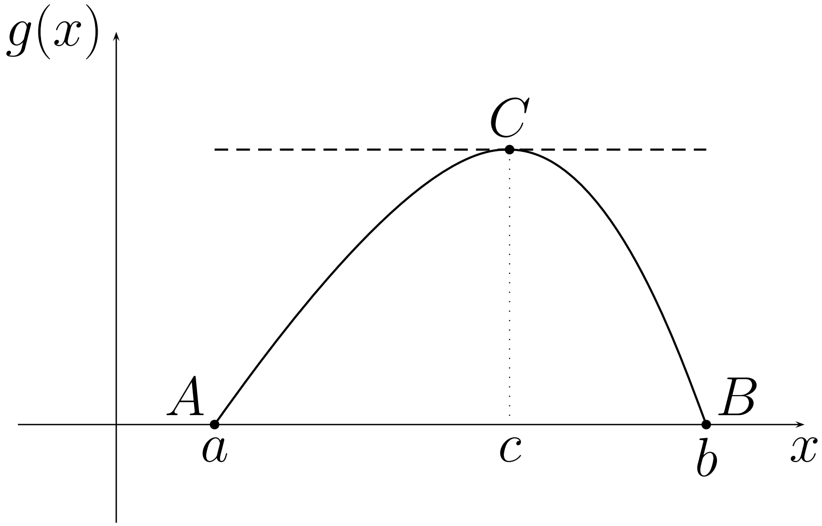
**Example.** Find counterexamples which demonstrate that the continuity and differentiability requirements must be met.

**Application.** A car enters a tunnel at a speed of 30km/h and after 1 minute leaves the tunnel at a speed of 40 km/h. The length of the tunnel is 1km. Did the driver break the speed limit of 50km/h?

**Example.** Apply the mean value theorem to the function $f$ defined by $f(x) = x^5$ and $a = -1,\ b = 4.$ Find the value(s) of $c$ which satisfy the conclusion of the theorem.

## 5.2 Proof of the mean value theorem

If $f(a) = f(b)$ then the mean value theorem reduces to Rolle's theorem.




**Rolle's theorem.** Suppose that a function $g$ is continuous on $[a, b]$ and differentiable on $(a, b)$. If, in addition,

$$g(a) = g(b) = 0$$

then there exists a $c$ in $(a, b)$ such that $g'(c) = 0$.

**Proof of Rolle's theorem.**

Case 1: Suppose that
$$g(x) = 0.$$
Then, $g'(c) = 0$ for every $c$ in $(a, b)$.

Case 2: Suppose that there exists a point $d$ in $(a, b)$ such that
$$g(d) > 0.$$

By the max-min theorem, $g$ attains a maximum value at some point $c$ in $[a, b]$.

Moreover,
$$g(c) \geq g(d) > g(a) = g(b) = 0$$
so that $c$ must lie in $(a, b)$.

Since $g$ is differentiable on $(a, b)$, we know that
$$g'(c) = 0.$$

Case 3: Suppose that

$$g(x) \le 0$$

for all $x$ in $[a, b]$ and that $g$ is not constant on $[a, b]$.

Then, $g$ attains a minimum at a point $c$ in $(a, b)$ and hence $g'(c) = 0$.

In order to prove the mean value theorem, we merely 'deform' the graph of $f$ in such a way that Rolle's theorem applies:

**Proof of the mean value theorem.** Suppose that $f$ is continuous on $[a, b]$ and differentiable on $(a, b)$.

We consider the function $g$ defined by

$$g(x) = f(x) - \left[\frac{f(b) - f(a)}{b - a}(x - a) + f(a)\right].$$

It is evident that $g$ is continuous on $[a, b]$, differentiable on $(a, b)$ and that $g(a) = g(b) = 0$.

By virtue of Rolle's theorem, there exists a $c$ in $(a, b)$ such that $g'(c) = 0$, that is, such that

$$f'(c) - \left[\frac{f(b) - f(a)}{b - a}\right] = 0.$$

Hence,

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

as required.

## 5.3 Proving inequalities using the mean value theorem

An important application of the mean value theorem is in proving inequalities.

**First example.** Show that

$$e^x > 1 + x \quad \text{for all} \quad x > 0.$$

Hint: Make the identification $[a, b] = [0, x]$ for any fixed $x > 0$.

**General 'philosophy.'** Apply the mean value theorem to an appropriate function $f$ and find a lower or upper bound for $f'(c)$ on $(a, b)$.

**Example.** It is known that any polynomial 'grows faster than' the logarithm. For, instance, show that

$$\ln x < x - 1 \quad \text{for all} \quad x > 1.$$

Suppose that $x > 1$ and consider the closed interval $[1, x]$. We define a function

$$f : [1, x] \to \mathbb{R}$$

by

$$f(t) = \ln t.$$

Now, $f$ is continuous on $[1, x]$ and differentiable on $(1, x)$ so that the mean value theorem implies that

$$\frac{f(x) - f(1)}{x - 1} = f'(c)$$

for some $c$ in $(1, x)$. Accordingly,

$$\frac{\ln x}{x - 1} = \frac{1}{c}$$

for some $c$ between $1$ and $x$.

An upper bound of $f'(c)$ on $(1, x)$ is given by

$$f'(c) = \frac{1}{c} < 1$$

since $c > 1$. We therefore conclude that

$$\frac{\ln x}{x - 1} < 1$$

or, equivalently,

$$\ln x < x - 1.$$

## 5.4 Error bounds

A second application of the mean value theorem is in calculating error bounds.

**Example.** How good an approximation of $\cos 1$ is $\cos\frac{\pi}{3}$?

Since $1 \approx \frac{\pi}{3}$, we are inclined to believe that $\cos 1 \approx \frac{1}{2}$.

We apply the mean value theorem to the function

$$f : [1, \tfrac{\pi}{3}] \to \mathbb{R}, \qquad f(x) = \cos x$$

which is continuous and differentiable. Thus,

$$\frac{\cos\frac{\pi}{3} - \cos 1}{\frac{\pi}{3} - 1} = f'(c) = -\sin c$$

for some $c$ in $(1, \frac{\pi}{3})$.

Accordingly,

$$0 < \cos 1 - \cos \tfrac{\pi}{3} = \left(\frac{\pi}{3} - 1\right) \sin c \leq \frac{\pi}{3} - 1 < 0.05$$

since $\pi < 3.15$ so that

$$0.5 < \cos 1 < 0.55.$$

Conclusion: The error of approximating $\cos 1$ by $\cos \frac{\pi}{3}$ is at most $0.05$.

## 5.5 The sign of a derivative

**Definition.** Suppose that a function $f$ is defined on an interval $I$. We say that

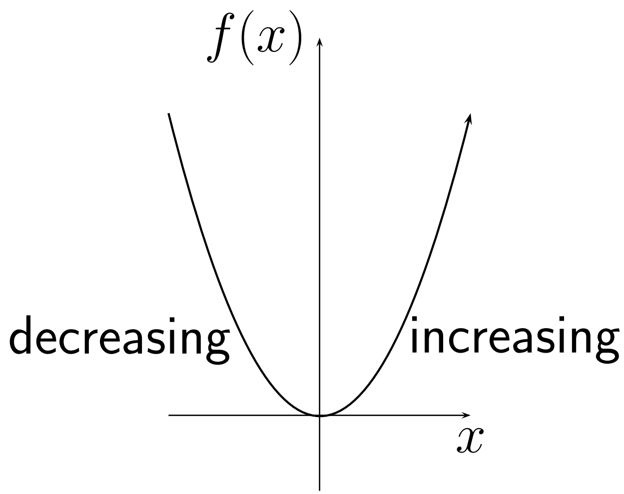
- $f$ is increasing on $I$ if for every two points $x_1$ and $x_2$ in $I$,

$$x_1 < x_2 \quad \text{implies that} \quad f(x_1) < f(x_2).$$

- $f$ is decreasing on $I$ if for every two points $x_1$ and $x_2$ in $I$,

$$x_1 < x_2 \quad \text{implies that} \quad f(x_1) > f(x_2).$$

**Example.** The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is increasing on $[0, 10)$ and decreasing on $[-5, 0]$.



**Theorem.** Suppose that $f$ is continuous on $[a, b]$ and differentiable on $(a, b)$.

- If $f'(x) > 0$ for all $x$ in $(a, b)$ then $f$ is increasing on $[a, b]$.
- If $f'(x) < 0$ for all $x$ in $(a, b)$ then $f$ is decreasing on $[a, b]$.
- If $f'(x) = 0$ for all $x$ in $(a, b)$ then $f$ is constant on $[a, b]$.

**Proof.** Suppose that $f'(x) > 0$ for all $x$ in $(a, b)$ and choose two points $x_1$ and $x_2$ in $[a, b]$ such that $x_1 < x_2$.

Since $f$ is differentiable on $I$, it is continuous on $[x_1, x_2]$ and differentiable on $(x_1, x_2)$. Hence, by the mean value theorem,

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c)$$

for some $c$ in $(x_1, x_2)$.

Accordingly, $$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) > 0.$$

The remaining two statements are proven in a similar manner.

The above theorem may be directly used to classify stationary points.

**Example.** Find and classify all stationary points of the function $f : \mathbb{R} \to \mathbb{R}$ whose derivative is given by

$$f'(x) = (x-4)(x-1)(x+5)^2.$$

Stationary points: Set $f'(x) = 0$.

Result: $x = 4$, $x = 1$ and $x = -5$.

Classification: Investigate $f'(x)$ in a ‘small’ neighbourhood of any stationary point.

| | $-5^-$ | $-5$ | $-5^+$ | $1^-$ | $1$ | $1^+$ | $4^-$ | $4$ | $4^+$ |
|---|---|---|---|---|---|---|---|---|---|
| $x-4$ | $-$ | $-$ | $-$ | $-$ | $-$ | $-$ | $-$ | $0$ | $+$ |
| $x-1$ | $-$ | $-$ | $-$ | $-$ | $0$ | $+$ | $+$ | $+$ | $+$ |
| $(x+5)^2$ | $+$ | $0$ | $+$ | $+$ | $+$ | $+$ | $+$ | $+$ | $+$ |
| $f'(x)$ | $+$ | $0$ | $+$ | $+$ | $0$ | $-$ | $-$ | $0$ | $+$ |
| Gradient | $\diagup$ | $-$ | $\diagup$ | $\diagup$ | $-$ | $\diagdown$ | $\diagdown$ | $-$ | $\diagup$ |

Result:

- $x = 4$: Local minimum point
- $x = 1$: local maximum point
- $x = -5$: horizontal point of inflexion

## 5.6 The second derivative and applications

Another (potential) method for classifying the stationary points of a function $f$ involves the second derivative of $f$, which is denoted by

$$f'' \quad \text{or} \quad \frac{d^2y}{dx^2}, \quad \text{or} \quad y''$$

if we set $y = f(x)$.

**Theorem (The second derivative test).** Suppose that a function $f$ is twice differentiable on $(a, b)$ and that $c \in (a, b)$.

- If $f'(c) = 0$ and $f''(c) > 0$ then $c$ is a local minimum point of $f$.
- If $f'(c) = 0$ and $f''(c) < 0$ then $c$ is a local maximum point of $f$.

The proof of this theorem uses the following result.

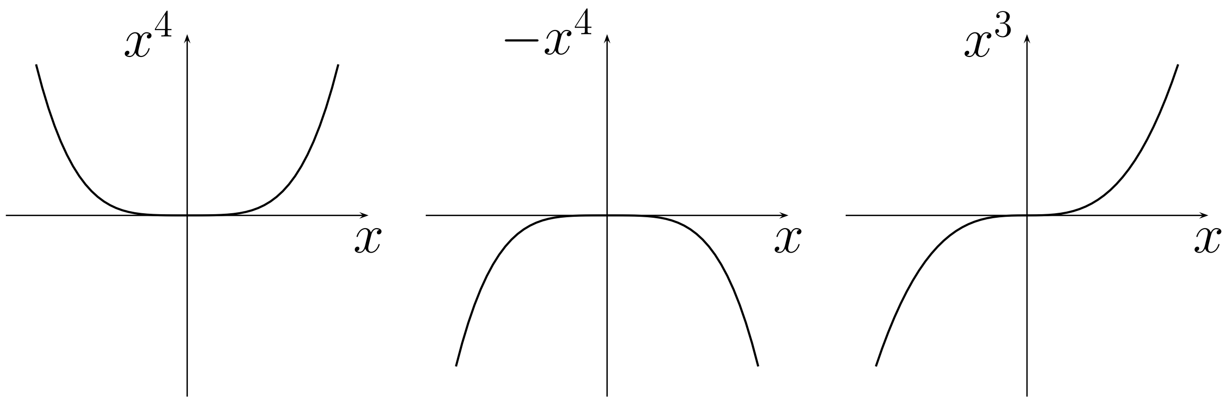
**Lemma.** Suppose that $g$ is differentiable at a point $c$.

- If $g'(c) > 0$ then $g(c-h) < g(c) < g(c+h)$ for all positive $h$ sufficiently small.
- If $g'(c) < 0$ then $g(c+h) < g(c) < g(c-h)$ for all positive $h$ sufficiently small.

**Exercise.** Find and classify the stationary points of the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = x^3 - 6x^2 + 9x - 5.$$

**Remark.** If $f'(c) = f''(c) = 0$, no conclusion may be drawn!




- If $f(x) = x^4$ then $f'(0) = f''(0) = 0$ and there is a local minimum at $0$.
- If $f(x) = -x^4$ then $f'(0) = f''(0) = 0$ and there is a local maximum at $0$.
- If $f(x) = x^3$ then $f'(0) = f''(0) = 0$ and there is a horizontal point of inflexion at $0$.

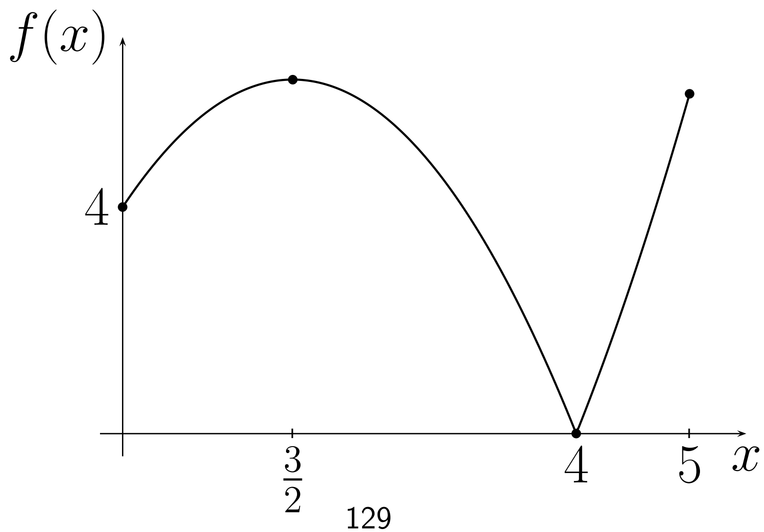
Hence if $f'(c) = f''(c) = 0$ then it is best to classify the stationary point $c$ by examining the sign of the derivative on either side of $c$!

## 5.7 Critical points, maxima and minima

**Question.** How does one find global maxima or minima?

**Definition.** Suppose that $f$ is defined on $[a, b]$. We say that a point $c$ in $[a, b]$ is a critical point for $f$ on $[a, b]$ if $c$ satisfies one of the following properties:

- $c$ is an endpoint $a$ or $b$ of the interval $[a, b]$
- $f$ is not differentiable at $c$
- $f$ is differentiable at $c$ and $f'(c) = 0$.



**Theorem.** Suppose that $f$ is continuous on $[a, b]$. Then, $f$ has a global maximum and global minimum on $[a, b]$. Moreover, the global maximum point and the global minimum point are both critical points for $f$ on $[a, b]$.

**Example.** Suppose that the function $f : \mathbb{R} \to \mathbb{R}$ is given by the rule

$$f(x) = |x^2 - 3x - 4|.$$

Find the absolute maximum and absolute minimum values of $f$ on the interval $[0, 5]$ (cf. previous graph).

## 5.8 Counting zeros

**Example.** Determine the number of real zeros of

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^4 - x^3 - 3x^2 - 8x - 5$$

and give an approximate location for each zero.

Differentiation yields

$$f'(x) = (x-2)(4x^2 + 5x + 4)$$

so that $f'(2) = 0$ and

$$f'(x) < 0 \text{ on } (-\infty, 2) \quad \text{and} \quad f'(x) > 0 \text{ on } (2, \infty).$$

Now, since $f(2) = -25$, we can roughly sketch the graph (exercise!) and conclude the following:

- $f$ is decreasing on $(-\infty, 2)$ and therefore cannot have more than one zero on this interval.

  But $f(-1) = 2$ and $f(0) = -5$ and, hence, the intermediate value theorem implies that $f$ has at least one zero on $(-1, 0)$.

- $f$ is inreasing on $(2, \infty)$ and therefore cannot have more than one zero on this interval.

  But $f(3) = -2$ and $f(4) = 107$ and, hence, the intermediate value theorem implies that $f$ has at least one zero on $(3, 4)$.

Result: $f$ has two real zeros, one in the interval $(-1, 0)$ and one in the interval $(3, 4)$.

## 5.9 Antiderivatives

**Remark.** The velocity of a particle is the time-derivative of its position. Accordingly, the position of a particle may be regarded as an antiderivative of its velocity.

**Definition.** Suppose that $f$ is continuous on an open interval $I$. A function $F$ is said to be an antiderivative (or a primitive) of $f$ on $I$ if

$$F'(x) = f(x) \quad \text{for all} \quad x \in I.$$

The process of finding an antiderivative of a function is called antidifferentiation.

**Remark.** It is evident that if $F$ is and antiderivative of $f$ then $G$ defined by

$$G(x) = F(x) + C,$$

where $C$ is an arbitrary constant, is also an antiderivative of $f$.

**Example.** Let

$$f_\alpha(x) = x^\alpha e^{x^2}, \qquad \alpha \geq 0.$$

Show that if $F_{\alpha-2}$ is an antiderivative of $f_{\alpha-2}$ for some $\alpha \geq 2$ then $F_\alpha$ defined by

$$F_\alpha(x) = \frac{x^{\alpha-1}}{2} e^{x^2} - \frac{\alpha-1}{2} F_{\alpha-2}(x)$$

is an antiderivative of $f_\alpha$.

**Theorem.** Suppose that $f$ is a continuous function on an open interval $I$ and that $F$ and $G$ are two antiderivatives of $f$ on $I$. Then, there exists a real constant $C$ such that

$$G(x) = F(x) + C$$

for all $x$ in $I$.

**Proof.** Let $H$ denote the function given by

$$H(x) = G(x) - F(x)$$

for all $x$ in $I$. Then, $H$ is differentiable on $I$ and

$$\begin{aligned} H'(x) &= G'(x) - F'(x) \\ &= f(x) - f(x) \\ &= 0 \end{aligned}$$

for all $x$ in $I$. Hence, there exists a constant $C$ such that $H(x) = C$ for all $x$ in $I$ so that

$$G(x) = F(x) + C$$

for all $x$ in $I$.

Some well-known antiderivatives are given below.

| Function | Antiderivative |
|---|---|
| $x^r$, where $r$ is rational and $r \neq -1$ | $\frac{1}{r+1}x^{r+1} + C$ |
| $\sin x$ | $-\cos x + C$ |
| $\cos x$ | $\sin x + C$ |
| $e^{ax}$ | $\frac{1}{a}e^{ax} + C$ |
| $\dfrac{f'(x)}{f(x)}$ | $\ln\lvert f(x)\rvert + C$ |

## 5.10 L'Hôpital's rule

**Question.** What is the limit of the 'indeterminate expression'

$$\lim_{x \to 0} \frac{x \sin x}{1 - \cos x}\,?$$

**Theorem (l'Hôpital's rule).** Suppose that $f$ and $g$ are both differentiable functions in a neighbourhood of some $a \in \mathbb{R}$ and that either one of the two following conditions hold:

- $f(x) \to 0 \quad$ and $\quad g(x) \to 0 \quad$ as $\quad x \to a$
- $f(x) \to \infty \quad$ and $\quad g(x) \to \infty \quad$ as $\quad x \to a$.

If

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

exists then

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

**Remark.** The theorem also holds for

- limits as $x \to \infty$ or $x \to -\infty$
- one-sided limits (as $x \to a^+$ or $x \to a^-$).

L'Hôpital's rule is proved using the mean value theorem!

**Example.** Determine the limit

$$\lim_{x \to 0} \frac{x \sin x}{1 - \cos x}.$$

**Remark.** l'Hôpital's rule may be applied iteratively.

**Example.**

$$\lim_{x\to 0}\frac{\sin x - x\cos x}{x - \sin x} = \lim_{x\to 0}\frac{x\sin x}{1-\cos x} = \ldots \text{ see above.}$$

**Remark.** It is important that the limit exists after a finite number of applications of l'Hôpital's rule!

**Exercise.** What is the limit

$$\lim_{x\to\infty}\frac{2x-\sin x}{3x+\sin x}$$

and why can l'Hôpital's rule not be applied?

# Chapter 6

# Inverse functions

**Problem.** Given a function

$$f : A \to B,$$

if we set

$$y = f(x),$$

under what circumstances is it possible to express $x$ as a function of $y$, that is, to find a function

$$g : C \to A$$

such that

$$x = g(y)\,?$$

**Remark.** The above process of expressing the 'input' of a function in terms of its 'output' is commonly referred to as inverting a function.

## 6.1 Some preliminary examples

**Examples.** Is it possible to invert the following functions?

(a)

$$f : [0, \infty) \to \mathbb{R}, \qquad y = f(x) = \frac{x}{x+1}$$

(b)

$$f : [-1, 1] \to \mathbb{R}, \qquad y = f(x) = \sqrt{1 - x^2}$$

**Question.** How do we make sure that a function is invertible and what is the domain of any inverse function?

**Remark.** It is evident that the domain and range of a function play a crucial role in answering this question! For example, formal inversion of $c = c(t)$ given by

$$c = c_m e^{-\kappa(t-t_d)}$$

leads to

$$t = t_d + \frac{1}{\kappa} \ln \frac{c_m}{c}.$$

But for which $t$ and $c$ does the latter make sense?

**Standard example.** Consider the rule

$$y = x^2.$$

Whether any function defined by this rule is invertible depends on the domain:

- $f_1 : [0, \infty) \to \mathbb{R}, \qquad y = f_1(x) = x^2$

  If we take into account that $\mathrm{Range}(f_1) = [0, \infty)$ then the inverse function is given by

  $$g_1 : [0, \infty) \to [0, \infty), \qquad x = g_1(y) = \sqrt{y}.$$

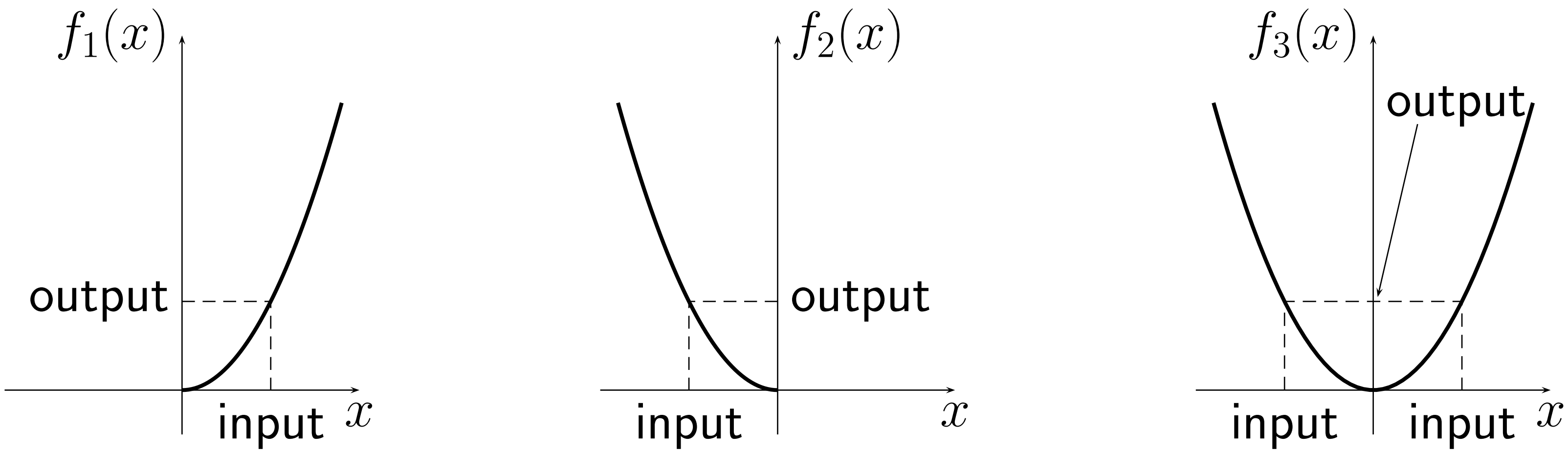
- $f_2 : (-\infty, 0] \to \mathbb{R}, \qquad y = f_2(x) = x^2$

  If we take into account that, again, $\mathrm{Range}(f_2) = [0, \infty)$ then the inverse function is given by

  $$g_2 : [0, \infty) \to (-\infty, 0], \qquad x = g_2(y) = -\sqrt{y}.$$

- $f_3 : \mathbb{R} \to \mathbb{R}, \qquad y = f_3(x) = x^2$

  The latter is not invertible since for any 'output' $y \neq 0$ there exist two 'inputs' $x = \sqrt{y}$ and $x = -\sqrt{y}$.




**Remark.** It is evident that it might be possible to construct an invertible function by restricting the domain of a given function.

**Conclusion.** The main criterion for invertibility is the existence of a one-to-one correspondence between 'inputs' and 'outputs'.

## 6.2 One-to-one functions

**Idea.** A function is one-to-one if every ‘output’ corresponds to a unique ‘input’.

**Definition.** A function $f$ is said to be one-to-one

$$\text{if} \quad f(x_1) = f(x_2) \quad \text{implies that} \quad x_1 = x_2$$

for all $x_1, x_2 \in \mathrm{Dom}(f)$.

**Terminology.** One-to-one functions are also called injective functions.

**Remark.** An injective function is equivalently characterised by

$$f(x_1) \neq f(x_2) \quad \text{for all} \quad x_1 \neq x_2$$

in the domain of $f$.

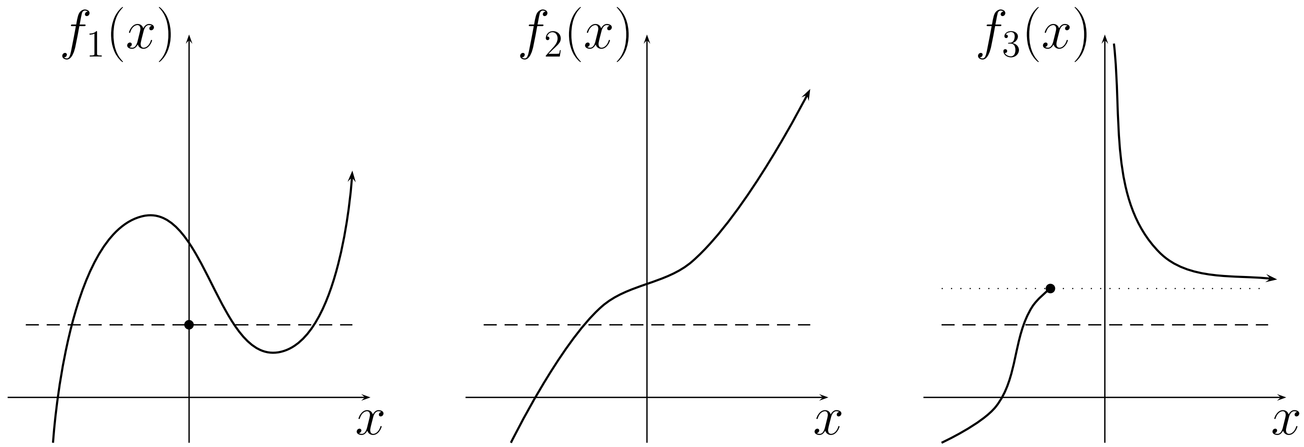
**Example.** For which $\alpha \in \mathbb{R}$ is the function

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^3 + \alpha x + 1$$

one-to-one?

**The horizontal line test.** Suppose that $f$ is a real-valued function defined on some subset of $\mathbb{R}$. Then, $f$ is one-to-one if and only if every horizontal line in the Cartesian plane intersects the graph of $f$ at most once.

**Examples.**




- $f_1$ is not one-to-one.
- $f_2$ is one-to-one.
- $f_3$ is one-to-one (even though it is neither strictly increasing nor strictly decreasing).

Although not every one-to-one function is strictly increasing (or strictly decreasing), it is true that every strictly increasing function is one-to-one.

**Theorem.** If a function $f$ is either strictly increasing or strictly decreasing then $f$ is one-to-one.

**Previous example.** Show that the function

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^3 + \alpha x + 1$$

is one-to-one for $\alpha > 0$?

Since

$$f'(x) = 3x^2 + \alpha > 0$$

for $\alpha > 0$, $f$ is strictly increasing and is therefore one-to-one for $\alpha > 0$.

Exercise: How does one argue in the case $\alpha \leq 0$?

**Example.** Is the function $f : (-1, 1) \to \mathbb{R}$ defined by

$$f(x) = 3 + x^2 + 2\tan\left(\frac{\pi}{2}x\right)$$

one-to-one?

**Remark.** Not every function whose derivative is only positive (or only negative) on its domain is one-to-one. For example,

$$\frac{d}{dx}\tan x = \sec^2 x \geq 1$$

but $\tan$ is not one-to-one on its maximal domain!

## 6.3 Inverse functions

**Theorem.** Suppose that $f$ is a one-to-one function. Then, there exists a unique function $g$ satisfying

$$g(f(x)) = x \quad \text{for all} \quad x \in \mathrm{Dom}(f)$$

and

$$f(g(y)) = y \quad \text{for all} \quad y \in \mathrm{Range}(f).$$

Moreover,

$$\mathrm{Dom}(g) = \mathrm{Range}(f), \qquad \mathrm{Range}(g) = \mathrm{Dom}(f)$$

and $g$ is one-to-one.

**Proof.** Set $D = \mathrm{Dom}(f)$ and $R = \mathrm{Range}(f)$ and define the function

$$g : R \to D$$

by choosing as $g(y)$ the unique $x \in D$ for which $y = f(x)$.

It is then left as an exercise to show that $g$ has the properties listed above.

The theorem allows us to define the term inverse function.

**Definition.** Suppose that $f$ is a one-to-one function. Then the inverse function of $f$ is the unique function $g$ given by the above theorem. The inverse function for $f$ is often denoted by $f^{-1}$.

**Remark.** If $f^{-1}$ denotes the inverse function of a one-to-one function $f$ then the relations in the above theorem may be expressed as

$$f^{-1}(f(x)) = x \quad \text{for all} \quad x \in \mathrm{Dom}(f)$$

and

$$f(f^{-1}(y)) = y \quad \text{for all} \quad y \in \mathrm{Range}(f)$$

so that $f$ may also be interpreted as the inverse of the function $f^{-1}$.

**Note.** $f^{-1}(y)$ does not mean $1/f(y)$!

**Remark.** Since $f^{-1}$ is a function just like any other function, we regard it as a function

$$x \mapsto f^{-1}(x)$$

so that we can graph $f^{-1}$ in the usual manner.

**Example.** Determine $f^{-1}$, where

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = 4 - \frac{1}{3}x^3.$$

Set

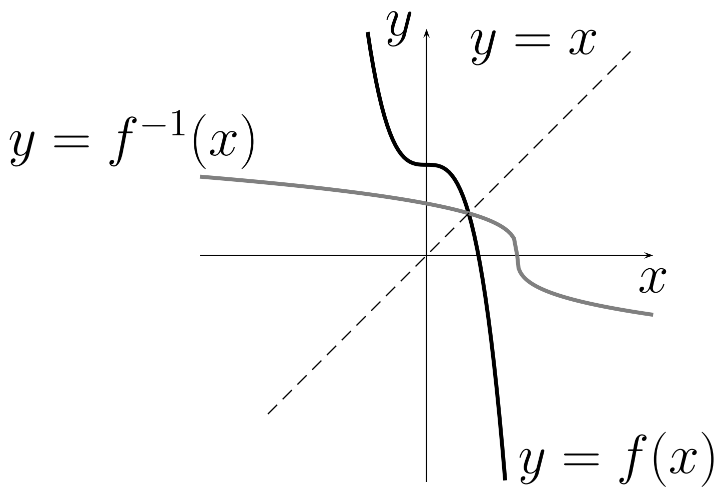
$$y = 4 - \frac{1}{3}x^3$$

so that

$$x = \sqrt[3]{12 - 3y}.$$

Hence, (interchanging $x$ and $y$),

$$f^{-1} : \mathbb{R} \to \mathbb{R}, \qquad f^{-1}(x) = \sqrt[3]{12 - 3x}.$$



### 6.4 The inverse function theorem

**Question.** If the derivative of an invertible function exists, under what circumstances is the inverse function likewise differentiable?

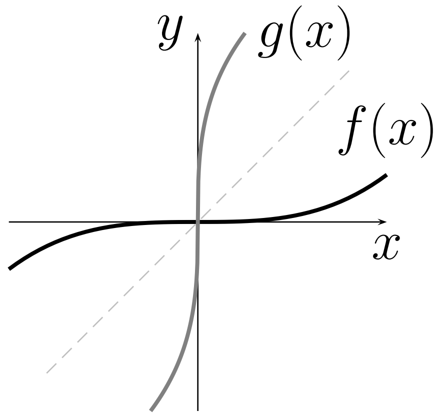
**Subtlety.** Consider the function

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^3.$$

Its inverse is given by

$$g : \mathbb{R} \to \mathbb{R}, \qquad g(x) = \sqrt[3]{x}$$

but $g$ is not differentiable at $x = 0$!




**Conclusion.** Points at which the derivative of $f$ vanishes must be excluded.

**The inverse function theorem.** Suppose that $I$ is an open interval, $f : I \to \mathbb{R}$ is differentiable and

$$f'(x) \neq 0$$

for all $x$ in $I$. Then,

- $f$ is one-to-one and has an inverse function

$$g : \mathrm{Range}(f) \to \mathrm{Dom}(f)$$

- $g$ is differentiable at all points in $\mathrm{Range}(f)$
- The derivative of $g$ is given by

$$g'(y) = \frac{1}{f'(g(y))}$$

  for all $y \in \mathrm{Range}(f)$.

**Proof.**

- Since $f'(x) \neq 0$ on $I$, $f$ is one-to-one (mean value theorem!).
- $g$ is differentiable ... too hard!
- Differentiation of
$$f(g(y)) = y$$
with respect to $y$ yields
$$f'(g(y)) \times g'(y) = 1.$$

**Remark.** Once again, we usually write the derivative of the inverse function $g$ as

$$g'(x) = \frac{1}{f'(g(x))}$$

for $x \in \mathrm{Range}(f)$.

**Example.** Consider the function

$$f : (0, \infty) \to (0, \infty), \qquad f(x) = x^3$$

and its inverse

$$g : (0, \infty) \to (0, \infty), \qquad g(x) = \sqrt[3]{x}.$$

Then,

$$g'(x) = \frac{1}{f'(g(x))} = \frac{1}{3[g(x)]^2} = \frac{1}{3x^{2/3}}$$

as expected.

**Previous example.** Determine the derivative of the inverse of the function $f : (-1, 1) \to \mathbb{R}$ defined by

$$f(x) = 3 + x^2 + 2\tan\left(\frac{\pi}{2}x\right)$$

at the point $f(x) = 3$.

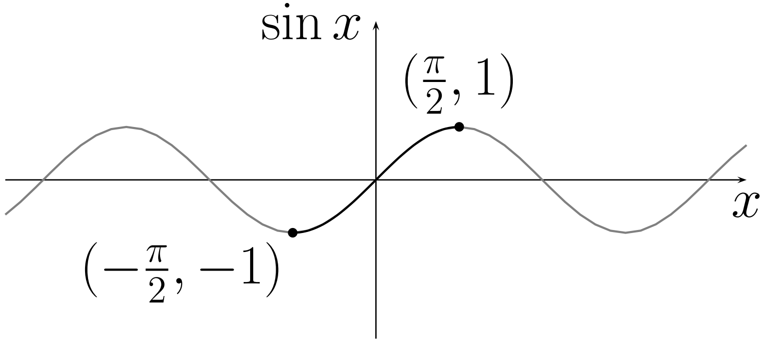
## 6.5 Applications to the trigonometric functions

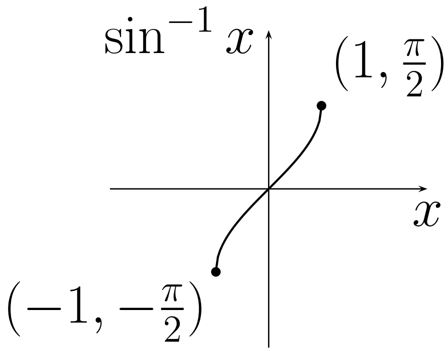
**The inverse sine function.** We consider the restricted sine function

$$\sin : [-\tfrac{\pi}{2}, \tfrac{\pi}{2}] \to [-1, 1].$$

This function is one-to-one and therefore has an inverse

$$\sin^{-1} : [-1, 1] \to [-\tfrac{\pi}{2}, \tfrac{\pi}{2}].$$



On $(-1, 1)$, according to the inverse function theorem, the derivative of $\sin^{-1}$ is given by

$$\frac{d}{dx}(\sin^{-1} x) = \frac{1}{\cos(\sin^{-1} x)}.$$

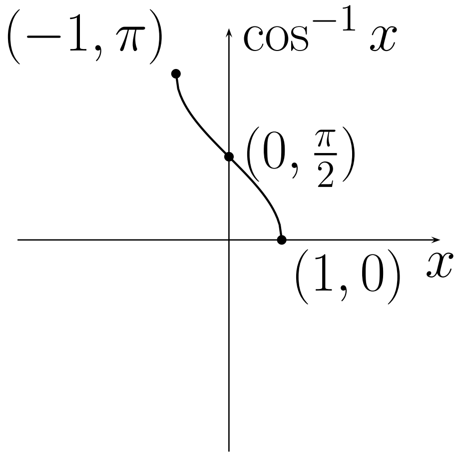
Since $\cos$ is positive on $(-\frac{\pi}{2}, \frac{\pi}{2})$, we conclude that

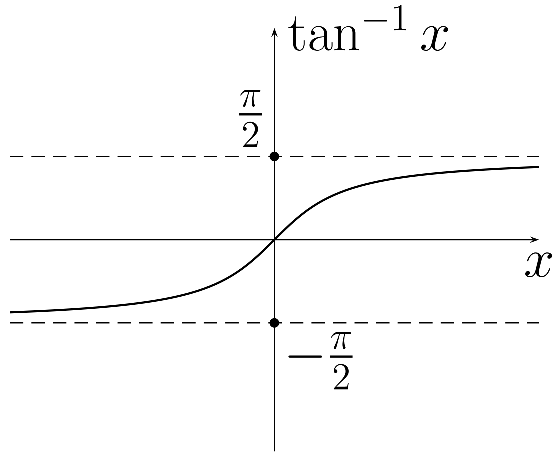
$$\frac{d}{dx}(\sin^{-1} x) = \frac{1}{\sqrt{1 - \sin^2(\sin^{-1} x)}} = \frac{1}{\sqrt{1 - x^2}}.$$

**Note.** $\dfrac{d}{dx}(\sin^{-1} x) \to \infty$ as $x \to \pm 1$.

## Table of inverse trigonometric functions.

| Function | Domain | Range | Derivative |
|---|---|---|---|
| $\sin$ | $[-\frac{\pi}{2}, \frac{\pi}{2}]$ | $[-1, 1]$ | $\frac{d}{dx}(\sin x) = \cos x$ |
| $\sin^{-1}$ | $[-1, 1]$ | $[-\frac{\pi}{2}, \frac{\pi}{2}]$ | $\frac{d}{dx}(\sin^{-1} x) = \frac{1}{\sqrt{1-x^2}}$ |
| $\cos$ | $[0, \pi]$ | $[-1, 1]$ | $\frac{d}{dx}(\cos x) = -\sin x$ |
| $\cos^{-1}$ | $[-1, 1]$ | $[0, \pi]$ | $\frac{d}{dx}(\cos^{-1} x) = -\frac{1}{\sqrt{1-x^2}}$ |
| $\tan$ | $(-\frac{\pi}{2}, \frac{\pi}{2})$ | $(-\infty, \infty)$ | $\frac{d}{dx}(\tan x) = \sec^2 x$ |
| $\tan^{-1}$ | $(-\infty, \infty)$ | $(-\frac{\pi}{2}, \frac{\pi}{2})$ | $\frac{d}{dx}(\tan^{-1} x) = \frac{1}{1+x^2}$ |




**Remark.** Even though

$$\sin(\sin^{-1}x) = x$$

for $x \in [-1,1]$, in general,

$$\sin^{-1}(\sin x) \neq x$$

unless $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

**Example.** Determine

(a)

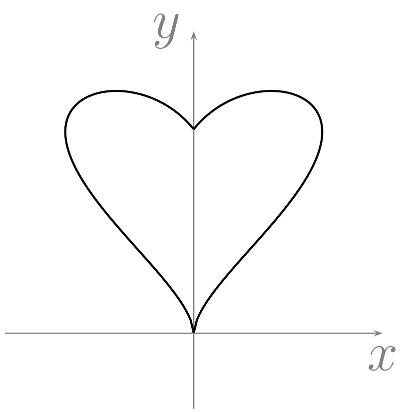
$$\cos\left(2\sin^{-1}\frac{3}{5}\right).$$

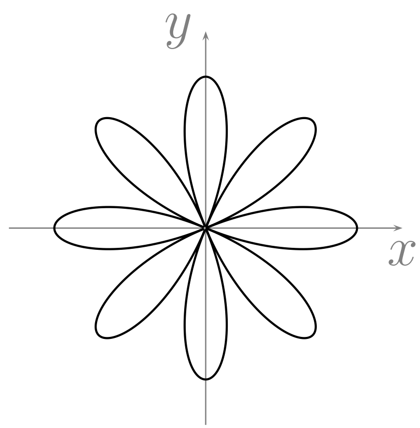
(b)

$$\sin^{-1}\left(\sin\frac{5\pi}{6}\right).$$

# Chapter 7

# Curve sketching



In this chapter, we study curves in a plane which are given in terms of

- a Cartesian equation such as

$$y = x^2 - 4 \quad \text{or} \quad y = \frac{1}{\sqrt{1-x^2}}.$$

- a parameter such as

$$x(t) = \sin t \cos t \ln |t|, \quad y(t) = \sqrt{|t|}\, \cos t$$

$$t \in [-1, 1], \quad t \neq 0.$$

- polar coordinates such as

$$x = r\cos\theta, \quad y = r\sin\theta, \quad \text{where} \quad r = \cos 4\theta.$$

## 7.1 Curves defined by a Cartesian equation

In this section, we survey techniques for sketching curves which are described by a Cartesian equation of the form $y = f(x)$.

**Checklist.** Consider the following items:

- Identify the domain of $f$.
- Identify any symmetries:
  - $f$ is even if
$$f(-x) = f(x) \quad \text{for all} \quad \pm x \in \mathrm{Dom}(f)$$
  - $f$ is odd if
$$f(-x) = -f(x) \quad \text{for all} \quad \pm x \in \mathrm{Dom}(f)$$
  - $f$ is periodic of period $T$ if
$$f(x+T) = f(x) \quad \text{for all} \quad x, x+T \in \mathrm{Dom}(f)$$

- Find $x$- and $y$-axis intercepts.
- Identify vertical asymptotes.
- Examine the behaviour of $f(x)$ as $x \to \pm\infty$ and identify existing (oblique) asymptotes.
- If necessary, identify stationary points and other features using calculus.

In the following, the above checklist is illustrated by various examples.

**Symmetries.** Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

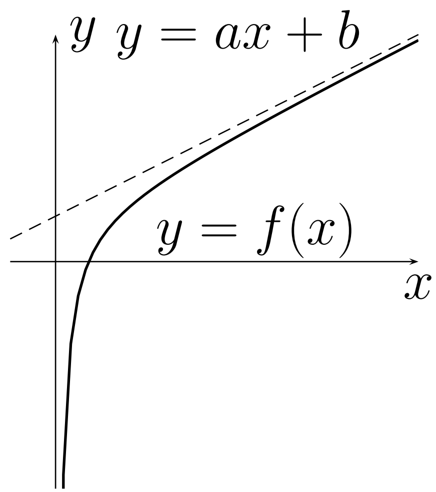
$$f(x) = \frac{|\sin x|}{2 + \cos(2x)}.$$

- $f$ is even since $f(-x) = f(x)$ for all $x \in \mathbb{R}$.
- $f$ is of period $\pi$ since $f(x + \pi) = f(x)$ for all $x \in \mathbb{R}$.

**Intercepts.** Sketch the graph of the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \dfrac{\sin x}{x} & \text{for } x \neq 0 \\ 1 & \text{for } x = 0. \end{cases}$$

**Oblique asymptotes.** The graph of a function $f$ may be asymptotic to a line $y = ax + b$ as $x \to \infty$ or $x \to -\infty$.




**Definition.** Suppose that $a \neq 0$ and $b$ are real numbers. We say that a straight line given by the equation

$$y = ax + b$$

is an oblique asymptote for a function $f$ if

$$\lim_{x \to \infty} \big(f(x) - (ax + b)\big) = 0$$

or

$$\lim_{x \to -\infty} \big(f(x) - (ax + b)\big) = 0.$$

**Remark.** If $f$ is a rational function with

$$f(x) = \frac{p(x)}{q(x)}, \quad \deg(p) = \deg(q) + 1$$

then the oblique asymptotes of $f$ may be determined by polynomial division.

**Example.** Find the oblique asymptotes to the function $f$ defined by

$$f(x) = \frac{(x-2)|x| + x}{x-2}$$

for all $x \neq 2$.

If $x > 0$ then

$$\begin{aligned} f(x) &= \frac{(x-2)x + x}{x-2} \\ &= x + \frac{(x-2) + 2}{x-2} \\ &= x + 1 + \frac{2}{x-2}. \end{aligned}$$

Thus, the line $y = x + 1$ is an oblique asymptote.

If $x < 0$ then

$$\begin{aligned} f(x) &= \frac{-(x-2)x + x}{x-2} \\ &= -x + \frac{(x-2)+2}{x-2} \\ &= -x + 1 + \frac{2}{x-2}. \end{aligned}$$

Thus, the line $y = -x + 1$ is an oblique asymptote.

**Example.** Sketch the function $f$ defined by

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

on the maximal domain.

## 7.2 Parametrically defined curves

Parametrically defined curves in a plane

$$\big(x(t), y(t)\big), \quad t \in A,$$

where $t$ is the parameter and $A$ is a given domain arise in various contexts such as

- the planar motion of a point particle (or body) under the influence of forces.

$$\text{Newton's 2nd law:} \quad m\big(\ddot{x}(t), \ddot{y}(t)\big) = F\big(x(t), y(t)\big)$$

- a suitable parametrisation of a curve defined implicitly.

$$\text{Ellipse:} \quad \big(x(t), y(t)\big) = \big(a\cos(t), b\sin(t)\big)$$

In principle, the analysis of parametrically defined curves is analogues to that outlined in the previous section.

**Example.** Sketch the curve

$$\gamma(t) = \big(x(t), y(t)\big) = (t^2 - 1, t^3 - 1).$$

<u>Possible values of $x(t)$ and $y(t)$:</u>

It is evident that $x(t)$ and $y(t)$ assume any value in the 'intervals'

$$x(t) \in [-1, \infty), \quad y(t) \in (-\infty, \infty)$$

respectively.

<u>Intercepts:</u>

$x(t) = 0$ if and only if $t = \pm 1$ so that we obtain the points

$$\gamma(-1) = (0, -2), \quad \gamma(1) = (0, 0).$$

In addition, $y(t) = 0$ if and only if $t = 1$.

Vector derivatives:

Consider the 'tangent vector'

$$\gamma'(t) = \big(x'(t), y'(t)\big).$$

If one interprets $\gamma(t)$ as the position of a particle at the time $t$ then $\gamma'(t)$ is the velocity of the particle at that time.

We will justify later that the slope of a parametrised curve $\gamma(t) = \big(x(t), y(t)\big)$ is given by

$$\frac{dy}{dx} = \frac{y'(t)}{x'(t)}$$

at all points with $x'(t) \neq 0$.

Here,

$$\gamma'(t) = (2t, 3t^2), \quad \frac{dy}{dx} = \frac{y'(t)}{x'(t)} = \frac{3}{2}t$$

so that

$$\gamma'(t) \text{ points } \nwarrow \qquad \text{for} \quad t < 0$$

and

$$\gamma'(t) \text{ points } \nearrow \qquad \text{for} \quad t > 0.$$

Note that $\gamma'(0) = 0$ so that the 'particle stops' at $t = 0$!

In fact, there exists a cusp at $\gamma(0) = (-1, -1)$ since the normalised tangent vector has the property

$$\lim_{t \to 0^\pm} \frac{\gamma'(t)}{|\gamma'(t)|} = \lim_{t \to 0^\pm} \frac{(2t, 3t^2)}{\sqrt{4t^2 + 9t^4}} = (\pm 1, 0).$$

Conclusion: The curve does not have a 'proper' tangent vector at the point $(-1, -1)$!

Limiting behaviour:

As $t \to \pm\infty$,
$$\lim_{t\to\pm\infty} \frac{\gamma'(t)}{|\gamma'(t)|} = \lim_{t\to\pm\infty} \frac{(2t, 3t^2)}{\sqrt{4t^2+9t^4}} = (0,1).$$

Accordingly, the normalised tangent vector becomes 'vertical at infinity'.

**Justification of slope formula:** By definition,

$$
\begin{aligned}
\frac{dy}{dx} &= \lim_{h\to 0} \frac{y(t+h) - y(t)}{x(t+h) - x(t)} \\
&= \lim_{h\to 0} \left[ \left( \frac{y(t+h) - y(t)}{h} \right) \Big/ \left( \frac{x(t+h) - x(t)}{h} \right) \right] \\
&= \left( \lim_{h\to 0} \frac{y(t+h) - y(t)}{h} \right) \Big/ \left( \lim_{h\to 0} \frac{x(t+h) - x(t)}{h} \right) \\
&= \frac{y'(t)}{x'(t)}
\end{aligned}
$$

provided that $x(t)$ and $y(t)$ are differentiable and $x'(t) \neq 0$.

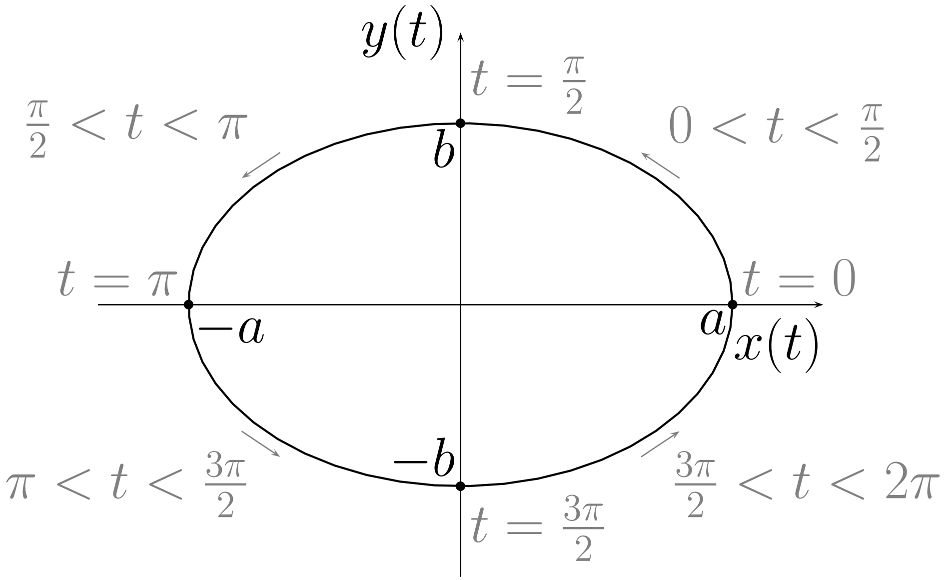
### 7.2.1 Parametrisation of conic sections

The ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

with semi-axes $a$ and $b$ admits the parametrisation

$$x(t) = a\cos t, \quad y(t) = b\sin t, \qquad 0 \le t < 2\pi.$$




Each point $(x, y)$ of the ellipse corresponds to a unique $t \in [0, 2\pi)$.

The table below lists some commonly used parametrisations of conic sections.

| Conic section | Cartesian equation | Parametric equation |
|---|---|---|
| Parabola | $4ay = x^2$ | $x(t) = 2at$<br>$y(t) = at^2$ |
| Circle | $x^2 + y^2 = a^2$ | $x(t) = a\cos t$<br>$y(t) = a\sin t$ |
| Ellipse | $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ | $x(t) = a\cos t$<br>$y(t) = b\sin t$ |
| Hyperbola | $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ | $x(t) = a\sec t$<br>$y(t) = b\tan t$ |

Note that there exist other useful parametrisations for each of these curves.

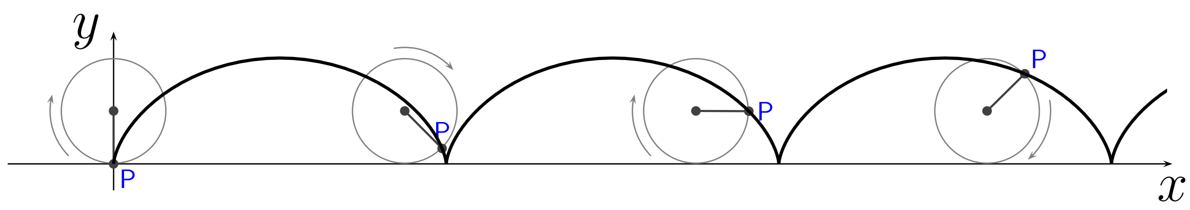
### 7.2.2 The cycloid and curve of fastest descent

**Question.** Find the shape of a curve that a particle should follow if it is to 'slide' without friction in the minimum time from a higher point $A$ to a lower point $B$ (not directly beneath it) under the influence of gravity.


Curve of fastest descent.

Such a curve is known as a curve of fastest descent or a brachistochrone (which, in Greek, means 'shortest time').

**Answer.** The curve of fastest descent from $A$ to $B$ is the unique arc of an (inverted) cycloid whose tangent at $A$ is vertical.




The cycloid.

**Exercise.** A circle of radius $r$ rolls along the $x$-axis, starting from the origin as shown above. Show that the locus $(x(t), y(t))$ of the point $P$ on the edge of the circle which satisfies $(x(0), y(0)) = (0, 0)$ is given by

$$x(t) = r(t - \sin t)$$
$$y(t) = r(1 - \cos t),$$

where $t \geq 0$.

**Remark.** If the particle moves on an inclined plane ('surfer') then the trajectory is still given by an arc of a cycloid!

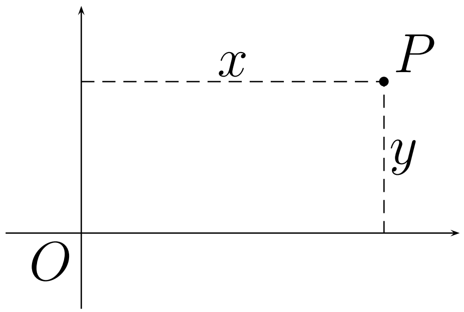
## 7.3 Curves defined by polar coordinates

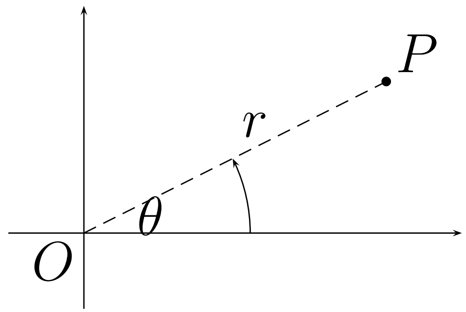
Many problems in mathematics are easier to solve if one chooses a suitable coordinate system. Here, we focus on polar coordinates.

### 7.3.1 Polar coordinates

Every point $P$ in a plane can be specified by, for example,

- $(x, y)$, where $x$ is the horizontal distance of $P$ from the origin, and $y$ is the vertical distance, or
- $(r, \theta)$, where $r$ is the distance of $P$ from the origin and $\theta$ is the angle (taken in the anticlockwise direction) between $OP$ and the positive horizontal axis.



The pair $(x, y)$ is called Cartesian coordinates of $P$.

The pair $(r, \theta)$ is called polar coordinates of $P$.

**Note.** If $P$ is the origin then $r = 0$ and $\theta$ is not defined.

Polar coordinates $(r, \theta)$ and Cartesian coordinates $(x, y)$ of a point $P$ are related by

$$x = r\cos\theta, \quad y = r\sin\theta$$

and

$$r = \sqrt{x^2 + y^2}, \quad \tan\theta = \frac{y}{x}$$

provided that $x \neq 0$.

**Note.** Finding Cartesian coordinates of a point $P$ given in terms of polar coordinates is easy but care must be taken in the opposite case.

**Example.** Find all polar coordinates of the point $P$ with Cartesian coordinates

$$P = (x, y) = (-3, \sqrt{3}).$$

**Remark.** Polar forms of equations may be simpler or more involved compared to their Cartesian counterparts.

**Examples.** Find the polar forms of the

(a) straight line

$$y = \sqrt{3}\,x.$$

(b) circle

$$x^2 + y^2 = 4.$$

### 7.3.2 Basic sketches of polar curves

Many curves can be described by equations of the form

$$r = f(\theta) \quad \text{or} \quad \theta = g(r)$$

so that we obtain the parametrically defined curves

$$\gamma(\theta) = (r\cos\theta, r\sin\theta) = \big(f(\theta)\cos\theta, f(\theta)\sin\theta\big)$$

or

$$\gamma(r) = (r\cos\theta, r\sin\theta) = \big(r\cos g(r), r\sin g(r)\big),$$

where $\theta$ or $r$ plays the role of the parameter.

**Previous examples.**

(a) $\theta = \frac{\pi}{3}$ corresponds to $y = \sqrt{3}\,x$ for $x \geq 0$.

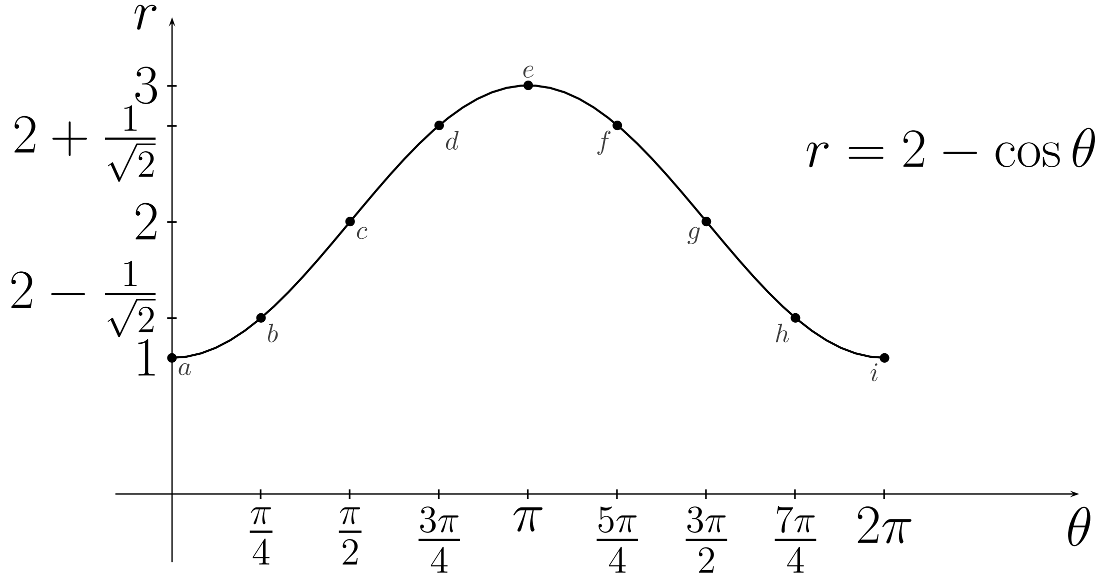
(b) $r = 2$ corresponds to $x^2 + y^2 = 4$.

**Remark.** In order to sketch a curve represented by an equation in polar form, it may be helpful to begin with an $r$ vs $\theta$ sketch.

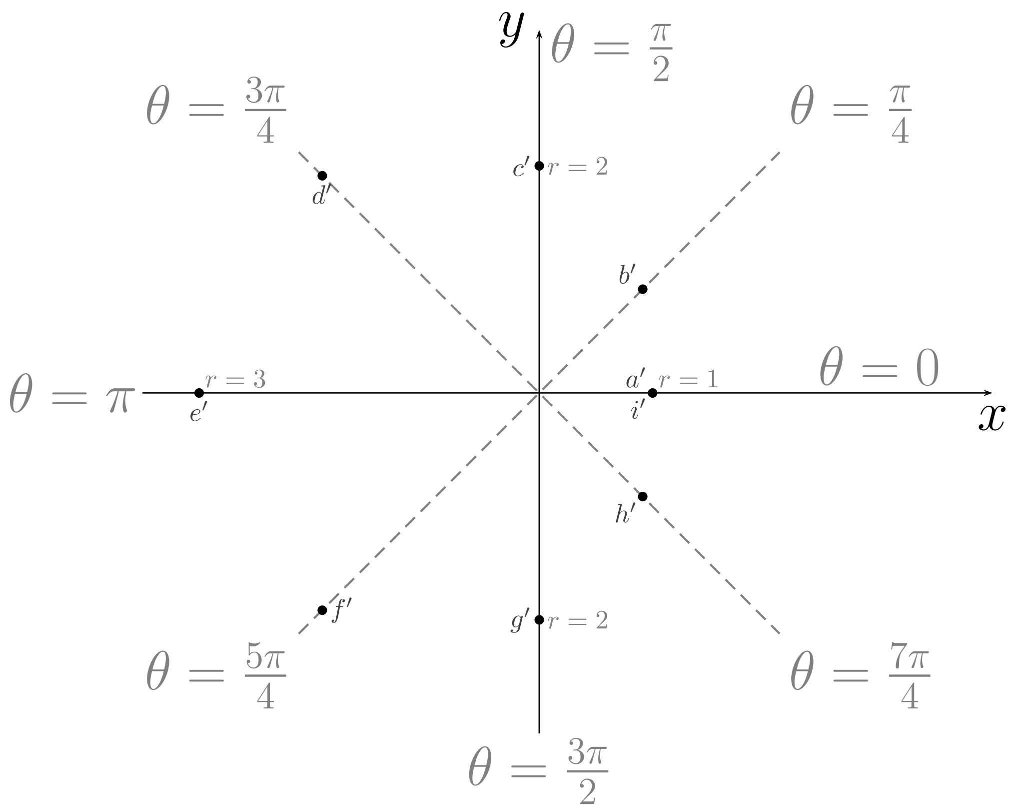
**Example.** Sketch the polar curve defined by

$$r = 2 - \cos\theta, \qquad 0 \leq \theta \leq 2\pi.$$

We first graph $r$ against $\theta$ ...



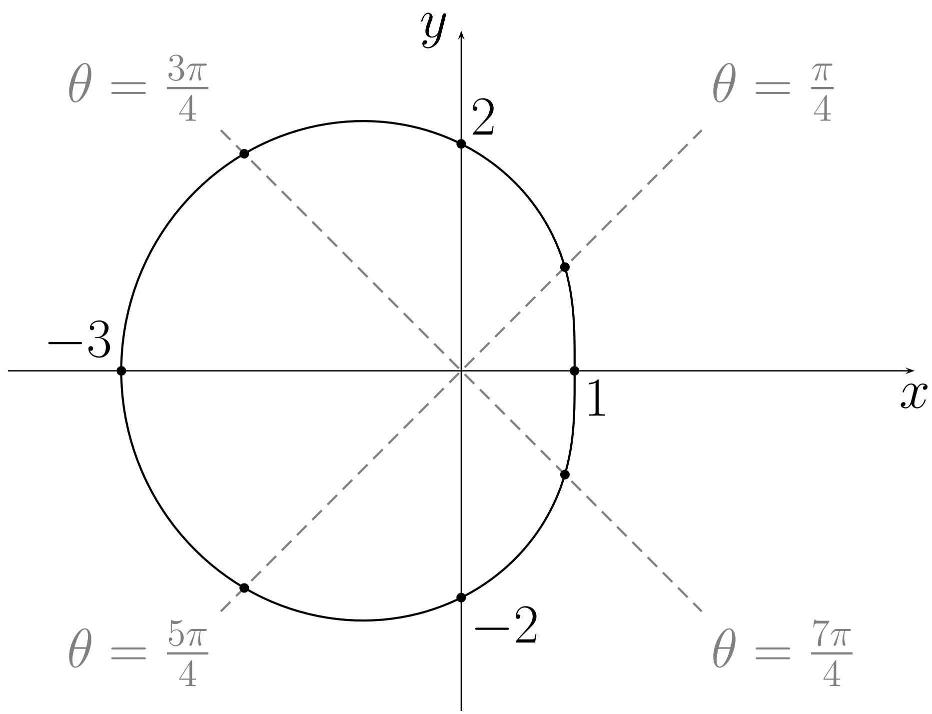
... and then mark the corresponding points on the $(x, y)$-plane:



Now:

- As $\theta$ increases from $0$ to $\pi$, $r$ increases from $1$ to $3$. Hence, the distance $r$ from the origin to points on the curve increases from $1$ to $3$.
- As $\theta$ increases from $\pi$ to $2\pi$, $r$ decreases from $3$ to $1$. Hence, the distance $r$ from the origin to points on the curve decreases from $3$ to $1$.

These considerations lead to the final sketch.




$r = 2 - \cos\theta$

**Symmetries.**

- If $f(-\theta) = f(\theta)$ then the polar curve is symmetric about the $x$-axis.
- If $f(\pi - \theta) = f(\theta)$ then the polar curve is symmetric about the $y$-axis.
- If $f$ is $2\pi$–periodic then it suffices to consider $\theta$ in the range $0 \leq \theta < 2\pi$.

**Exercise.** Sketch the curve described by the polar equation

$$r = 2\sin\theta, \qquad 0 \leq \theta \leq \pi$$

and show that it constitutes a circle.

### 7.3.3 Sketching polar curves using calculus

Suppose that a curve can be expressed in polar form as

$$r = f(\theta).$$

Since the curve's parametric form is given by

$$\gamma(\theta) = \big(x(\theta), y(\theta)\big) = (r\cos\theta, r\sin\theta),$$

the tangent vector reads

$$\begin{aligned}\gamma'(\theta) &= \big(x'(\theta), y'(\theta)\big) \\ &= \left(\frac{dr}{d\theta}\cos\theta - r\sin\theta, \frac{dr}{d\theta}\sin\theta + r\cos\theta\right).\end{aligned}$$

Thus, horizontal tangents are obtained by solving

$$y'(\theta) = 0 \quad \text{but} \quad x'(\theta) \neq 0,$$

while vertical tangents correspond to

$$x'(\theta) = 0 \quad \text{but} \quad y'(\theta) \neq 0.$$

**Example.** Sketch the curve described by the polar equation

$$r = 1 + \cos\theta.$$

**Concluding remark.** In connection with representing and sketching curves in terms of polar coordinates, it is often convenient to allow negative values of $r$. Thus, if $r$ is negative, we make the identification

$$(r, \theta) \quad \leftrightarrow \quad (-r, \theta + \pi)$$

and hence $(r, \theta)$ is obtained from $(-r, \theta)$ by reflection in the origin. This is consistent with the formulae

$$x = r\cos\theta = (-r)\cos(\theta + \pi)$$
$$y = r\sin\theta = (-r)\sin(\theta + \pi).$$

**Example.** Sketch the curve described by the polar equation

$$r = \frac{1}{2} + \cos\theta.$$

# Chapter 8

# Integration

**Problem.** How does one find/measure/define areas (of regions with curved boundaries)?

The main ideas can be traced back to

- Archimedes (the great Greek mathematician of antiquity!) and then
- Isaac Barrow (Isaac Newton's mentor)
- Isaac Newton (one of the greatest mathematicians and physicists of all time)
- Gottfried Leibniz (Newton's contemporary)
- Bernhard Riemann (a German mathematician of the nineteenth century)
- Henri Lebesgue (a French mathematician of the early twentieth century).

One (but not the only) method for calculating the area of regions with curved boundaries is known as Riemann integration.

**Note.** A priori, calculating areas (integration) and antidifferentiation are two separate problems but the remarkable fundamental theorem of calculus shows that these are essentially the same!

This theorem was actually known to Barrow but its implications were developed by Newton, Leibniz and their disciples.

**Remark.** The problem of calculating area is much the same as that of calculating mass, volume, work and probability. The unifying feature is the so-called integral calculus.

Here, we confine ourselves to the determination of 'the area under the graph of a function' via the Riemann integral. This may then be generalised to the area of sets the boundary of which cannot be represented in this simple manner.

## 8.1 Area and the Riemann integral

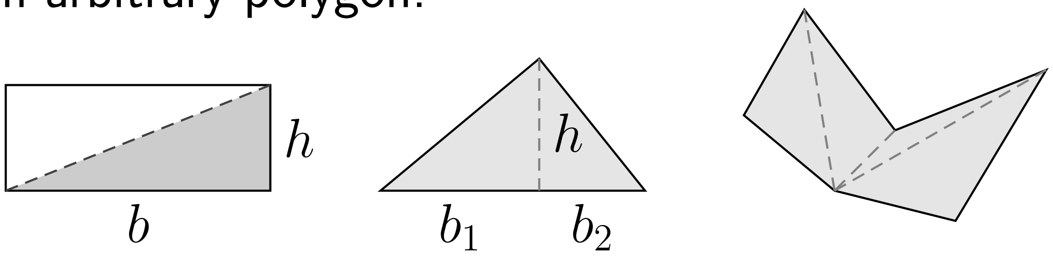
We all (think that we) know how to calculate the area of a rectangle, a triangle or even general polygons by partitioning the polygon into triangles.

But what are the rules that we take for granted?

- The area of a rectangle is the product of its length and height.
- Areas of congruent regions are equal.
- The area of a whole region is the sum of the areas of its 'parts'.

Thus, from the area of a rectangle ($bh$), we can 'derive' the formula for

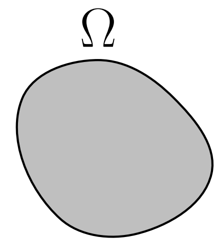
- the area of a right-angled triangle ($\frac{1}{2}bh$), then
- the area of an arbitrary triangle ($\frac{1}{2}(b_1 + b_2)h$) and then
- the area of an arbitrary polygon.



Formally, we demand that any definition of an area satisfy the following axioms:

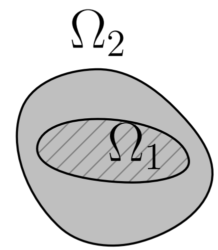
(A1) If $\Omega$ is a region of the plane then

$$\text{area}(\Omega) \geq 0.$$



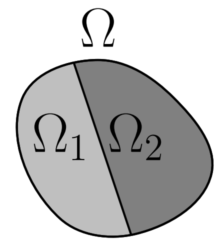

(A2) If one region $\Omega_1$ is contained in another region $\Omega_2$, then

$$\text{area}(\Omega_1) \leq \text{area}(\Omega_2).$$



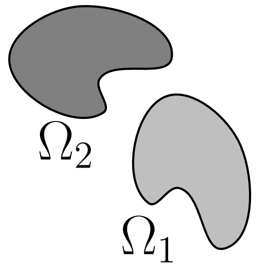

(A3) If the area of a region $\Omega$ is partitioned into two smaller disjoint regions $\Omega_1$ and $\Omega_2$, then

$$\text{area}(\Omega) = \text{area}(\Omega_1) + \text{area}(\Omega_2).$$



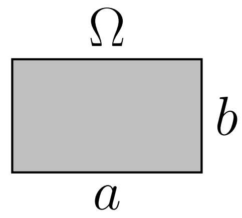
(A4) If $\Omega_1$ and $\Omega_2$ are congruent regions then

$$\text{area}(\Omega_1) = \text{area}(\Omega_2).$$



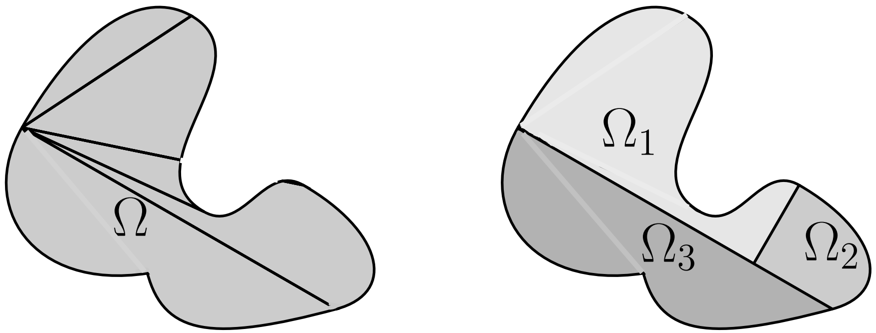

(A5) If $\Omega$ is a rectangle of length $a$ and height $b$ then

$$\text{area}(\Omega) = ab.$$



### 8.1.1 Area of regions with curved boundaries

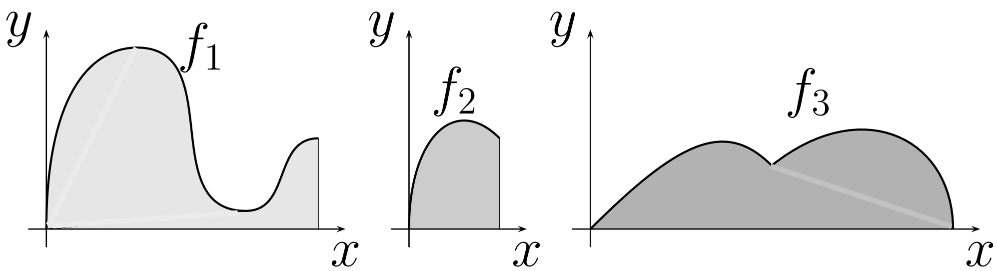
**Question.** How could one calculate the area of the region below?




**Answer.** Apply axiom (A3) and conclude that

$$\mathrm{area}(\Omega) = \mathrm{area}(\Omega_1) + \mathrm{area}(\Omega_2) + \mathrm{area}(\Omega_3).$$

Then, rotate and translate each subregion, so that application of axiom (A4) implies that each of $\mathrm{area}(\Omega_1)$, $\mathrm{area}(\Omega_2)$ and $\mathrm{area}(\Omega_3)$ is equal to the area under the graph of a function.



This procedure can be done for any region in the plane with a ‘reasonable’ boundary.

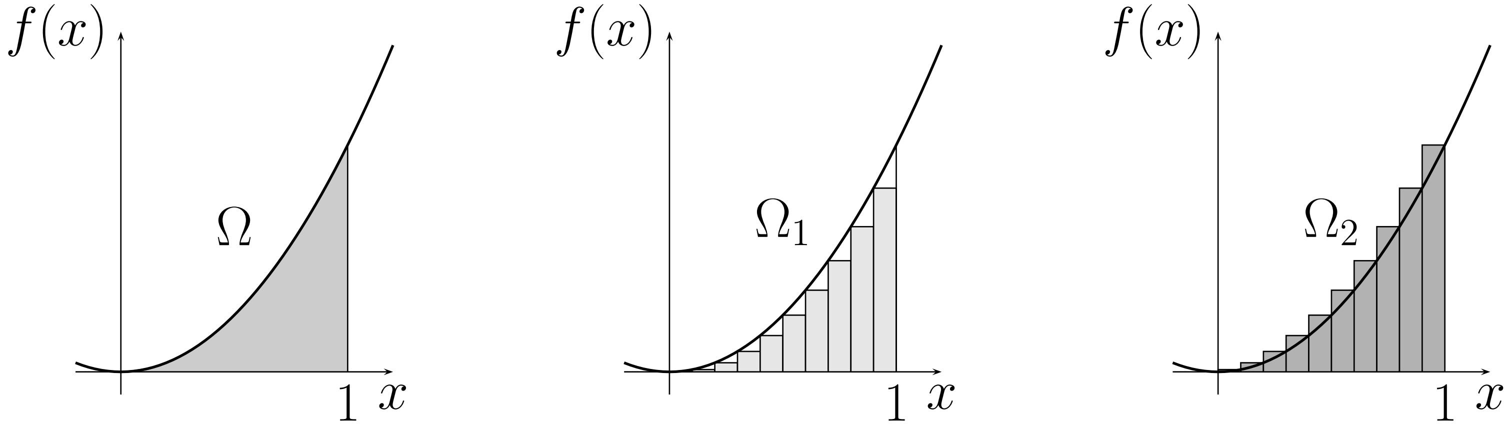
Hence, we are left with defining what is meant by ‘the area under the graph of a function’.

### 8.1.2 Approximations of area using Riemann sums

**Example.** Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by the rule

$$f(x) = x^2.$$

Let $\Omega$ denote the region bounded by the graph of $f$, the $x$-axis and the lines $x = 0$ and $x = 1$.




**Idea.** Find lower and upper bounds for $\mathrm{area}(\Omega)$ by choosing appropriate 'approximations' $\Omega_1$ and $\Omega_2$ of the region $\Omega$ in terms of $n$ rectangles.

It is evident that

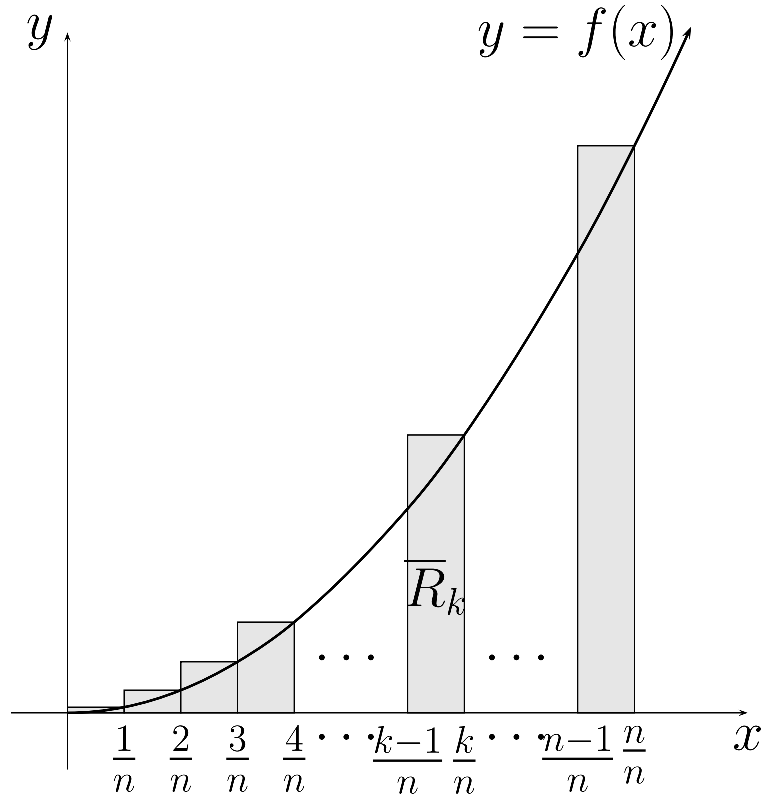
$$\text{area}(\Omega_1) \leq \text{area}(\Omega) \leq \text{area}(\Omega_2).$$

If $\lim_{n\to\infty} \text{area}(\Omega_1) = \lim_{n\to\infty} \text{area}(\Omega_2)$ then

$$A = \text{area}(\Omega) = \lim_{n\to\infty} \text{area}(\Omega_1) = \lim_{n\to\infty} \text{area}(\Omega_2).$$

**Remark.** One of Archimedes' best ideas was to estimate areas in circles and parabolas by approximating them by 'inner' and 'outer' polygons!

**Explict evaluation of the bounds.**




We begin by subdividing the interval $[0, 1]$ into $n$ subintervals

$$[0, \tfrac{1}{n}], [\tfrac{1}{n}, \tfrac{2}{n}], [\tfrac{2}{n}, \tfrac{3}{n}], \ldots, [\tfrac{n-1}{n}, 1].$$

The set $\mathcal{P}_n$ given by

$$\mathcal{P}_n = \{0, \tfrac{1}{n}, \tfrac{2}{n}, \tfrac{3}{n}, \ldots, \tfrac{n-1}{n}, 1\}$$

which divides the interval $[0, 1]$ into these subintervals is called a partition of $[0, 1]$. Let $\overline{R}_k$ denote the area of the $k$th rectangle. Then

$$\overline{R}_k = \text{width} \times \text{height} = \frac{1}{n} \times f(\tfrac{k}{n}) = \frac{k^2}{n^3}.$$

If $\overline{S}_{\mathcal{P}_n}(f)$ denotes the total area of the shaded region then

$$\overline{S}_{\mathcal{P}_n}(f) = \sum_{k=1}^{n} \overline{R}_k = \frac{1}{n^3} \sum_{k=1}^{n} k^2.$$

One may show by induction (exercise!) that

$$\sum_{k=1}^{n} k^2 = \frac{1}{6} n(n+1)(2n+1)$$

and hence

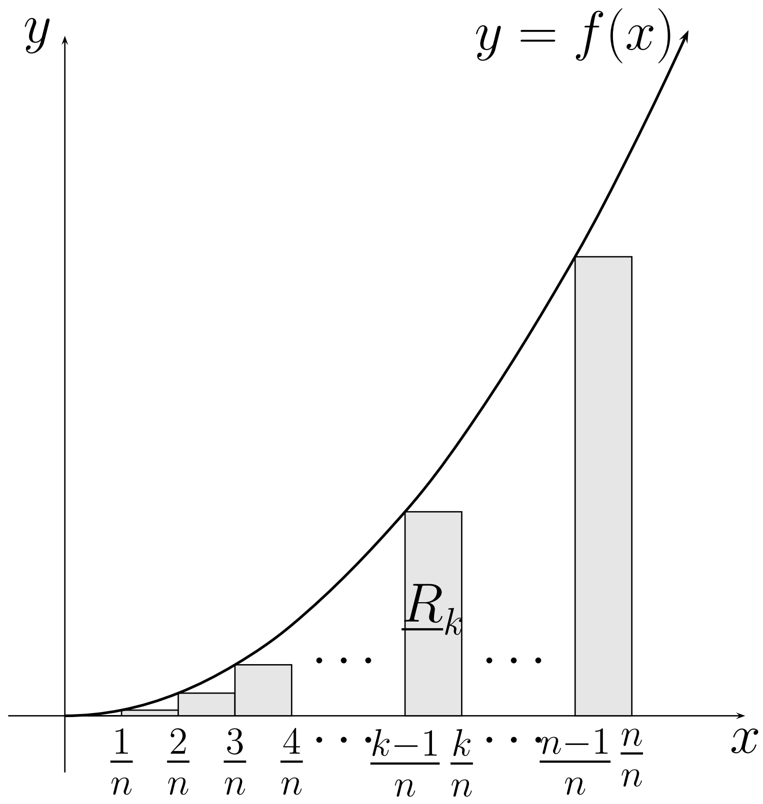
$$\overline{S}_{\mathcal{P}_n}(f) = \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}.$$

The quantity $\overline{S}_{\mathcal{P}_n}(f)$ is called the upper Riemann sum of $f$ with respect to the partition $\mathcal{P}_n$.

Axiom (A2) now implies that

$$A \leq \overline{S}_{\mathcal{P}_n}(f) = \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}.$$

In a similar manner, a lower bound is obtained:




The area $\underline{R}_k$ of the $k$th rectangle is given by

$$\underline{R}_k = \frac{1}{n} \times f(\tfrac{k-1}{n}) = \frac{(k-1)^2}{n^3}.$$

The sum of all the areas of the rectangles is called the lower Riemann sum for the function $f$ over the partition $\mathcal{P}_n$ and is denoted by $\underline{S}_{\mathcal{P}_n}(f)$. We obtain

$$\underline{S}_{\mathcal{P}_n}(f) = \sum_{k=1}^{n} \underline{R}_k = \frac{1}{n^3} \sum_{k=1}^{n} (k-1)^2$$

so that

$$\underline{S}_{\mathcal{P}_n}(f) = \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}.$$

Axiom (2) therefore implies that

$$A \geq \underline{S}_{\mathcal{P}_n}(f) = \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}.$$

Hence, for every positive integer $n$, the inequality

$$\frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2} \leq A \leq \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}$$

gives an upper and a lower bound for $A$.

**Conclusion.** In the limit $n \to \infty$, we obtain

$$A = \frac{1}{3},$$

regardless of the actual definition of $A$ as long as it is compatible with the axioms (A1)-(A5)!

**Remark.** The process of calculating upper and lower Riemann sums and taking a limit of the above type (provided it exists) is called integration.

### 8.1.3 The definition of area under the graph of a function and the Riemann integral

Suppose that $f$ is a bounded function on $[a, b]$ and that $f(x) \geq 0$ for all $x$ in $[a, b]$.

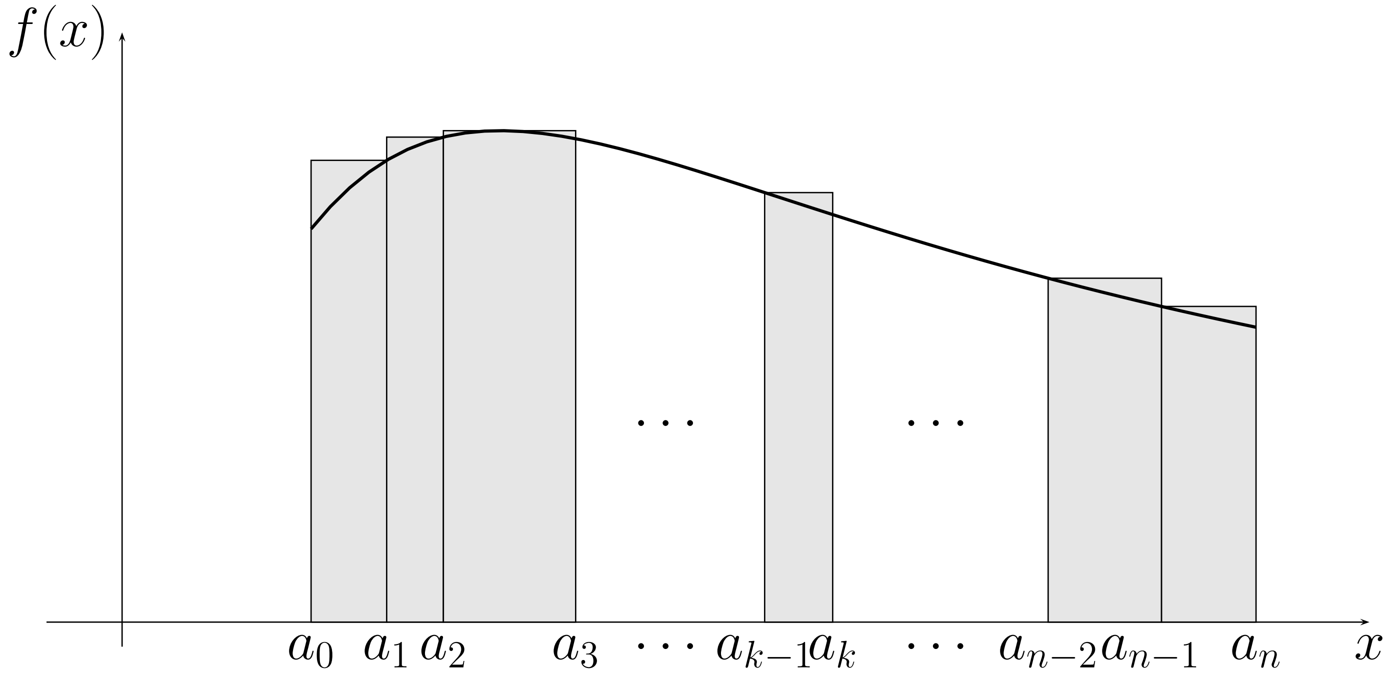
**Definition.** A finite set $\mathcal{P}$ of points in $\mathbb{R}$ is said to be a partition of $[a, b]$ if

$$\mathcal{P} = \{a_0, a_1, a_2, \ldots, a_n\}$$

and

$$a = a_0 < a_1 < a_2 < \ldots < a_n = b.$$

Suppose that $\mathcal{P}$ is a partition of $[a, b]$:




**Note.** The points of $\mathcal{P}$ need not be evenly spaced.

The area of the $k$th rectangle in the above figure is

$$\text{width} \times \text{height} = (a_k - a_{k-1}) \times \overline{f}_k,$$

where

$$\overline{f}_k = \text{the maximum value of } f \text{ on the subinterval } [a_k, a_{k-1}].$$

**Note.** Even though $f$ is bounded, the maximum value of $f$ on a given subinterval may not exist. Then, $\overline{f}_k$ is the least upper bound (MATH1241) of $f$ on that subinterval.

The upper Riemann sum $\overline{S}_{\mathcal{P}}(f)$ for $f$ with respect to the partition $\mathcal{P}$ is defined by

$$\overline{S}_{\mathcal{P}}(f) = \sum_{k=1}^{n} (a_k - a_{k-1})\overline{f}_k$$

which is the total area of the rectangles in the above figure.

Likewise, the lower Riemann sum $\underline{S}_{\mathcal{P}}(f)$ for $f$ with respect to the partition $\mathcal{P}$ is defined by

$$\underline{S}_{\mathcal{P}}(f) = \sum_{k=1}^{n} (a_k - a_{k-1})\underline{f}_k, \tag{8.1}$$

where

$$\underline{f}_k = \text{the minimum value of } f \text{ on the subinterval } [a_k, a_{k-1}]$$

or the greatest lower bound (MATH1241).

**Definition.** Suppose that a function $f$ is bounded on $[a, b]$ and that $f(x) \geq 0$ for all $x$ in $[a, b]$. If there exists a unique real number $A$ such that

$$\underline{S}_{\mathcal{P}}(f) \leq A \leq \overline{S}_{\mathcal{P}}(f)$$

for every partition $\mathcal{P}$ of $[a, b]$ then we say that $A$ is the area under the graph of $f$ from $a$ to $b$.

In general, if we remove the condition $f(x) \geq 0$ then the following definition is still sensible:

**Definition.** Suppose that a function $f$ is bounded on $[a,b]$. If there exists a unique real number $I$ such that

$$\underline{S}_{\mathcal{P}}(f) \leq I \leq \overline{S}_{\mathcal{P}}(f)$$

for every partition $\mathcal{P}$ of $[a,b]$ then we say that $f$ is Riemann integrable on the interval $[a,b]$.

The unique real number $I$ is called the definite integral of $f$ from $a$ to $b$ and we write

$$I = \int_a^b f(x)\,dx.$$

**Remark.** The function $f$ is called the integrand of the definite integral, while the points $a$ and $b$ are called the limits of the definite integral.

**Exercise.** Ponder about why this definition of area is consistent with the axioms (A1)-(A5).

**Historical remark.** The notation

$$\int_a^b f(x)\,dx$$

is due to Leibniz. It evolved from a slightly different way of writing down lower and upper Riemann sums. For example, $\overline{S}_{\mathcal{P}}(f)$ may be written as

$$\overline{S}_{\mathcal{P}}(f) = \sum_{k=1}^{n} f(\overline{x}_k)\Delta x_k,$$

where $\Delta x_k = a_k - a_{k-1}$ and $f$ attains its maximum value on $[a_{k-1}, a_k]$ at the point $\overline{x}_k$.

When taking a limit as before, $\Delta x_k$ was replaced with $dx$ and the symbol $\sum$ was replaced with an elongated 'S' ('S' stands for 'sum').

## 8.2 Integration with Riemann sums

Even in the case $f(x) = C$, where $C$ is a positive constant, the actual determination of the Riemann integral is somewhat tedious.

Indeed, let $\mathcal{P} = \{a_0 = a, a_1, \ldots, a_n = b\}$ be a partition of $[a, b]$. Since $f$ is constant, we have

$$\overline{f}_k = \underline{f}_k = C$$

for every $k$ between $1$ and $n$ and hence

$$\begin{aligned}
\overline{S}_{\mathcal{P}}(f) &= \sum_{k=1}^{n} \overline{f}_k (a_k - a_{k-1}) \\
&= C \sum_{k=1}^{n} (a_k - a_{k-1}) \\
&= C(b - a).
\end{aligned}$$

A similar calculation leads to

$$\underline{S}_{\mathcal{P}}(f) = C(b - a).$$

Accordingly,

$$\underline{S}_{\mathcal{P}}(f) = C(b-a) = \overline{S}_{\mathcal{P}}(f)$$

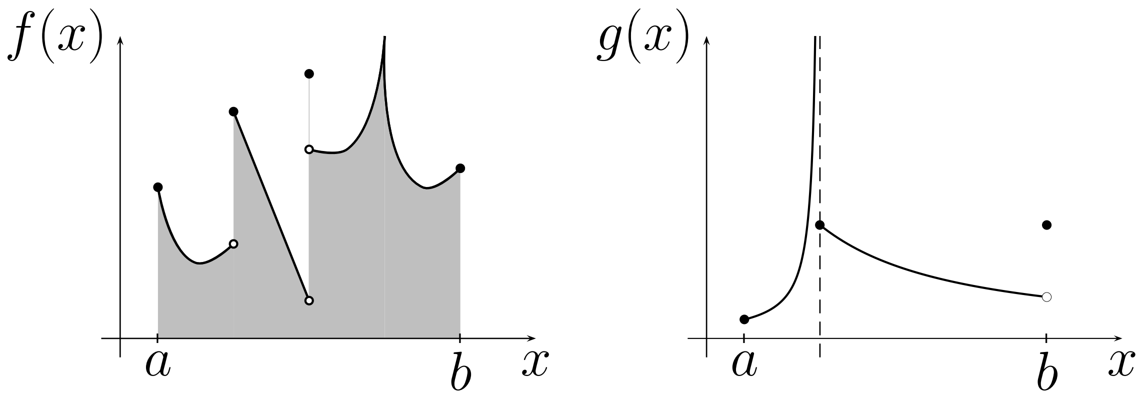
for every partition $\mathcal{P}$ of $[a, b]$ so that we conclude that $f$ is Riemann integrable and

$$\int_a^b f(x)\, dx = C(b-a).$$

**Question.** Can we find simple sufficient conditions which guarantee that a Riemann integral exists?

**Definition.** A function $f : [a, b] \to \mathbb{R}$ is said to be piecewise continuous if it is continuous on $[a, b]$ at all except perhaps a finite number of points.

**Examples.**




Both functions $f$ and $g$ are piecewise continuous but $f$ is bounded while $g$ is not!

**Theorem.** If $f$ is bounded and piecewise continuous on $[a, b]$ then $f$ is Riemann integrable on $[a, b]$.

**Proof.** Difficult!

**Remark.** If we know for some reason that $f$ is Riemann integrable on $[a,b]$ (e.g. because the above conditions are satisfied) then it is sufficient to show that

$$\lim_{n\to\infty} \overline{S}_{\mathcal{P}_n}(f) = \lim_{n\to\infty} \underline{S}_{\mathcal{P}_n}(f)$$

for some sequence of partitions $\mathcal{P}_n$ of $[a,b]$. If $I$ denotes this common limit, then

$$\int_a^b f(x)\,dx = I.$$

**Example (Fermat).** Find

$$\int_1^2 x^4\,dx,$$

using a 'clever' partition, i.e. rectangles of varying width!

**Exercise.** Find a Riemann integrable function which has infinitely many discontinuities.

**Example.** The function $f : [0, 1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

is not Riemann integrable.

**Proof.** Let $\mathcal{P}_n$ be an arbitrary partition of $[0, 1]$. Then,

$$\overline{f}_k = 1, \qquad \underline{f}_k = 0$$

for every $k$ between $1$ and $n$ so that

$$\overline{S}_{\mathcal{P}_n}(f) = 1, \qquad \underline{S}_{\mathcal{P}_n}(f) = 0.$$

Hence, there is no unique number $I$ satisfying

$$\underline{S}_{\mathcal{P}}(f) \leq I \leq \overline{S}_{\mathcal{P}}(f)$$

so that $f$ is not Riemann integrable.

**Remark.** There exist more sophisticated ways of 'measuring' areas, volumes etc. such as Lebesque integration.

- If $f$ is Riemann integrable then it is Lebesque integrable.
- If $f$ is not Riemann integrable, it may still be Lebesque integrable. The Lebesque integral of the above example is 0!
- However, there exist regions in $\mathbb{R}^2$ to which it is impossible to assign an area in any meaningful way! (Cf. Banach-Tarski paradox in three dimensions).

## 8.3 The Riemann integral and signed area

If $f$ is a function which is Riemann integrable but not necessarily non-negative then it is natural to refer to

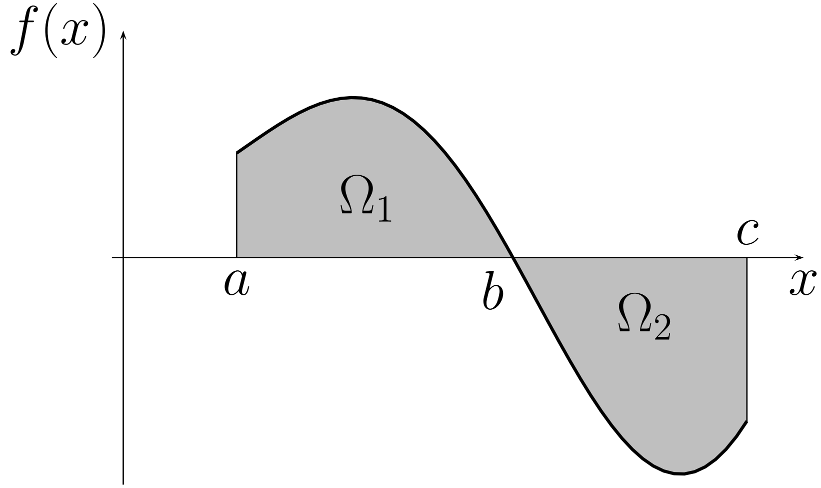
$$\int_a^c f(x)\,dx$$

as the signed area under the graph of the function $f$ from $a$ to $c$ and to

$$\text{area}(\Omega) = \int_a^c |f(x)|\,dx$$

as the unsigned area under the graph of the function $f$ from $a$ to $c$ provided that the latter integral exists.

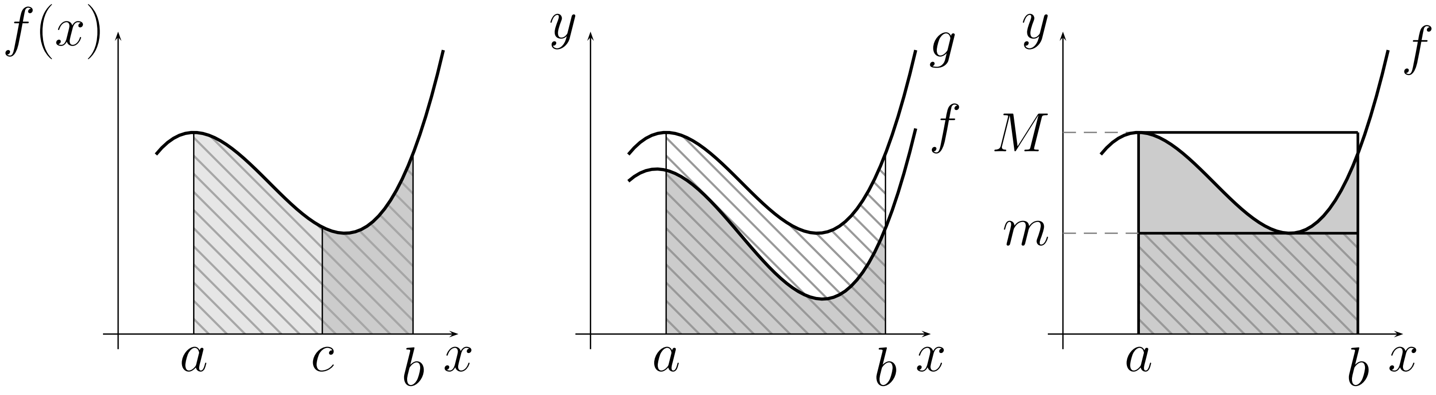
Consider the following example:




In this case, the unsigned area is

$$\text{area}(\Omega) = \int_a^b f(x)\,dx - \int_b^c f(x)\,dx.$$

## 8.4 Basic properties of the Riemann integral




For brevity, from now on, we refer to Riemann integrable functions as merely 'integrable'.

**Theorem.** Suppose that $f, g : [a, b] \to \mathbb{R}$ are integrable. Then,

(i) (Linearity) $\alpha f + \beta g$ is integrable for any $\alpha, \beta \in \mathbb{R}$ with

$$\int_a^b (\alpha f + \beta g)(x)\, dx = \alpha \int_a^b f(x)\, dx + \beta \int_a^b g(x)\, dx.$$

(ii) If $a < c < b$ then

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

(iii) If $f(x) \geq 0$ for all $x$ in $[a, b]$ then

$$\int_a^b f(x)\,dx \geq 0.$$

(iv) If $f(x) \leq g(x)$ for all $x$ in $[a, b]$ then

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

(v) If $m \leq f(x) \leq M$ for all $x$ in $[a, b]$ then

$$m(b-a) \leq \int_a^b f(x)\,dx \leq M(b-a).$$

(vi) If $|f|$ is integrable on $[a, b]$ then

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx.$$

**Sketch of proof.**

(iii) If $f(x) \geq 0$ then

$$0 \leq \underline{S}_{\mathcal{P}}(f) \leq \int_a^b f(x)\,dx.$$

(iv) Apply (iii) to $h = g - f$.

(v) From (iv), it follows that

$$\int_a^b m\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b M\,dx.$$

(vi) From

$$-|f(x)| \leq f(x) \leq |f(x)| \qquad \forall x \in [a, b]$$

and (iv), we deduce that

$$-\int_a^b |f(x)|\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b |f(x)|\,dx.$$

It is convenient to introduce the following definition ...

**Definition.** Suppose that $b < a$ and that $f$ is integrable on $[b, a]$. Then,

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$$

and

$$\int_a^a f(x)\,dx = 0.$$

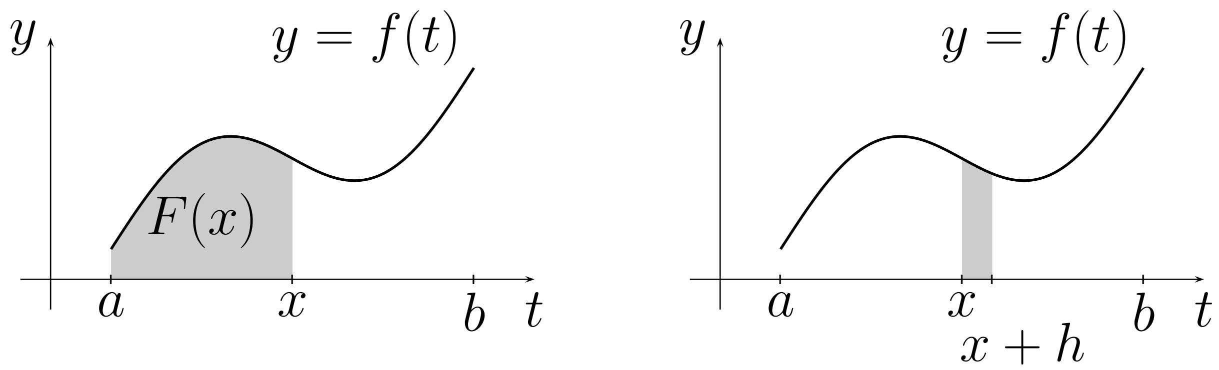
... since, for instance, the following version of (ii) is still valid.

**Remark.** Suppose that $a$, $b$ and $c$ are real numbers and that $f$ is integrable over some interval containing $a$, $b$ and $c$. Then

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

## 8.5 The first fundamental theorem of calculus

**Idea (Leibniz, Newton).** How does the integral (area) change as a boundary changes?




Suppose that a function $f$ is continuous and therefore integrable on an interval $[a, b]$.

We define the area function $F$ by

$$F(x) = \int_a^x f(t)\,dt$$

for all $x \in [a, b]$.

Then,
$$\frac{F(x+h)-F(x)}{h} \approx f(x),$$
where $h$ is a 'small' positive number.

Thus, we are tempted to believe that $F$ is differentiable and
$$F'(x) = \lim_{h\to 0} \frac{F(x+h)-F(x)}{h} = f(x).$$

**The first fundamental theorem of calculus.** Suppose that $f : [a, b] \to \mathbb{R}$ is a continuous function. Then, the function $F : [a, b] \to \mathbb{R}$ defined by
$$F(x) = \int_a^x f(t)\,dt \tag{8.2}$$
is continuous on $[a, b]$ and differentiable on $(a, b)$ with
$$F'(x) = f(x)$$
for all $x$ in $(a, b)$.

**Implications.**

- Every continuous function $f : [a, b] \to \mathbb{R}$ has an antiderivative $F$ on $(a, b)$ given by

$$F(x) = \int_a^x f(t)\, dt.$$

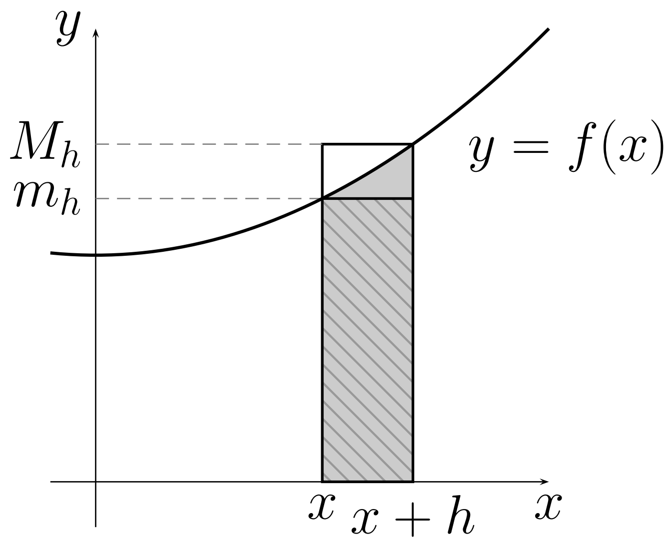
- Since two antiderivatives of $f$ differ by a constant, integration and antidifferentiation are essentially the same procedures!

- Differentiation undoes what integration does to $f$ since

$$f(x) = \frac{d}{dx} \left( \int_a^x f(t)\, dt \right).$$

  Is the converse true?

**Proof of the 1. FTC.**




Differentiability.

Suppose that $x \in (a,b)$ and choose an $h > 0$ such that $x + h \in (a,b)$. Then,

$$F(x+h) - F(x) = \int_x^{x+h} f(t)\,dt.$$

Since $f$ is continuous on $[a,b]$, it attains a minimum value $m_h$ and a maximum value $M_h$ on $[x, x+h]$, that is,

$$m_h \leq f(t) \leq M_h$$

for all $t \in [x, x+h]$.

We therefore conclude that

$$m_h h \leq \int_x^{x+h} f(t)\,dt \leq M_h h$$

and hence

$$m_h h \leq F(x+h) - F(x) \leq M_h h$$

so that

$$m_h \leq \frac{F(x+h) - F(x)}{h} \leq M_h$$

since $h > 0$.

Now, since $f$ is continuous on $[x, x+h]$, we know that

$$\lim_{h \to 0^+} m_h = f(x) = \lim_{h \to 0^+} M_h$$

and hence

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

by virtue of the pinching theorem.

In a similar manner, one can show that

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

Hence, $F$ is differentiable on $(a, b)$ and

$$F'(x) = f(x)$$

for all $x$ in $(a, b)$.

Continuity.

$F$ is differentiable on $(a, b)$ and is therefore continuous on $(a, b)$.

Since $f$ is continuous on $[a, b]$, it is bounded on $[a, b]$, that is, there exists an $M > 0$ such that

$$|f(x)| \leq M$$

for all $x$ in $[a, b]$.

Hence, for $x \in (a, b)$, we obtain:

$$\begin{aligned}
|F(x) - F(a)| &= \left| \int_a^x f(t)\,dt - \int_a^a f(t)\,dt \right| \\
&= \left| \int_a^x f(t)\,dt \right| \\
&\leq \int_a^x |f(t)|\,dt \\
&\leq \int_a^x M\,dt \\
&= M|x - a|.
\end{aligned}$$

Accordingly,

$$|F(x) - F(a)| \leq M|x - a| \to 0$$

as $x \to a^+$. Thus, $F$ is continuous at $a$.

Similarly, $F$ is also continuous at the endpoint $b$.

**Remark.** The above proof of continuity of $F$ at the endpoints could be applied to any point on $[a, b]$. Thus, for continuity of $F$ on $[a, b]$, it is actually only required that $f$ is integrable and bounded!

**Note.** Many important functions are actually defined as area functions such as

$$\ln(x) = \int_1^x \frac{1}{t}\,dt$$

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,dt$$

$$\operatorname{Si}(x) = \int_0^x \frac{\sin t}{t}\,dt.$$

The first fundamental theorem of calculus then implies that these functions are continuous and differentiable on the respective (closed or open) intervals.

## 8.6 The second fundamental theorem of calculus

There exists a fast way of calculating an integral if an explicit antiderivative is known.

**The second fundamental theorem of calculus.** Suppose that $f$ is a continuous function on $[a, b]$. If $F$ is an antiderivative of $f$ on $[a, b]$, that is,

$$F'(x) = f(x)$$

for all $x \in [a, b]$, then

$$\int_a^b f(t)\,dt = F(b) - F(a).$$

**Proof.** We know that two antiderivatives differ by a constant and hence

$$F(x) = \int_a^x f(t)\,dt + C$$

for some constant $C$. Accordingly,

$$F(b) - F(a) = \int_a^b f(t)\,dt + C - 0 - C = \int_a^b f(t)\,dt.$$

**Notation.** We frequently use the notation

$$F(x)\Big|_a^b \qquad \text{or} \qquad \Big[F(x)\Big]_a^b$$

for the expression $F(b) - F(a)$.

**Example.** Calculate the area under the curve

$$y = f(x) = 3x^2 e^{x^3}$$

between $x = 0$ and $x = 1$.

It is easy to verify that $F$ given by

$$F(x) = e^{x^3}$$

is an antiderivative of $f$, that is, $F'(x) = f(x)$, and hence

$$A = F(1) - F(0) = e - 1.$$

**Question.** How did we find the antiderivative of the above function $f$?

## 8.7 Indefinite integrals

**Notation.** We denote the complete class of antiderivatives of a function $f$ defined on a suitable interval by

$$\int f(x)\,dx.$$

The latter is called the indefinite integral of $f$. If $F$ is an antiderivative of the function $f$ then

$$\int f(x)\,dx = F(x) + C,$$

where $C$ is an arbitrary constant of integration.

**Examples.** It is easy to verify that

$$\int x^n\,dx = \frac{1}{n+1}x^{n+1} + C, \qquad n \neq -1$$

$$\int \sin x\,dx = -\cos x + C$$

$$\int e^{ax}\,dx = \frac{1}{a}e^{ax} + C, \qquad a \neq 0$$

$$\int \frac{2x+1}{\sqrt{3x+2}}\,dx = \frac{2}{27}\sqrt{3x+2}\,(1+6x) + C$$

$$\int \frac{f'(x)}{f(x)}\,dx = \ln|f(x)| + C.$$

## 8.8 Integration by substitution

**Question.** What is the analogue of the chain rule

$$\frac{d}{dx}F(g(x)) = F'(g(x))g'(x)$$

in the context of integration?

Formally, if $F$ is an antiderivative of a function $f$ then

$$\int f(g(x))g'(x)\,dx = F(g(x)) + C,$$

since

$$\frac{d}{dx}F(g(x)) = F'(g(x))g'(x) = f(g(x))g'(x).$$

**Previous example.**

$$\int 3x^2 e^{x^3}\,dx = \int e^{x^3}\left(\frac{d}{dx}x^3\right)dx = e^{x^3} + C.$$

**Example.** Find
$$\int \frac{1}{x \ln x}\, dx.$$

It is not always easy to recognise an integrand as being of the form
$$F'(g(x))g'(x).$$

In this case, the 'mechanical' procedure which might lead to a successful integration is called integration by substitution.

Thus, if we set

$$u = g(x)$$

then

$$\begin{aligned}\int f(g(x))g'(x)\,dx &= F(g(x)) + C \\ &= F(u) + C \\ &= \int f(u)\,du.\end{aligned}$$

The shorthand for the above operation is

$$u = g(x), \qquad du = g'(x)\,dx$$

or, if appropriate,

$$x = h(u), \qquad dx = h'(u)\,du.$$

**Example.** Find

$$I = \int \frac{2x+1}{\sqrt{3x+2}}\,dx.$$

Set

$$u = \sqrt{3x+2} \quad \Rightarrow \quad du = \frac{3}{2\sqrt{3x+2}}\,dx, \qquad x = \frac{u^2-2}{3}.$$

Hence,

$$\begin{aligned} I &= \int \left( \frac{2(u^2-2)}{3} + 1 \right) \frac{2}{3}\,du \\ &= \int \frac{4}{9}u^2 - \frac{2}{9}\,du \\ &= \frac{4}{27}u^3 - \frac{2}{9}u + C \\ &= \frac{4}{27}(3x+2)^{3/2} - \frac{2}{9}\sqrt{3x+2} + C. \end{aligned}$$

**Note.** The last step always consists of rewriting everything in terms of the original variable!

The precise statement in the case of definite integrals is the following:

**Theorem.** Suppose that $g$ is a differentiable function such that $g'$ is continuous on $[a, b]$. If $f$ is continuous on any interval $I$ containing $g([a, b])$ then the change of variables formula

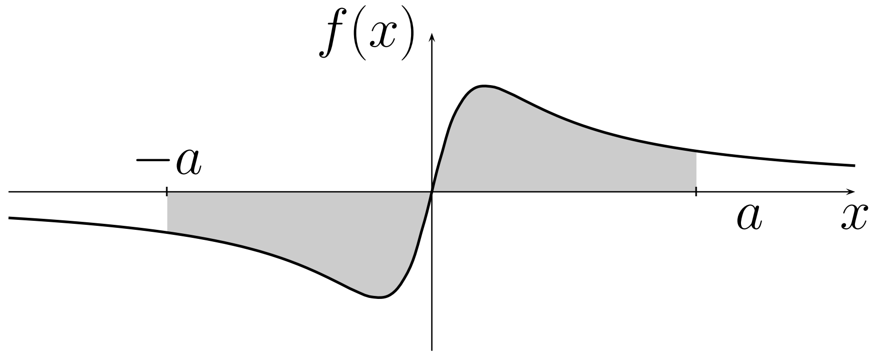
$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du$$

holds.

**Proof.** Exercise.

**Example.** Evaluate the definite integral

$$I = \int_0^1 \frac{1}{\sqrt{4 - x^2}}\,dx.$$




**Application.** Suppose that $f$ is a continuous function and $a$ is a real number.

(i) If $f$ is even then

$$\int_{-a}^{a} f(x)\,dx = 2\int_{0}^{a} f(x)\,dx.$$

(ii) If $f$ is odd then

$$\int_{-a}^{a} f(x)\,dx = 0.$$

(iii) If $f$ is periodic with period $T$ then

$$\int_{a}^{a+T} f(x)\,dx = \int_{0}^{T} f(x)\,dx.$$

**Example.** Find

$$\int_{\frac{\pi}{17}}^{\frac{37\pi}{51}} \sin^{99} 3x \, dx.$$

## 8.9 Integration by parts

**Question.** What is the analogue of the product rule

$$[f(x)g(x)]' = f'(x)g(x) + f(x)g'(x)$$

in the context of integration?

Integration of both sides leads to

$$f(x)g(x) = \int f'(x)g(x)\,dx + \int f(x)g'(x)\,dx.$$

Hence, we obtain the integration by parts formula

$$\int f(x)g'(x)\,dx = f(x)g(x) - \int f'(x)g(x)\,dx.$$

The corresponding formula for definite integrals reads

$$\int_a^b f(x)g'(x)\,dx = \Big[f(x)g(x)\Big]_a^b - \int_a^b f'(x)g(x)\,dx.$$

Sometimes, it may be relatively easy to apply the above formulae...

**Example.** Find the indefinite integral

$$I = \int x \cos x \, dx.$$

It is evident that

$$\begin{aligned} I &= \int x \left( \frac{d}{dx} \sin x \right) \, dx \\ &= x \sin x - \int \left( \frac{d}{dx} x \right) \sin x \, dx \\ &= x \sin x + \cos x + C. \end{aligned}$$

... on other occasions, it may not be so easy:

**Example.** Find the definite integral

$$I = \int_1^e \ln x \, dx.$$

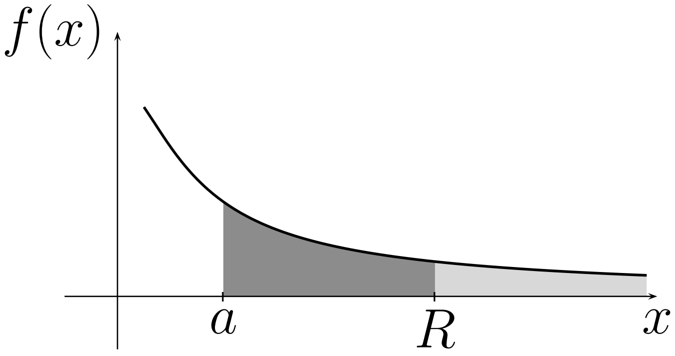
## 8.10 Improper integrals

**Question.** In many applications (e.g. probability theory, quantum mechanics etc.), the class of integrals of bounded functions over finite intervals is too restrictive.

For instance, we could consider the improper integral

$$\int_a^\infty f(x)\,dx,$$

but what is its exact definition?



**Definition.**

(a) Suppose that there exists a real number $L$ such that

$$\int_a^R f(x)\,dx \to L$$

as $R \to \infty$. Then, $f$ is said to be integrable over $[a, \infty)$. We say that the improper integral

$$\int_a^\infty f(x)\,dx$$

is convergent and write

$$\int_a^\infty f(x)\,dx = L.$$

(b) Suppose that

$$\int_a^R f(x)\,dx$$

does not have a limit as $R \to \infty$. Then, we say that $f$ is not integrable over $[a, \infty)$ and the improper integral

$$\int_a^\infty f(x)\,dx$$

is said to be divergent.

**Remark.** A similar definition for improper integrals of the form $\int_{-\infty}^{b} f(x)\,dx$ applies.

**Example.** The volume of a solid bounded by a curve $y = f(x)$ rotated about the $x$-axis and the planes $x = a$ and $x = b$ is given by

$$V = \pi \int_{a}^{b} [f(x)]^2\,dx.$$

Show that the volume of the body generated by the curve $y = 1/x$ between $1$ and "$\infty$" is

$$V = \pi \int_{1}^{\infty} \frac{1}{x^2}\,dx = \pi$$

and, hence, finite.

**A "paradox".** The surface area of a solid of the above type (excluding the 'vertical' boundaries) is given by

$$A = 2\pi \int_a^b f(x)\sqrt{1 + [f'(x)]^2}\, dx.$$

Hence, for the above example, we obtain

$$\begin{aligned} A &= 2\pi \int_1^b \frac{1}{x}\sqrt{1 + \frac{1}{x^4}}\, dx \\ &= \pi \left[ \ln\left(x^2 + \sqrt{x^4+1}\right) - \frac{\sqrt{x^4+1}}{x^2} \right]_1^b \end{aligned}$$

[Verify this!] $\quad \to \infty \quad$ as $\quad b \to \infty.$

Thus, the improper integral

$$\int_1^\infty \frac{1}{x}\sqrt{1 + \frac{1}{x^4}}\, dx$$

diverges and the surface area is infinite!

**Question.** How would one define the improper integral

$$\int_{-\infty}^{\infty} f(x)\,dx\,?$$

**Definition.** We say that $f$ is integrable over $(-\infty, \infty)$ if $f$ is integrable over both $(-\infty, 0]$ and $[0, \infty)$. In this case, we write

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx.$$

If $f$ is not integrable on either of the intervals $(-\infty, 0]$ or $[0, \infty)$ then we say that the improper integral

$$\int_{-\infty}^{\infty} f(x)\,dx$$

diverges.

**Example.** Does the improper integral

$$\int_{-\infty}^{\infty} x\,dx$$

converge?

Since

$$\int_0^R x\,dx = \left[\frac{1}{2}x^2\right]_0^R = \frac{R^2}{2} \to \infty$$

as $R \to \infty$, the improper integral

$$\int_{-\infty}^{\infty} x\,dx$$

diverges by definition.

**Remark.** The fact that

$$\lim_{R\to\infty} \int_{-R}^{R} x\,dx = \lim_{R\to\infty} 0 = 0$$

is irrelevant

**Theorem (Convergence and divergence of $p$-integrals).** The improper integral

$$\int_1^\infty \frac{1}{x^p}\,dx$$

is convergent if $p > 1$ and divergent if $p \le 1$.

**Proof.** If $p \neq 1$ then

$$\begin{aligned}
\int_1^R x^{-p}\,dx &= \left[\frac{x^{1-p}}{1-p}\right]_1^R \\
&= \frac{R^{1-p}-1}{1-p} \\
&\to \begin{cases} \frac{1}{p-1} & \text{when } 1-p<0 \\ \infty & \text{when } 1-p>0 \end{cases}
\end{aligned}$$

as $R \to \infty$.

If $p = 1$ then

$$\int_1^R \frac{1}{x}\,dx = \Big[\ln x\Big]_1^R = \ln R - \ln 1 \to \infty \quad \text{as } R \to \infty.$$

**Remark.** If the logarithm is defined as an area function then it is required to prove (using Riemann sums) that it is unbounded!

## 8.11 Comparison tests for improper integrals

Some important improper integrals such as

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx$$

(probability theory, quantum mechanics) do not admit 'elementary' antiderivatives but one may still be able to discuss convergence/divergence by comparing them with known improper integrals.

**Theorem (Comparison test).** Suppose that $f$ and $g$ are integrable functions and that

$$0 \leq f(x) \leq g(x)$$

for $x \geq a$.

(i) If $\displaystyle\int_a^{\infty} g(x)\,dx$ converges then $\displaystyle\int_a^{\infty} f(x)\,dx$ converges.

(ii) If $\displaystyle\int_a^{\infty} f(x)\,dx$ diverges then $\displaystyle\int_a^{\infty} g(x)\,dx$ diverges.

**Sketch of proof.** This follows from

$$0 \leq \int_a^R f(x)\,dx \leq \int_a^R g(x)\,dx$$

and the Least Upper Bound Axiom (MATH1241).

**Example.** Discuss the convergence of

$$I = \int_{-\infty}^{\infty} e^{-x^2}\, dx.$$

The integrand is even and hence

$$I = 2\int_0^{\infty} e^{-x^2}\, dx = 2\int_0^{1} e^{-x^2}\, dx + 2\int_1^{\infty} e^{-x^2}\, dx.$$

On the other hand, we know that

$$e^{x^2} > x^2 \qquad \Rightarrow \qquad e^{-x^2} < \frac{1}{x^2}$$

for $x \neq 0$ and that the $p$-integral

$$\int_1^{\infty} \frac{1}{x^2}\, dx$$

converges.

We therefore conclude that $I$ converges.

**Note.** One may show (2nd year) that $I = \sqrt{\pi}$.

**Example.** Discuss the convergence of

$$I = \int_2^\infty \frac{1}{\ln x}\, dx.$$

Instead of using inequalities to estimate integrands, one often uses a ‘dominant term analysis’ such as

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

‘behaves like’

$$g(x) = \frac{1}{2x^3}$$

for large $x$ and hence one expects the convergence of the two associated improper integrals to be the same.

The precise formulation of this idea is as follows:

**Theorem (Limit form of the comparison test.)** Suppose that $f$ and $g$ are nonnegative and bounded on $[a, \infty)$. If

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

and $0 < L < \infty$ then either

$$\text{both} \quad \int_a^\infty f(x)\,dx \quad \text{and} \quad \int_a^\infty g(x)\,dx \quad \text{converge}$$

or

$$\text{both} \quad \int_a^\infty f(x)\,dx \quad \text{and} \quad \int_a^\infty g(x)\,dx \quad \text{diverge.}$$

**Proof.** See calculus notes.

**Example.** Discuss the convergence of

$$I = \int_2^\infty \frac{\sqrt{\sin x + x^2}}{2x^4 - 1}\, dx.$$

Set

$$f(x) = \frac{\sqrt{\sin x + x^2}}{2x^4 - 1}, \qquad g(x) = \frac{1}{x^3}.$$

Since

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = \frac{1}{2}$$

and the $p$-integral

$$\int_2^\infty \frac{1}{x^3}\, dx$$

converges, we conclude that $I$ converges.

**Example.** Discuss the convergence of

$$I = \int_3^\infty (\sqrt{x} - \sqrt{x-3}\,)\, dx.$$

# Chapter 9

# The logarithmic and exponential functions

In the preceding, we have manipulated functions such as

$$\ln x, \quad e^x, \quad x^\pi$$

even though we have not defined them formally.

In particular, we are familiar with the important formula

$$\ln(xy) = \ln x + \ln y.$$

**Question.** Consider the functional equation

$$f(st) = f(s) + f(t), \tag{9.1}$$

where $s$ and $t$ are independent variables. It is evident that $f = \ln$ constitutes one solution of this functional equation. Are there other functions $f$ obeying this functional equation?

**Answer.** We first note that (9.1) evaluated at $s = t = 1$ yields

$$f(1) = 0.$$

Moreover, differentiation with respect to $s$ leads to

$$tf'(st) = f'(s)$$

so that

$$f'(t) = \frac{1}{t}f'(1)$$

at $s = 1$.

If we now demand that $f'(1) = 1$ then $f$ is uniquely determined via integration since $f(1) = 0$.

**Conclusion.** A function $f$ is uniquely defined by the functional equation

$$f(st) = f(s) + f(t),$$

subject to

$$f'(1) = 1.$$

It is given by

$$f(x) = \int_1^x \frac{1}{t}\, dt.$$

## 9.1 Powers and logarithms

See calculus notes.

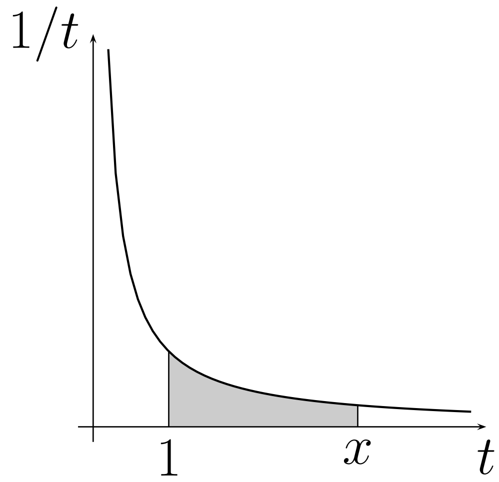
## 9.2 The natural logarithm function

**Definition.** The natural logarithm function

$$\ln : (0, \infty) \to \mathbb{R}$$

is defined by the formula

$$\ln x = \int_1^x \frac{1}{t}\, dt.$$




It is evident that $\ln x$ is simply the area of the shaded region shown below in the case when $x \geq 1$.

**Theorem.** The function $\ln : (0, \infty) \to \mathbb{R}$ has the following properties:

(i) $\ln$ is differentiable on $(0, \infty)$ and

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

(ii) $\ln x > 0 \quad \text{for} \quad x > 1$,

$\ln 1 = 0$,

$\ln x < 0 \quad \text{for} \quad 0 < x < 1.$

(iii) $\ln x \to -\infty \quad \text{as} \quad x \to 0^+$,

$\ln x \to \infty \quad \text{as} \quad x \to \infty.$

(iv) For all $x, y > 0$: $\qquad \ln(xy) = \ln x + \ln y.$

(v) For all $x, y > 0$: $\qquad \ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y).$

(vi) For all $x > 0$ and $r \in \mathbb{Q}$: $\qquad \ln(x^r) = r \ln x.$

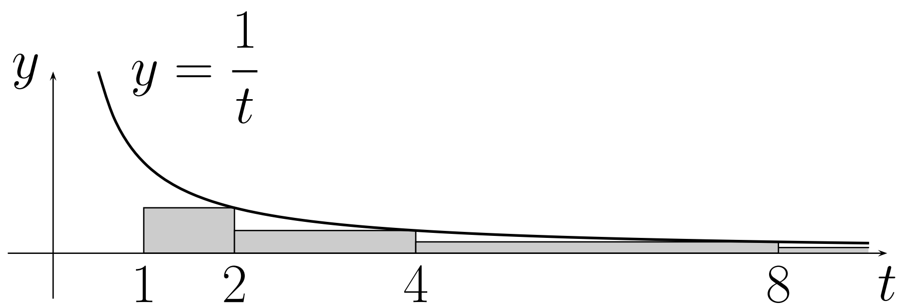
**Proof.**

(i) Apply the first fundamental theorem of calculus to the definition of $\ln$.

(ii) This follows from the definition of $\ln$ and the fact that $\frac{1}{t} > 0$ when $t > 0$.

(iii) The diagram below shows that

$$\int_1^2 \frac{dt}{t} \geq 1 \times \frac{1}{2}, \quad \int_2^4 \frac{dt}{t} \geq 2 \times \frac{1}{4}, \quad \int_4^8 \frac{dt}{t} \geq 4 \times \frac{1}{8}.$$



In general,

$$\int_1^{2^n} \frac{dt}{t} \geq \underbrace{\frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2}}_{n \text{ terms}}$$
$$= \frac{n}{2} \to \infty$$

as $n \to \infty$.

Hence the improper integral

$$\int_1^{\infty} \frac{1}{t}\, dt$$

'diverges to infinity' and, therefore, $\ln x \to \infty$ as $x \to \infty$.

This argument can be adapted to show that

$$\int_{2^{-n}}^{1} \frac{dt}{t} \to \infty$$

as $n \to \infty$. Hence $\ln x \to -\infty$ as $x \to 0^+$.

(iv) Suppose that $y$ is some fixed positive number and that $x > 0$. Then, the chain rule implies that

$$\frac{d}{dx}[\ln(xy)] = \frac{1}{xy}\frac{d}{dx}(xy) = \frac{y}{xy} = \frac{1}{x} = \frac{d}{dx}\ln x.$$

Accordingly,

$$\ln(xy) = \ln(x) + C$$

for some constant $C$.

Evaluation at $x = 1$ leads to

$$\ln(y) = C$$

and hence

$$\ln(xy) = \ln(x) + \ln(y).$$

(v) Similar technique.

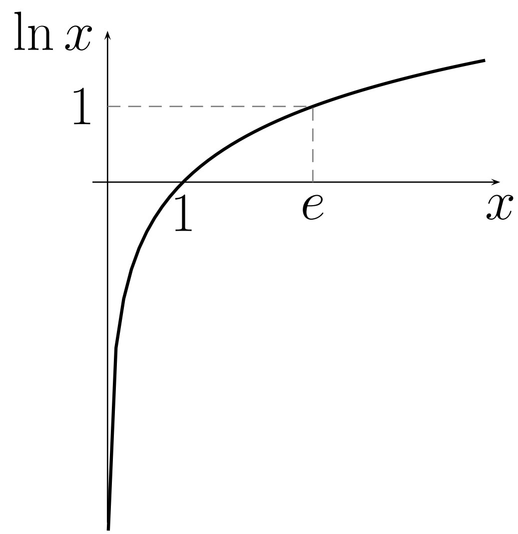
(vi) Similar technique.

**Remark.** The above properties imply that

- $\mathrm{Range}(\ln) = \mathbb{R}$, and
- $\ln$ is increasing and hence invertible so that
- $\ln x = 1$ has a unique solution.

**Definition.** The real number $e$ is defined to be the unique number $x$ satisfying

$$\int_1^x \frac{1}{t}\,dt = 1.$$




**Note.** $\dfrac{d}{dx}\ln x = \dfrac{1}{x} \to 0 \quad$ as $\quad x \to \infty$.

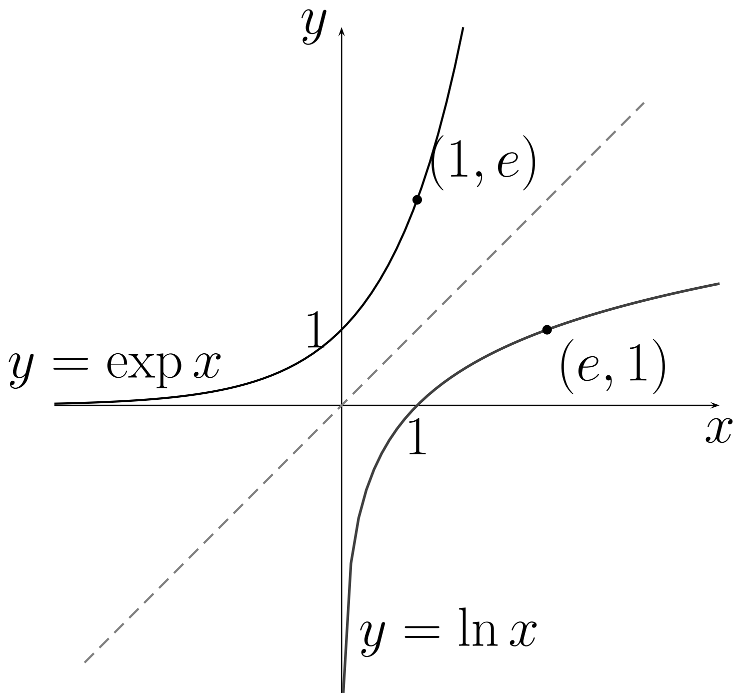
## 9.3 The exponential function

**Question.** It is recalled that we have already defined $x^r$ for $r \in \mathbb{Q}$. Can this definition be extended to $r \notin \mathbb{Q}$?

**Definition.** The function

$$\exp : \mathbb{R} \to (0, \infty)$$

is defined to be the inverse function of $\ln : (0, \infty) \to \mathbb{R}$.



**Remark.** For any rational number $r$, we can evaluate both

$$\exp r \qquad \text{and} \qquad e^r$$

but are these two numbers the same?

**Theorem.** The function $\exp : \mathbb{R} \to (0, \infty)$ has the following properties:

(i) $\exp(\ln x) = x \quad$ for all $\quad x \in (0, \infty)$,
$\ln(\exp x) = x \quad$ for all $\quad x \in \mathbb{R}$.

(ii) $\exp(1) = e$,
$\exp(0) = 1$.

(iii) $\exp x \to \infty \quad$ as $\quad x \to \infty$,
$\exp x \to 0 \quad$ as $\quad x \to -\infty$.

(iv) $\exp$ is differentiable on $\mathbb{R}$ with

$$\frac{d}{dx}\exp x = \exp x.$$

(v) For all $x, y \in \mathbb{R}$:

$$\exp(x+y) = \exp x\ \exp y.$$

(vi) For all $x > 0$ and $r \in \mathbb{Q}$:

$$\exp(rx) = (\exp x)^r.$$

**Proof.**

(i) - (iii) Follows from the definition of $\exp$.

(iv) The function $\exp$ is differentiable on $\mathbb{R}$ by virtue of the inverse function theorem and differentiation of

$$\ln(\exp x) = x$$

produces

$$\frac{1}{\exp x} \frac{d}{dx} \exp x = 1$$

as required.

(v) For any $x$ and $y$, we have

$$\begin{aligned} \exp(x+y) &= \exp\Big( \ln(\exp x) + \ln(\exp y) \Big) \\ &= \exp\Big( \ln(\exp x \, \exp y) \Big) \\ &= \exp x \, \exp y \end{aligned}$$

as required.

(vi) Suppose that $r$ is a rational number and $x$ is a real number. Then,

$$\begin{aligned}
\exp(rx) &= \exp\Big(r \ln \exp x)\Big) \\
&= \exp\Big(\ln\big((\exp x)^r\big)\Big) \\
&= (\exp x)^r.
\end{aligned}$$

**Remark.** In particular, the above theorem implies that

$$\exp(r) = (\exp 1)^r = e^r.$$

for every rational number $r$. It is therefore consistent to make the following definition.

**Definition.** For any $x \notin \mathbb{Q}$, we define the number $e^x$ to be

$$e^x = \exp x.$$

**Note.** The above definition 'merely' means that $e^x$ is the unique real number $R$ such that

$$\int_1^R \frac{1}{t}\, dt = x$$

for any $x \in \mathbb{R}$. By construction, the function

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = \exp x = e^x$$

is differentiable (and continuous) and is called the exponential function.

## 9.4 Exponentials and logarithms with other bases

**Question.** How would one define $b^x$ for $x \notin \mathbb{Q}$ and $b > 0$?

Since, for any rational number $r$,

$$b^r = \exp\big(\ln(b^r)\big) = \exp\big(r \ln b\big) = e^{r \ln b},$$

the following definition is natural.

**Definition.** Suppose that $b > 0$ and $x \notin \mathbb{Q}$. Then, the number $b^x$ is defined by

$$b^x = \exp(x \ln b) = e^{x \ln b}.$$

**Note.** By combining the defintions of $b^x$ for rational $x$ and irrational $x$, we now obtain a well-defined function

$$f_b : \mathbb{R} \to (0, \infty), \quad f_b(x) = b^x = \exp(x \ln b)$$

for any $b > 0$.

**Example.** It is seen that

$$f_3(2) = 3^2 = \exp(2\ln 3) = \exp(\ln 3)\ \exp(\ln 3) = 3 \times 3 = 9$$

as one would expect!

**Remark.** Since $f_b$ is a combination of continuous and differentiable functions, it is also continuous and differentiable with

$$f_b'(x) = (\ln b)\exp(x \ln b) = (\ln b) b^x.$$

Accordingly,

- if $\quad b > 1 \quad$ then $\quad f_b'(x) > 0 \quad$ for all $\quad x \in \mathbb{R}$,
- if $\quad 0 < b < 1 \quad$ then $\quad f_b'(x) < 0 \quad$ for all $\quad x \in \mathbb{R}$

so that $f_b$ is invertible for $b \neq 1$.

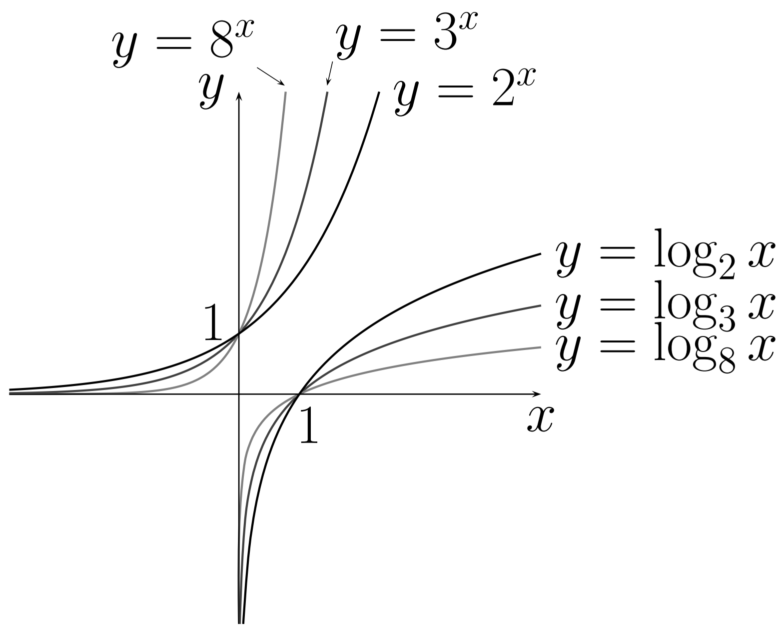
**Definition.** Suppose that $b$ is a positive real number with $b \neq 1$. Then, the logarithm function to the base $b$

$$\log_b : (0, \infty) \to \mathbb{R}$$

is defined to be the inverse of the function

$$f_b : \mathbb{R} \to (0, \infty), \quad f_b(x) = b^x = \exp(x \ln b).$$

In particular, $\log_e x = \ln x$.




**Remark.** The above definition may be stated as

$$y = b^x \quad \Leftrightarrow \quad x = \log_b y.$$

The following theorem demonstrates that all logarithm functions are just scaled versions of the natural logarithm function ...

**Theorem.** Suppose that $b$ is a positive real number with $b \neq 1$. Then

$$\log_b x = \frac{\ln x}{\ln b}$$

for all $x > 0$.

**Proof.** Since

$$x = b^{\log_b x} = \exp(\log_b x \ln b),$$

we conclude that

$$\ln x = \log_b x \ln b.$$

... and, hence, they share all the properties of $\ln$ such as

$$\frac{d}{dx} \log_b x = \frac{d}{dx} \left( \frac{\ln x}{\ln b} \right) = \frac{1}{x \ln b}$$

or

$$\log_b(x^y) = y \log_b x.$$

**Remark.** The above theorem implies that the class of logarithm functions represents the general solution of the functional equation (9.1).

## 9.5 Integration and the $\ln$ function

Since

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

for $x < 0$, the function $\ln(-x)$ is also an antiderivative of $1/x$. Thus,

$$\int \frac{1}{x}\,dx = \ln|x| + C$$

provided that $x$ is restricted to an interval which does not contain $0$.

Moreover, if a function $f$ is positive and differentiable then

$$\frac{d}{dx}\big(\ln f(x)\big) = \frac{f'(x)}{f(x)}$$

so that

$$\int \frac{f'(x)}{f(x)}\,dx = \ln f(x) + C.$$

This may be generalised to

$$\int \frac{f'(x)}{f(x)}\,dx = \ln|f(x)| + C$$

provided that $f$ is differentiable and does not vanish on the interval of integration.

**Example.** On any interval not including zeros of $\cos x$, we have

$$\int \tan x\,dx = -\int \frac{-\sin x}{\cos x}\,dx = -\ln|\cos x| + C.$$

**Example.** Find the indefinite integral

$$\int \frac{1}{2\sec x + \tan x}\,dx.$$

## 9.6 Logarithmic differentiation

Logarithms are powerful in that they 'transform' powers into products, products into sums and quotients into differences.

**Example.** Find the derivative of

$$y = \left(\frac{(3x^2+4)(x+2)}{x^3+5x}\right)^{3/5}.$$

The idea is to take $\ln$ of both sides of the equation to obtain

$$\begin{aligned}\ln y &= \frac{3}{5}\ln\left(\frac{(3x^2+4)(x+2)}{x^3+5x}\right)\\ &= \frac{3}{5}\big(\ln(3x^2+4)+\ln(x+2)-\ln(x^3+5x)\big).\end{aligned}$$

Differentiating both sides with respect to $x$ is relatively easy and leads to

$$\frac{1}{y}\frac{dy}{dx} = \frac{3}{5}\left(\frac{6x}{3x^2+4}+\frac{1}{x+2}-\frac{3x^2+5}{x^3+5x}\right).$$

Hence, we obtain

$$\frac{dy}{dx} = \frac{3}{5}\left(\frac{(3x^2+4)(x+2)}{x^3+5x}\right)^{3/5}\left(\frac{6x}{3x^2+4} + \frac{1}{x+2} - \frac{3x^2+5}{x^3+5x}\right).$$

**Remark.** The above procedure is only valid for intervals on which $y > 0$.

**Example.** Consider the function

$$f : (0,\pi) \to \mathbb{R}, \quad f(x) = x^{\sin x}.$$

Determine its derivative.

## 9.7 Indeterminate forms with powers

**Example.** Find

$$L = \lim_{x \to 0^+} x^{\sin x}.$$

Note that the above limit is of the indeterminate form "$0^0$".

Since the logarithm function is continuous, we conclude that

$$\ln L = \ln \left( \lim_{x \to 0^+} x^{\sin x} \right) = \lim_{x \to 0^+} \ln \left( x^{\sin x} \right).$$

Now,

$$\begin{aligned} \lim_{x \to 0^+} \left| \ln \left( x^{\sin x} \right) \right| &= \lim_{x \to 0^+} |\sin x \ln x| \\ &\leq \lim_{x \to 0^+} |x \ln x| \\ &= 0 \end{aligned}$$

so that

$$\ln L = 0 \quad \Rightarrow \quad L = 1.$$

The following example is of the indeterminate form "$1^\infty$".

**Example.** Show that

$$\lim_{x\to\infty}\left(1+\frac{t}{x}\right)^x = e^t,$$

where $t$ is a constant real parameter.

# Chapter 10

# The hyperbolic functions

## 10.1 Hyperbolic sine and cosine functions

**Definition.** The hyperbolic cosine function $\cosh : \mathbb{R} \to \mathbb{R}$ is defined by

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

**Definition.** The hyperbolic sine function $\sinh : \mathbb{R} \to \mathbb{R}$ is defined by

$$\sinh x = \frac{1}{2}(e^x - e^{-x}).$$

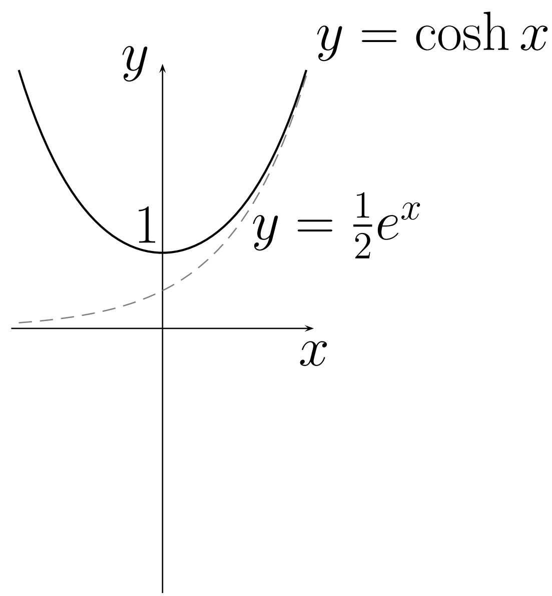
**Note.** $\cosh$ and $\sinh$ are differentiable with

$$\frac{d}{dx}(\sinh x) = \cosh x, \qquad \frac{d}{dx}(\cosh x) = \sinh x$$

so that $\cosh x$ and $\sinh x$ obey the differential equation

$$\frac{d^2y}{dx^2} = y.$$

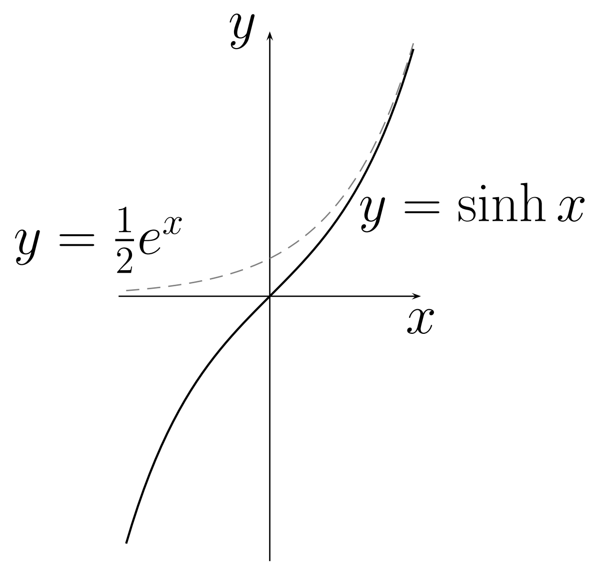
**Properties of the $\cosh$ function.**




- $\cosh$ is an even function.
- $\cosh 0 = 1$.
- $\cosh$ is decreasing on $(-\infty, 0)$, stationary at $0$ and increasing on $(0, \infty)$.
- $\cosh x \geq 1$ for all $x$ in $\mathbb{R}$.
- $\cosh x$ gets arbitrarily close to $\frac{1}{2}e^{\pm x}$ as $x \to \pm\infty$.

**Properties of the $\sinh$ function.**

- $\sinh$ is an odd function.
- $\sinh 0 = 0$.
- $\sinh$ is increasing on $(-\infty, \infty)$ and has a point of inflexion at $0$.
- $\sinh x < 0$ for $x < 0$ and $\sinh x > 0$ for $x > 0$.
- $\sinh x$ gets arbitrarily close to $\pm\frac{1}{2}e^{\pm x}$ as $x \to \pm\infty$.



**Theorem.** The hyperbolic functions are related by

$$\cosh^2 x - \sinh^2 x = 1.$$

**Proof.** By definition,

$$\begin{aligned}\cosh^2 x - \sinh^2 x &= \left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2 \\ &= \frac{1}{4}\left[(e^{2x} + 2 + e^{-2x}) - (e^{2x} - 2 + e^{-2x})\right] \\ &= 1.\end{aligned}$$

**Remark.** The similarity to relations such as

$$\cos^2 x + \sin^2 x = 1, \quad \frac{d}{dx}\cos x = -\sin x, \quad \frac{d}{dx}\sin x = \cos x$$

explains the words cosine and sine in the hyperbolic functions.

The term hyperbolic is motivated in the following manner:

**Example.** Sketch the curve $\gamma(t)$ defined by

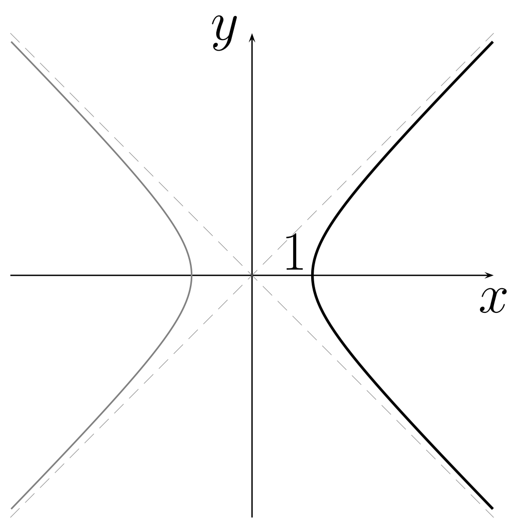
$$\gamma(t) = \big(x(t), y(t)\big) = (\cosh t, \sinh t).$$

Elimination of the parameter $t$ leads to

$$[x(t)]^2 - [y(t)]^2 = \cosh^2 t - \sinh^2 t = 1$$

so that $\gamma$ parametrises the branch of the hyperbola

$$x^2 - y^2 = 1, \qquad x > 0.$$




The other branch of the hyperbola is parametrised by

$$\big(x(t), y(t)\big) = (-\cosh t, \sinh t).$$

## 10.2 Other hyperbolic functions

Other hyperbolic functions are defined in analogy with the trigonometric functions according to

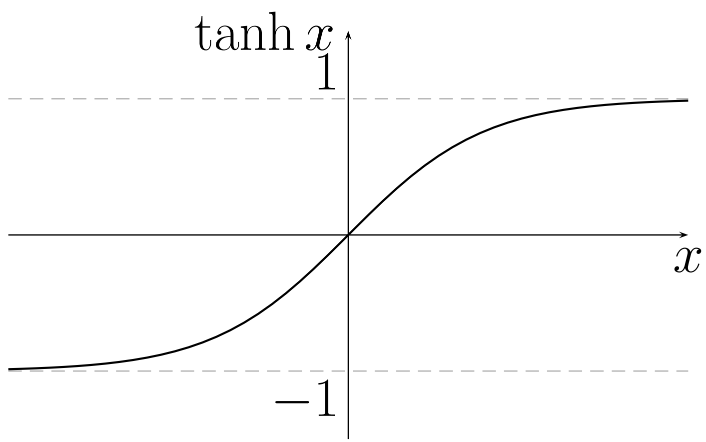
$$\tanh x = \frac{\sinh x}{\cosh x}, \qquad \coth x = \frac{\cosh x}{\sinh x},$$
$$\operatorname{sech} x = \frac{1}{\cosh x}, \qquad \operatorname{cosech} x = \frac{1}{\sinh x}.$$

Recall that

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

**Properties of the $\tanh$ function.**

- $\tanh$ is an odd function.
- $\tanh 0 = 0$.
- $\tanh$ is increasing on $(-\infty, \infty)$ and has a point of inflexion at $0$.
- $\tanh x < 0$ for $x < 0$ and $\tanh x > 0$ for $x > 0$.
- $\lim\limits_{x \to \pm\infty} \tanh x = \pm 1$.
- $\dfrac{d}{dx} \tanh x = \operatorname{sech}^2 x > 0$.



**Proof of the derivative of $\tanh$.** By definition,

$$
\begin{aligned}
\frac{d}{dx}(\tanh x) &= \frac{d}{dx}\left(\frac{\sinh x}{\cosh x}\right) \\
&= \frac{\cosh x \dfrac{d}{dx}\sinh x - \sinh x \dfrac{d}{dx}\cosh x}{\cosh^2 x} \\
&= \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} \\
&= \frac{1}{\cosh^2 x} \\
&= \operatorname{sech}^2 x.
\end{aligned}
$$

**Note.** The slope at the point of inflexion is

$$\left.\frac{d}{dx}\tanh x\right|_{x=0} = \operatorname{sech}^2 0 = 1.$$

## 10.3 Hyperbolic identities

**'Difference of squares' identities.**

$$\cosh^2 x - \sinh^2 x = 1$$
$$1 - \tanh^2 x = \operatorname{sech}^2 x$$
$$\coth^2 x - 1 = \operatorname{cosech}^2 x$$

**'Sum and difference' formulae.**

$$\sinh(x \pm y) = \sinh x \cosh y \pm \cosh x \sinh y$$
$$\cosh(x \pm y) = \cosh x \cosh y \pm \sinh x \sinh y$$
$$\tanh(x \pm y) = \frac{\tanh x \pm \tanh y}{1 \pm \tanh x \tanh y}$$

**'Double-angle' formulae.**

$$\begin{aligned}
\sinh(2x) &= 2\sinh x\cosh x\\
\cosh(2x) &= \cosh^2 x + \sinh^2 x\\
\tanh(2x) &= \frac{2\tanh x}{1+\tanh^2 x}.
\end{aligned}$$

**Exercise.** Prove the first two 'sum and difference' formulae and, hence, derive the third.

## 10.4 Hyperbolic derivatives and integrals

The following derivatives may be readily verified:

$$\frac{d}{dx}\sinh x = \cosh x, \qquad \frac{d}{dx}\cosh x = \sinh x$$

$$\frac{d}{dx}\tanh x = \operatorname{sech}^2 x, \qquad \frac{d}{dx}\coth x = -\operatorname{cosech}^2 x$$

$$\frac{d}{dx}\operatorname{sech} x = -\operatorname{sech} x\ \tanh x$$

$$\frac{d}{dx}\operatorname{cosech} x = -\operatorname{cosech} x\ \coth x.$$

Corresponding indefinite integrals are, for instance,

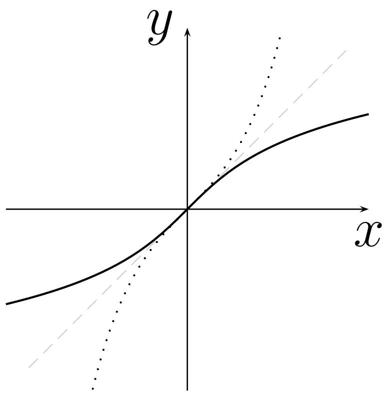
$$\int \sinh x\,dx = \cosh x + C, \qquad \int \operatorname{sech}^2 x\,dx = \tanh x + C.$$

**Example.** Determine the definite integral

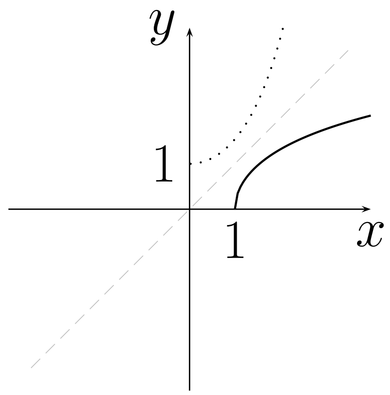
$$I = \int_0^{(\ln 2)^2} \frac{\operatorname{sech}^2 \sqrt{x}}{\sqrt{x}}\, dx.$$

## 10.5 The inverse hyperbolic functions

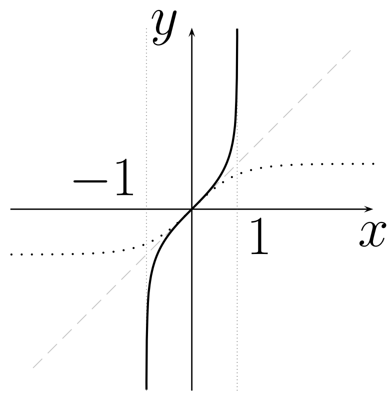
The hyperbolic sine and tan functions are increasing and therefore invertible. In the case of the hyperbolic cosine function, one usually restricts the domain to $[0, \infty)$ to guarantee invertibility.




$\sinh^{-1}$




$\cosh^{-1}$




$\tanh^{-1}$

**Definition.** The functions

- $\cosh^{-1} : [1, \infty) \to [0, \infty)$
- $\sinh^{-1} : \mathbb{R} \to \mathbb{R}$
- $\tanh^{-1} : (-1, 1) \to \mathbb{R}$

are defined to be the inverse functions of the (restricted) hyperbolic functions

- $\cosh : [0, \infty) \to [1, \infty)$
- $\sinh : \mathbb{R} \to \mathbb{R}$
- $\tanh : \mathbb{R} \to (-1, 1)$.

**Example.** Evaluate

$$\sinh\left(\cosh^{-1}\tfrac{4}{3}\right).$$

Evaluation of

$$\sinh^2 t = \cosh^2 t - 1$$

at $t = \cosh^{-1}\frac{4}{3}$ yields

$$\sinh^2\left(\cosh^{-1}\tfrac{4}{3}\right) = \left(\tfrac{4}{3}\right)^2 - 1 = \tfrac{7}{9}.$$

Accordingly,

$$\sinh\left(\cosh^{-1}\tfrac{4}{3}\right) = +\tfrac{\sqrt{7}}{3}$$

since $t > 0$.

**Remark.** Note that

$$\cosh^{-1}(\cosh x) = x$$

for $x \geq 0$ only!

**Example.** Simplify

$$\cosh^{-1}\left(\frac{x}{2} + \frac{1}{2x}\right), \quad x > 0.$$

The relation

$$y = \sinh x = \frac{e^x - e^{-x}}{2}$$

may be written as

$$(e^x)^2 - 2ye^x - 1 = 0$$

so that

$$e^x = \frac{2y + \sqrt{4y^2 + 4}}{2}$$

since $e^x > 0$. Hence,

$$x = \ln\left(y + \sqrt{y^2 + 1}\right).$$

In a similar manner, the remaining identities of the following theorem may be proven.

**Theorem.** The following identities hold:

$$\sinh^{-1} x = \ln\left(x + \sqrt{x^2+1}\right)$$

$$\cosh^{-1} x = \ln\left(x + \sqrt{x^2-1}\right)$$

$$\tanh^{-1} x = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right).$$

The derivatives of these functions are therefore given by

$$\frac{d}{dx}\sinh^{-1} x = \frac{1}{\sqrt{x^2+1}}$$

$$\frac{d}{dx}\cosh^{-1} x = \frac{1}{\sqrt{x^2-1}}$$

$$\frac{d}{dx}\tanh^{-1} x = \frac{1}{1-x^2}.$$

**Example.** Use the inverse function theorem to confirm that

$$\frac{d}{dx}\sinh^{-1} x = \frac{1}{\sqrt{x^2+1}}.$$

If we set $y = \sinh^{-1} x$ then the inverse function theorem implies that

$$\frac{dy}{dx} = \frac{1}{\cosh y} = \frac{1}{\cosh(\sinh^{-1} x)}.$$

On the other hand, since

$$\cosh t = \sqrt{1+\sinh^2 t},$$

we obtain

$$\cosh(\sinh^{-1} x) = \sqrt{1+x^2}$$

so that

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

## 10.6 Integration leading to the inverse hyperbolic functions

From the previous considerations, it follows that

$$\begin{aligned}\int \frac{dx}{\sqrt{x^2+a^2}} &= \sinh^{-1}\frac{x}{a} + C \\ &= \ln\left(x+\sqrt{x^2+a^2}\right) + \tilde{C}, \qquad\qquad a > 0\end{aligned}$$

$$\begin{aligned}\int \frac{dx}{\sqrt{x^2-a^2}} &= \cosh^{-1}\frac{x}{a} + C \\ &= \ln\left(x+\sqrt{x^2-a^2}\right) + \tilde{C}, \qquad\qquad x \geq a > 0\end{aligned}$$

$$\int \frac{dx}{a^2 - x^2} = \begin{cases} \dfrac{1}{a}\tanh^{-1}\dfrac{x}{a} + C, & |x| < a \\ \dfrac{1}{a}\coth^{-1}\dfrac{x}{a} + C, & |x| > a > 0 \end{cases}$$

$$= \frac{1}{2a}\ln\left|\frac{a+x}{a-x}\right| + C, \qquad x^2 \neq a^2.$$

[These formulae are included in the table of standard integrals which is issued at the final examination.]

**Example.** Determine the indefinite integral

$$\int \frac{dx}{\sqrt{x^2 - 2x + 10}}.$$

If we 'complete the square', that is,

$$x^2 - 2x + 10 = (x-1)^2 + 9 = u^2 + 3^2,$$

where $u = x - 1$, then

$$\begin{aligned}\int \frac{dx}{\sqrt{x^2 - 2x + 10}} &= \int \frac{du}{\sqrt{u^2 + 3^2}} \\ &= \sinh^{-1} \frac{u}{3} + C \\ &= \sinh^{-1} \frac{x-1}{3} + C.\end{aligned}$$